Private Politics and Public Interest: NGOs, Corporate Campaigns, and Social Welfare[1]

Jose Miguel Abito[2]
David Besanko[3]
Daniel Diermeier[4]

November 22, 2014

Abstract

We present a theory of corporate campaigns and private regulation, with a focus on welfare implications. An activist induces private regulation through campaigns aimed at harming the firm's reputation. In equilibrium, campaigns have both static and dynamic effects, which have fairly different but complementary effects on social welfare. The canonical case in which an activist campaign increases welfare involves a firm that cares intensely about protecting itself against reputation loss, an activist that is not excessively "passionate," and a campaign aimed at addressing a high-stakes negative externality. Finally, campaigns resulting in private regulation that is essentially redistributive necessarily reduce welfare. JEL L31, D60

[1] We are grateful to Davin Raiha for his helpful comments, and we thank participants at the 14th Annual Strategy and the Environment Conference, Kellogg School of Management, Northwestern University, April 2014 for their useful suggestions and questions.

[2] University of Pennsylvania, Wharton School, Department of Business Economics & Public Policy, Philadelphia, PA 19104, abito@wharton.upenn.edu.

[3] Northwestern University, Kellogg School of Management, Department of Management & Strategy, 2001 Sheridan Rd., Evanston, IL. 60208, d-besanko@kellogg.northwestern.edu (corresponding author).

[4] Northwestern University, Department of Managerial Economics and Decision Sciences (MEDS), Kellogg School of Management, Northwestern University, 2001 Sheridan Rd., Evanston, IL. 60208, d-diermeier@kellogg.northwestern.edu and Canadian Institute for Advanced Research, CIFAR Program on Institutions, Organizations, and Growth, CIFAR, Toronto, ON M5G 1Z8.

1 Introduction

The regulation of economic activity is one of the main arenas of political competition. The impetus for changes to regulatory regimes frequently originates with concerned citizens, often motivated by social or ethical concerns. Traditionally, concerned citizens have used public institutions such as legislatures, executive agencies, and courts to advance their agenda. But in recent decades *private politics* has emerged as a new regulatory mechanism (e.g. Baron 2001, Baron 2003, Baron 2012, Baron and Diermeier 2007, Feddersen and Gilligan 2001, Ingram, Yue, and Rao, 2010, King and McDonnell 2012, King and Pearce 2010).[5] Private politics does not operate through public institutions. Instead, it is characterized by the interaction of private entities such as firms, activists and NGOs.

One of the key outcomes of private politics is *private regulation*, the voluntary adoption of rules that constrain certain company conduct without the involvement of public agents.[6] Private regulation plays a growing role in the regulation of global commerce. For example, fourteen percent of the world's temperate forests and seven percent of global fisheries are governed by private certification systems (Vandenbergh 2013). Issues that have given rise to various forms and levels of private regulation include environmental protection, human rights, discrimination, working conditions, data privacy, product safety, endangered species, and animal welfare.

Private regulation is particularly widely used in cases where public institutions are missing or governance processes are underdeveloped or corrupt. Those include issues that transcend a single government—such as regulation of labor practices within a multinational supply chain or climate change—that require inter-governmental coordination among governments but easily undermined by the free-rider problem, the lack of an adequate framework for international law governing multinational corporations, and the inability of existing multilateral organizations to impose sanctions. Other examples can be found in cases of intra-state conflict or weak regulatory or state capacity. One such example is the attempt to reduce the availability of "conflict diamonds," which are used to fund civil wars in West Africa. In these settings private regulation often takes the form of explicit standards and rules-making processes that resemble public regulation but without any reference to governmental actors. Examples are the Equator Principles or the Sustainable Forestry Initiative.

But private regulation also occurs in countries with well-developed regulatory capacity. In the U.S., for example, more money is spent on private environmental inspections than federal enforcement efforts (Vandenbergh 2013). Activists and observers have argued that this shift towards private solutions may reflect the increasing difficulty to adopt new regulations. Vandenbergh (2013) points out that in the U.S., no major environmental statutes has been enacted since the Clean Air Act Amendments in 1990. Michael Brune, long-time executive director of the Rainforest Action Network (RAN) and currently executive director of the Sierra Club, commented that "Companies were more responsive to public opinion than certain legislatures were. We felt we could create more democracy in the marketplace than in the government." (Baron and Yurday 2004).

Private regulation can occur at the level of individual firms or entire industries. An example of the former

[5] Maxwell, Lyon, and Hackett (2001) call this "self regulation." Vogel (2010) presents the closely related idea of "civil regulation." Vandenberg (2013) uses the term "private governance."

[6] In some cases companies and industries adopt self-regulation to forestall public regulattion (Lyon and Maxwell 2002). Our focus will be on self-regulation to prevent harm from the actions of private agents (activists, NGOs) not public agents (regulators, legislators, courts). For a recent model that studies the interaction between private politics and public regulators see Egorov and Harstad (2012).

is Apple's agreement to improve labor conditions at the Chinese factories of its main supplier Foxconn in the wake of a global media outrage that involved allegations of illegal overtime, inadequate safety conditions and poor workers' housing, as well as a string of reported suicides. Another example is the variety of changes to business practices undertaken by McDonald's over the years—in areas ranging from safety, environmental stewardship, labor conditions at suppliers, animal welfare, the use of antibiotics in food animals, as well as obesity and healthy living—in response to activist campaigns and in anticipation of both public and private politics (Baron 2006). Given the scale of global supply-chains, private regulation involving a single firm can have a significant impact. Walmart, for example, has over 10,000 Chinese suppliers and would be China's eighth largest trading partner if it were a country (Vandenbergh 2013). Walmart uses supply-chain contracting on a large scale from labor conditions to energy and emission requirements, packaging, sustainable fisheries, or conflict-free diamonds, to name just a few. Vandenbergh (2007) studied the supply chain practices of 74 large (and mainly U.S.) firms in eight sectors (e.g., retailing and automobile manufacturing) and found that more than half the companies imposed environmental requirements on their suppliers that exceeded regulatory requirements.

A recent example of industry-wide private regulation was triggered by the collapse of the Rana Plaza factory building in Dhaka, Bangladesh on April 24, 2013. More than 1,000 garment workers, mostly women, died. The Rana Plaza factory was a supplier to many international retailers and brands including Walmart, Disney, The Gap, and H&M. Global media coverage quickly pointed to lax building construction and safety standards, and poor working conditions. After considerable media outrage and activist pressure, companies either withdrew from Bangladesh or agreed to improved safety standards.

Private regulation is based on agreements between private parties, and such agreements sometimes constitute legally binding contracts. For example, in the wake of the Rana Plaza collapse, more than 70 companies, mostly leading European retail brands, established the *Accord on Fire and Building Safety in Bangladesh,* a legally binding agreement to improve working conditions and safety standards among overseas suppliers. In other cases, the agreements are not legally binding. U.S. retailers and brands, for example, largely did not join the *Accord*, but joined in a separate agreement, the *Alliance of Bangladesh Worker Safety,* a voluntary agreement without legal force. To be effective, agreements that are not legally binding must be self-enforcing, e.g. through the credible activist threat to restart a campaign against a company that fails to comply. Whether agreements are designed to be legally binding or not depends on various factors including different liability exposures and legal traditions (e.g. Kaeb 2008). But whether enshrined in formal contracts or voluntary agreements, private regulatory mechanisms in practice "play the standard-setting, implementation, monitoring, enforcement, and adjudication roles traditionally played by public regulatory regimes" (Vandenbergh 2013; p. 105).

Private regulation can have substantial consequences. The 2011 protests over the price of cottage cheese in Israel led to a 20 percent immediate price reduction that was sustained into 2013 (Hendel, Lach, and Spiegel 2013). Similarly, an analysis of Indonesian suppliers in the textile, footwear, apparel sectors showed a 30 percent wage increase for suppliers to major brands such as Nike, Adidas, and Reebok compared to domestic manufacturers in the same sector (Harrison and Scorse 2010).[7] Major brands, especially Nike, had been accused since the 1990s for sweat-shop working conditions at their off-shore suppliers. After a sustained campaign, Nike and other retailers committed to codes of conduct for their suppliers and established monitoring programs. Harrison and Scorse (2010) also show that the rise in wages did not lead to a detectable loss of employment, but was effectively redistributive in nature, analogous to forced profit-sharing.

[7] Indonesian manufacturers in the textile, footwear, apparel sectors also have higher wage rates, about 10-20 percent, compared to other manufacturing sectors (Harrison and Scorse 2010).

More generally, as the scope of private regulation is growing, scholars have raised concerns whether it is an adequate response to the market and governance failures that can arise in a global economy (e.g. Haufler 2001, Bhagwati and Narlikar 2013). For example, in the context of labor conditions, some critics argue that private regulation will undermine the competitiveness of low-cost suppliers and reverse labor force participation for women in traditional societies (e.g. Bhagwati and Narlikar 2013). Others have argued that private regulation may undercut the pressure for public regulation, especially in the context of global environmental concerns such as climate change. Such claims are difficult to assess in the absence of a model. For example, any welfare analysis of private regulation must be clear about its comparison case. First-best comparisons are often unrealistic since the alternative to private politics is often the complete absence of regulation due to lack of governance capacity or a dysfunctional or corrupt political process that is incapable of implementing first-best regulatory solutions.

An additional dimension of the welfare impact of private regulation is the process by which private regulation often comes about. Rather than engaging through traditional public channels, activists have started to target companies directly via corporate campaigns. Such campaigns usually consist of a specific demand, e.g. increase wage rates, accompanied by an implicit or explicit threat to impose harm on the company if management does not comply with the demand (Baron and Diermeier 2007).[8] Though campaigns sometimes involve disruption of operations or consumer boycotts, the most common way that activists seek to harm firms is by affecting their reputations. To be successful, corporate campaigns need to attract media attention, either through coverage in the news media or by leveraging social media such as Facebook and Twitter. Campaigns affect private regulation (and thus social welfare) in potentially two different ways. Anticipating the threat posed by a campaign, a firm might increase its current level of private regulation to defuse the threat (e.g., by creating a buffer of goodwill or by shaping a "narrative" that counters an activist's claims). In addition, because reputation has persistence, if the campaign actually damages the firm's reputation, the firm's future path of private regulation may differ from what it would have been had the campaign not occurred.

This paper presents a theory of corporate campaigns and private regulation, with a focus on their welfare implications. To the extent that firms are highly motivated to protect their reputations, the corporate campaign is a potentially powerful incentive mechanism for inducing firms to take socially beneficial actions. If public regulation is weak or non-existent, campaigns could serve the social interest. However, the harm that a campaign can do to a company's reputation may be costly, e.g., by making it more difficult for the firm whose reputation is impaired to recruit talent or access capital. Moreover, the objectives of activists are not necessarily the objectives of society as a whole; activists may be much more passionate about abating a social harm than a welfare-maximizing planner would be. Thus, a campaign might induce firms to overdirect resources toward addressing problems that activists care passionately about. For these reasons, it is an open question to what extent corporate campaigns would be expected to increase social welfare. Indeed, it seems conceivable that they could even decrease welfare, in which case society would be better off without them, even in a context in which there is no viable public regulation.

Because campaigns can affect reputation and private regulation in both the present and the future, we analyze a dynamic model. Specifically, we study a three-period game with two players: a profit-maximizing firm and an activist whose objective is to achieve abatement of an unregulated social harm. Each period the firm can incur costs to voluntarily abate the social harm (i.e., engage in private regulation) as a way to stochastically improve its reputation.[9] At the same time, the activist can initiate a campaign in each

[8] Most corporate campaigns tend to rely exlcusively on threatened harm rather than promised rewards (e.g. endorsements). See Baron and Diermeier (2007) for details.

[9] Because the game terminates in the third period, the interesting action takes place in the first and second periods. However,

period that could undermine the firm's reputation. The firm's single-period profits are assumed to increase in reputation, but at a diminishing rate, what we call diminishing static returns to reputation. Because reputation enhancement is valuable, the firm would engage in some degree of private regulation even without activist pressure, though its amount may be less than the socially efficient level because it does not internalize the social benefit of abatement.

In equilibrium, the activist will be shown to initiate a campaign only in period 1. The first-period campaign arises because second-period private regulation decreases in reputation—an implication of diminishing static returns to reputation—and if the firm's reputation is tarnished, it will engage in more abatement activity, which is the activist's objective. If there were constant or increasing static returns to reputation, the interests of the activist and the firm would be aligned, and there would be no campaign.[10]

With a campaign by the activist in the first period, the firm not only faces headwinds in its efforts to improve its reputation, it also faces the risk that its reputation could decline between periods 1 and 2. We show that diminishing static returns to reputation "cascade back" to make the firm's (endogenous) value function at the beginning of the second period increasing and concave in reputation, what we call diminishing dynamic returns to reputation. With this endogenously induced risk aversion in reputation, the firm will choose a higher level of private regulation than it would have, had there been no activist. Private regulation in our model thus stems from two conceptually distinct sources: a baseline level aimed at burnishing reputation, and an additional level (arising when the activist launches a campaign) that serves to mitigate the risk of reputation loss created by the activist campaign. In this sense, private regulation in our model reflects two of the drivers identified by Haufler (2001): reputation enhancement and risk management.[11]

In our model, there are two distinct ways an activist campaign could increase the expected present value of social welfare relative to the benchmark in which there is no private regulation (the plausible benchmark in a global setting). First, if private regulation in the initial period is less than the static abatement optimum, the increase in first-period private regulation contributes to higher first-period social welfare. Second, despite the increase in first-period private regulation, the presence of the activist can be shown to stochastically reduce the firm's reputation in the second period. This unfavorable shift in the distribution of the firm's second-period reputation has two offsetting effects. On the one hand, it means that the firm is more likely to undertake higher levels of abatement in period 2 than it would have in the absence of the activist. On the other hand, it reduces the discounted present value of the firm's profit and possibly, too, the welfare of other constituencies, such as consumers or employees, who derive more value when the firm's reputation is higher rather than lower. The balance of these forces is summarized by a term we call the *social return to reputation*, the extent to which second-period social welfare goes up or down when the firm's reputation increases. When the social return to reputation is negative, the stochastic reduction in the firm's reputation contributes to higher social welfare. Depending on parameter values, the presence of the activist could increase or decrease the expected present value of social welfare relative to the no-activist benchmark. The typical case in which the activist plays a constructive role is one in which the social marginal benefit of abatement is high, the marginal cost curve of abatement is flat, diminishing static returns to reputation are pronounced, and the activist is not excessively passionate about the cause of abating the social harm. Indeed, we establish that a sufficiently high social marginal benefit of abatement is a necessary condition for the activist's presence to be socially beneficial, and we show that when private regulation is primarily redistributive, as found in

the third period matters because the prospect of improving profitabillity in the third period motivates the firm to engage in private regulation in period 2.

[10] In fact, with increasing static returns to reputation, the activist would want to assist the firm in building its reputation because that would induce the firm to undertake more abatement activity.

[11] The third driver identified by Haufler (2001) is the more rapid spread of best practices for how to voluntarily regulate.

Harrison and Scorse (2010), it must necessarily decrease social welfare.[12]

The organization of the remainder of the paper is as follows. Section 2 describes the model of competition between the firm and the activist. Section 3 analyzes the equilibrium of the three-period game. Section 4 explores the question of whether the activist's campaign benefits society. Section 5 presents a number of robustness checks and potential extensions of the model. Section 6 summarizes and concludes. Proofs of all propositions are in the Appendix.[13]

2 The model

We consider a model with two actors—a firm and an activist —and two key features. First, the firm's operations contribute to a visible, but unregulated social harm, such as a negative environmental externality or unsafe conditions in an upstream supplier's plant. Second, the extent to which the firm abates the harm—denoted by x— and the intensity of the activist's campaign against the firm—denoted by y—affects the evolution of the firm's reputation over time. Our notion of reputation is consistent with that employed in the sociology and management literatures: it is a social construct reflecting the public's subjective beliefs and attitudes about a company (Bermiss, Zajac, and King 2013; Clardy 2012). It is *not* a posterior belief about hidden information (such as product quality) as in a model with informed and uninformed players.[14]

Throughout the paper, we use the terms "abatement activity" and "private regulation" interchangeably. Moreover, we use those terms broadly: in our model, private regulation can encompass not only voluntary changes in production operations or product characteristics that reduce environmental externalities or safety risks, but also codes of conduct whose primary impact is to transfer surplus from the firm to workers or consumers in the form of higher wages or lower prices.

2.1 Reputational dynamics

Because reputation is fundamentally a dynamic phenomenon, our model is dynamic. Specifically, the interaction between the activist and the firm is modeled as a dynamic stochastic game that plays out over a finite horizon $t = 1, \ldots, T$, and for simplicity we assume $T = 3$, which (as will be seen below) is the smallest number of periods needed to generate interesting behavior by the activist.

The strength of the firm's reputation is r, where $r \in \mathcal{I} \in \{-\infty, \ldots, 0, 1, \ldots, \infty\}$ is integer-valued and is the state variable in the model. We summarize the impact of reputation on the firm's profitability by a reduced-form profit function $\pi(r)$, where $\pi(\cdot)$ is strictly increasing and strictly concave in r. We thus assume a stronger reputation is an asset to the firm, either because it translates into higher demand for the firm's products, gives the firm an edge in recruiting executive talent, allows it to economize on other brand investment activity such as advertising, or makes it easier to find partners for business deals that rely on some degree of trust (Fombrun 1996).

In what follows, we let $\pi_r = \pi(r)$, and $\Delta\pi_r \equiv \pi_{r+1} - \pi_r$. Thus, $\pi_{r-1} < \pi_r < \pi_{r+1}$, while $\Delta\pi_{r-1} > \Delta\pi_r > \Delta\pi_{r+1}$.We refer to this latter condition as diminishing *static* returns to reputation (DSRR), "static" because it pertains to the properties of the *single-period* profit function. When DSRR prevails the single-period

[12] As we discuss below, we use a standard Kaldor-Hicks welfare function of the type typically employed in welfare analyses in the regulation and industrial organization literatures. (consumer plus producer surplus plus/minus external benefits/costs). If instead we used a more general utilitarian welfare function with decreasing marginal utilities, even if private regulation is purely redistributive, the activist campaign could potentially increase social welfare.

[13] Additional derivations and calculations as noted below are in the On-line Appendix to this paper.

[14] See Board and Meyer-Ter-Vehn (2013) for an elegant model in which a firm's reputation is the market's belief about the firm's product quality. A challenge in adapting this type of model to the setting we study is how to develop a compelling mechanism by which an activist could distort the Bayesian learning process.

profit "hit" from reputational loss is more significant than the single-period profit "bump" from reputational improvement. Put another way, to assume DSRR is to assume that the firm is risk averse with respect to its reputation. This strikes us as a plausible assumption. That said, we illustrate below the implications of constant and increasing static returns to reputation.

The profit increment to reputation $\Delta\pi_r$ plays an important role in the analysis. In what follows, it will be useful to parameterize $\Delta\pi_r$ as follows:

$$\Delta\pi_r = \theta\Delta\pi_{r-1}, \tag{1}$$

where $\theta \in (0,1)$ indicates the extent of DSRR: the lower is θ, the more significant is DSRR, and as $\theta \to 1$, DSRR disappears in the limit.

Reputational dynamics are determined by the following process:

$$r_t = r_{t-1} + f_t - a_t,$$

where $f_t \in \{0,1\}$ is a positive shock to the firm's reputation, and $a_t \in \{0,1\}$ is a negative shock to the firm's reputation. Let $p = \Pr(f_t = 1) = P(x_t)$ and $q = \Pr(a_t = 1) = Q(y_t)$,where $P(0) = 0$, $P'(x) > 0$, $Q(0) = 0$, and $Q'(y) > 0$. Thus, more private regulation x in a given period increases the probability of a positive shock in that period, and greater campaign intensity y in that period increases the probability of a negative shock.[15] If a positive and negative shock occur in the same period, they offset each other, and the firm's reputation remains unchanged.

Because there is a one-to-one relationship between x and p and y and q, respectively, it will be analytically convenient to model the firm and activist as choosing p and q directly. Stretching the terminology somewhat, hereafter we refer to p and q as private regulation and campaign intensity, respectively. Letting $h_u(p,q)$, $h_s(p,q)$, and $h_d(p,q)$ denote the probabilities that the firm's reputation increases, stays the same, and decreases from one period to the next, we have $h_u(p,q) = p(1-q)$; $h_d(p,q) = (1-p)q$, and $h_s(p,q) = 1 - h_u(p,q) - h_d(p,q)$.

2.2 The firm and society

The social benefit $w(p) = \overline{w}(P^{-1}(p))$ of private regulation is given by $w(p) = \omega p$, where $\omega > 0$ is the social marginal benefit of abatement effort. This social marginal benefit reflects both the real external benefits of private regulation (e.g., the social benefits of reduced carbon emissions), as well as the net balance of any transfers of surplus among the firm and its stakeholders (e.g., the net change in worker plus producer surplus due to a code of conduct requiring the firm's upstream suppliers to pay higher wages). The total cost to the firm of private regulation—the sum of implementation, compliance, and administrative costs plus the costs of publicizing the firm's efforts so as to boost reputation—is given by $c(p) = \overline{c}(P^{-1}(p))$.[16] For simplicity we assume it is quadratic: $c(p) = \frac{c}{2}p^2$, where $c > 0$. Let $p^S = \min\left\{\frac{\omega}{c}, 1\right\} \leq 1$ denote the *static abatement optimum*, the level of private regulation that maximizes the (static) net benefit of abatement, $\omega p - \frac{c}{2}p^2$.

[15] An alternative formulation would allow for the possibility that current reputational shocks could depend on the entire history of private regulation and campaign intensity, i.e., $\Pr(f_t = 1) = P(x_t, x_{t-1}, \ldots, x_1)$ and $\Pr(a_t = 1) = Q(y_t, y_{t-1}, , \ldots, y_1)$. This formulation, which includes our setup as a special case, adds notational complexity but relatively little additional insight. For this reason, we focus on the case in which changes in reputation from the current to the next period depend only on current period private regulation and campaign intensity.

[16] The cost $c(p)$ does not include transfers of surplus (e.g., lower producer surplus due higher wages) from the firm to stakeholders. These are included, along with their offsetting benefits to recipients, in $w(p)$.

2.3 The activist

An activist campaign tries to draw public attention to the fact that the firm's business operations create a social harm. By choosing a higher level of campaign intensity (e.g., by devoting more resources to organize volunteers or working harder to draw attention to the campaign so that it is more likely to be covered in the media), the activist increases the likelihood that the firm's reputation will suffer a negative shock. The total cost to the activist of a campaign is given by $\gamma(q) = \overline{\gamma}(Q^{-1}(q)) = \frac{\gamma q^2}{2}$, where $\gamma \geq 0$.

We assume that the activist's campaign tactics have no *direct* cost to the firm. These tactics are used only to draw attention to the harm the firm creates in the traditional or social media in the hope of hurting the firm's reputation. We further assume that the activist does not derive a *direct benefit* from hurting the firm's reputation. That is, it is no happier when the firm has a lower reputation than it is when the firm has a higher reputation. Instead, we assume that the activist is *pragmatic*, i.e., its only source of benefit is the level of abatement activity the firm provides each period. We assume the activist maximizes the difference between $\psi w(p)$ and the cost of mounting a campaign $\gamma(q)$, where $\psi \geq 1$ is the activist's "passion" for the "abatement cause." A campaign benefits a pragmatic activist only if, by causing the firm's reputation to decrease, the firm subsequently increases its private regulation.

2.4 Equilibrium conditions

The interaction between the activist and the firm is modeled as a dynamic stochastic game. An equilibrium is described by $\{(p_{rt}, q_{rt}, u_{rt}, v_{rt}) | \, (r,t) \in \mathcal{I} \times \{1,2,3\}\}$, where u_{rt} and v_{rt} are the firm's and activist's values in state r, period t. Letting $\beta_F \in (0,1)$ and $\beta_A \in (0,1)$ denote the discount factors of the firm and activist, respectively, the values u_{rt} and v_{rt} are given by the Bellman equations:

$$u_{rt} = \max_{p_{rt} \in [0,1]} U_{rt}(p_{rt}, q_{rt}) \equiv \pi_r - \frac{c p_{rt}^2}{2} + \beta_F u_{r,t+1} + \beta_F \left\{ \Delta u_{r,t+1} h_u(p_{rt}, q_{rt}) - \Delta u_{r-1,t+1} h_d(p_{rt}, q_{rt}) \right\}; \quad (2)$$

$$v_{rt} = \max_{q_{rt} \in [0,1]} V_{rt}(p_{rt}, q_{rt}) = \psi w(p_{rt}) - \frac{\gamma q_{rt}^2}{2} + \beta_A v_{r,t+1} + \beta_A \left\{ \Delta v_{r,t+1} h_u(p_{rt}, q_{rt}) - \Delta v_{r-1,t+1} h_d(p_{rt}, q_{rt}) \right\}, \quad (3)$$

where $\Delta u_{rt} \equiv u_{r+1,t} - u_{rt}$ and $\Delta v_{rt} \equiv v_{r+1,t} - v_{rt}$, and it is understood that $u_{r4} = v_{r4} = 0$. $U_{rt}(p_{rt}, q_{rt})$ and $V_{rt}(p_{rt}, q_{rt})$ are strictly concave in p_{rt} and q_{rt}, respectively, so the Kuhn-Tucker first-order conditions are necessary and sufficient for a unique global optimum. Using the expression for the derivatives of $h_u(p,q)$ and $h_d(p,q)$ those conditions in state (r,t) can be expressed as:

$$-c p_{rt} + \beta_F \left[(1-q) \Delta u_{r,t+1} + q \Delta u_{r-1,t+1} \right] + \varsigma_{rt}^{p0} - \varsigma_{rt}^{p1} = 0; \left\{ \begin{array}{c} p_{rt} \in [0,1] \\ \varsigma_{rt}^{p0} p_{rt} = \varsigma_{rt}^{p1}(1-p_{rt}) = 0 \\ \varsigma_{rt}^{p0} \geq 0; \varsigma_{rt}^{p1} \geq 0 \end{array} \right\}; \quad (4)$$

$$-\gamma q_{rt} + \beta_A \left[\{ p(-\Delta v_{r,t+1}) + (1-p)(-\Delta v_{r-1,t+1}) \} \right] + \varsigma_{rt}^{q0} - \varsigma_{rt}^{q1} = 0; \left\{ \begin{array}{c} q_{rt} \in [0,1] \\ \varsigma_{rt}^{q0} q_{rt} = \varsigma_{rt}^{q1}(1-q_{rt}) = 0 \\ \varsigma_{rt}^{q0} \geq 0; \varsigma_{rt}^{q1} \geq 0 \end{array} \right\}. \quad (5)$$

where ς_{rt}^{p1}, ς_{rt}^{p1}, ς_{rt}^{q0}, and ς_{rt}^{q1} are Lagrange multipliers. Hereafter, we let $*$ denote equilibrium values.

We maintain the following assumption throughout our analysis:

Assumption 1 *For all $r \in \mathcal{I}$, $c > \beta_F(1+\beta_F)\Delta\pi_r$.*

This assumption says that the marginal cost of private regulation at $p = 1$, exceeds the discounted gain from improving reputation. The impact of the assumption is to eliminate equilibria (with or without the activist) which involve corner solutions for private regulation ($p = 1$).[17]

3 Equilibrium analysis

We proceed in two steps. We first describe what the firm would optimally do in the absence of an activist (the no-activist benchmark). This forms the foundation for analyzing the equilibrium with the activist, which follows.

3.1 No activist benchmark

We denote the no-activist benchmark by the superscript "0." In the Appendix, we derive the firm's private regulation and value function:[18]

$$p^0_{r3} = 0, \tag{6}$$

$$p^0_{r2} = \frac{\beta_F \Delta\pi_r}{c} \in (0,1), \tag{7}$$

$$p^0_{r1} = \frac{\beta_F \Delta u^0_{r2}}{c} \in (0,1). \tag{8}$$

$$u^0_{r3} = \pi_r. \tag{9}$$

$$u^0_{r2} = (1+\beta_F)\pi_r + \frac{\beta_F^2 (\Delta\pi_r)^2}{2c}. \tag{10}$$

$$u^0_{r1} = \pi_r + \beta_F u^0_{r2} - \frac{c}{2}\left(p^0_{r1}\right)^2 + \beta_F p^0_{r1} \Delta u^0_{r2}. \tag{11}$$

That is, there is no private regulation in the terminal period, but positive amounts in the penultimate and initial periods.

Key properties of the no-activist solution (established in the Appendix) are as follows.

- In period 2, the level of private regulation is decreasing in reputation, i.e.,

$$p^0_{r+1,2} < p^0_{r2}. \tag{12}$$

- The firm's value is strictly increasing in reputation in each period, i.e.,

$$\Delta u^0_{rt} > 0,\ r \in \mathcal{I}, t = 1,2,3. \tag{13}$$

- The increment to value is greater the more patient the firm, i.e., $\frac{\partial \Delta u^0_{r2}}{\partial \beta_F} > 0,\ r \in \mathcal{I}$.
- The firm's second period value function exhibits diminishing marginal returns to reputation, i.e., $\Delta u^0_{r2} < \Delta u^0_{r-1,2}$.
- As a consequence of the concavity of the second-period value function, the firm's first period private regulation is strictly decreasing in its reputation level, i.e., $p^0_{r+1,1} < p^0_{r1}$.

[17] For results where Assumption 1 is important, we discuss the implications of it not holding.

[18] Unless explicitly noted, the results in this and subsequently propositions that are expressed in terms of an arbitrary state r should be thought of as applying to all possible states in $\mathcal{I}$.

The concavity of the second-period value function is important for our subsequent analysis. It shows that the *assumed* property of DSRR *endogenously* "cascades backward" to make the firm's value function concave in reputation in period 2. For this reason, we refer to the result that $\Delta u^0_{r-1,2} > \Delta u^0_{r2}$ as diminishing *dynamic* returns to reputation (DDRR). With this property, a strong reputation is a deterrent to private regulation. Because DDRR plays an important role in our analysis with an activist, it is useful to know how it varies with key parameters. From (10) and using (1) we have:

$$\Delta u^0_{r2} = \theta \Delta \pi_{r-1} \left\{ (1+\beta_F) - \frac{\theta\left(1-\theta^2\right)\beta_F^2 \Delta \pi_{r-1}}{2c} \right\}; \tag{14}$$

so

$$\Delta u^0_{r-1,2} - \Delta u^0_{r2} = \Delta \pi_{r-1} \left\{ (1-\theta)(1+\beta_F) - \frac{\beta_F^2 \Delta \pi_{r-1}\left(1-\theta^2\right)^2}{2c} \right\}. \tag{15}$$

Thus

$$\frac{\partial\left[\Delta u^0_{r-1,2} - \Delta u^0_{r2}\right]}{\partial \theta} = \Delta \pi_{r-1} \left\{ -1 - \beta_F \left[1 - \frac{2\theta(1-\theta^2)\beta_F \Delta \pi_{r-1}}{c} \right] \right\} < 0,$$

where the inequality follows because Assumption 1 implies $\frac{\beta_F \Delta \pi_{r-1}}{c} < 1$ and since $\theta < 1$, $2\theta(1-\theta) \leq \frac{1}{2}$. Thus, the more pronounced is DSRR (lower θ), the more pronounced is DDRR. It is also straightforward to show that $\lim_{\theta \to 1}\left[\Delta u^0_{r-1,2} - \Delta u^0_{r2}\right] = 0$, so that DSRR is a necessary condition for DDRR.

Differentiating with respect to the discount factor β_F, we have

$$\frac{\partial\left[\Delta u^0_{r-1,2} - \Delta u^0_{r2}\right]}{\partial \beta_F} = (1-\theta)\Delta \pi_{r-1} \left\{ 1 - \frac{\beta_F \Delta \pi_{r-1}(1+\theta)\left(1-\theta^2\right)}{c} \right\}.$$

Given Assumption 1, a sufficient condition for this to be positive is $(1+\theta)\left(1-\theta^2\right) < 1+\beta_F$, which holds if the firm is sufficiently patient or sufficiently risk averse.[19] Thus, for sufficiently high discount factors and/or high degrees of risk aversion, DDRR becomes more pronounced the more patient the firm is.

3.2 The equilibrium with an activist

We now turn to the characterization of the equilibrium with an activist.

3.2.1 Preliminary results

Let us first establish results that immediately follow from the finite-horizon structure of the game. (Proofs are in the Appendix.) First, $p^*_{r3} = q^*_{r3} = q^*_{r2} = 0$, i.e., in the terminal period the firm does not engage in private regulation (as in the no-activist case), and in both the terminal and penultimate periods, the activist does not mount a campaign. Second, the firm's private regulation in period 2 equals the level in the no-activist case, i.e., $p^*_{r2} = p^0_{r2}$. Third, the equilibrium values for the firm and the activist in periods 2 and 3 are as follows:

$$\begin{aligned} u^*_{r3} &= u^0_{r3} = \pi_r, \\ u^*_{r2} &= u^0_{r2} = (1+\beta_F)\pi_r + \frac{\beta_F^2\left(\Delta \pi_r\right)^2}{2c}, \end{aligned}$$

[19] In fact, because $\max_{\theta \in (0,1)}(1+\theta)\left(1-\theta^2\right) = 32/27$, $\beta_F > 5/27 \simeq 0.19$ is sufficient for this inequality to hold for all $\theta \in [0,1]$.

$$
\begin{aligned}
v^*_{r3} &= 0, \\
v^*_{r2} &= \psi w(p^0_{r2}).
\end{aligned}
$$

The firm's value in periods 2 and 3 corresponds to the values in the no-activist benchmark which, as noted above, is strictly increasing in reputation level. The activist's value, by contrast, is decreasing in the firm's reputation in period 2, i.e., $\Delta v^*_{r2} < 0$.

Because the activist benefits only from a change in the firm's behavior in the subsequent period, and because the firm's behavior (trivially) does not vary in period 3, the activist gains no benefit from a campaign in period 2, which is why $q^*_{r2} = 0$. By contrast, the firm may want to engage in private regulation in period 2 because the potential improvement in reputation would result in higher profits in period 3.

An important implication of $\Delta v^*_{r2} < 0$ is that the interests of the firm and the activist in period 1 are opposed. Because the activist prefers the firm to be in a *lower* reputation state in period 2 to a higher one, it benefits from actions in period 1 that make a *decrease* in the firm's reputation more likely. The activist benefits by weakening the firm's reputation in period 2 because it motivates the firm to undertake more private regulation in that period. A weaker reputation keeps the firm "hungry" to improve its image. This preference by the activist sets the stage for a corporate campaign in period 1.

3.2.2 First-period equilibrium and implications for private regulation dynamics

Let

$$
\Delta w^0_{r2} \equiv w(p^0_{r2}) - w(p^0_{r+1,2}) = \frac{\omega \beta_F}{c} \left[\Delta \pi_r - \Delta \pi_{r+1}\right] = \frac{\omega \beta_F \Delta \pi_{r-1}}{c} \theta(1-\theta) > 0. \tag{16}
$$

$$
\Delta w^0_{r-1,2} \equiv w(p^0_{r-1,2}) - w(p^0_{r2}) = \frac{\omega \beta_F}{c} \left[\Delta \pi_{r-1} - \Delta \pi_r\right] = \frac{\omega \beta_F \Delta \pi_{r-1}}{c} (1-\theta) > \Delta w^0_{r2} > 0. \tag{17}
$$

From the characterization of v^*_{r2} above, $\Delta v^*_{i2} = -\psi \Delta w^0_{i2} < 0$ for $i = r-1, r$. Thus, $-\Delta v^*_{i2}$ and Δu^0_{i2} are positive for $i = r-1, r$, and the Kuhn-Tucker conditions in (4)-(5) imply that $\varsigma^{p0}_{r1} = \varsigma^{q0}_{r1} = 0$ and $p_{r1} > 0$ and $q_{r1} > 0$.[20] Moreover, because Assumption 1 implies $\frac{\beta_F \Delta u^0_{r2}}{c} < \frac{\beta_F \Delta u^0_{r-1,2}}{c} < 1$, we have $p_{r1} < 1$. Thus, the only possibility for a corner solution is for q to equal 1. From the Kuhn-Tucker conditions, we can therefore derive reaction functions for the firm and activist given by:

$$
p^{\mathcal{R}}_{r1}(q) = \frac{\beta_F}{c} \left[\Delta u^0_{r2} + \left(\Delta u^0_{r-1,2} - \Delta u^0_{r2}\right) q\right]. \tag{18}
$$

$$
q^{\mathcal{R}}_{r1}(p) = \min \left\{ \frac{\beta_A \psi}{\gamma} \left[\Delta w^0_{r-1,2} - \left(\Delta w^0_{r-1,2} - \Delta w^0_{r2}\right) p\right], 1 \right\}. \tag{19}
$$

Because $\Delta u^0_{r-1,2} > \Delta u^0_{r2}$, the firm's reaction function is increasing in campaign intensity. Moreover, to the extent that $\Delta u^0_{r-1,2}$ is large relative to Δu^0_{r2}, campaign intensity will have a stronger impact on private regulation on the margin. In other words, the activist provides the strongest incentives for the firm to engage in private regulation when DDRR is most significant. As shown above, DDRR tends to be large when the firm is patient and when DSRR is pronounced. Thus, patient firms for which the loss of reputation is far more consequential than gains to reputation tend to be most responsive to an activist campaign.

Because $\Delta w^0_{r-1,2} > \Delta w^0_{r2}$, $q^{\mathcal{R}}_{r1}(p)$ is non-increasing in p, and it is strictly decreasing when $q^{\mathcal{R}}_{r1}(p) < 1$. Thus, all things being equal, the smaller the firm's anticipated abatement activity, the greater the intensity

[20] To see why, suppose to the contrary that $p_{r1} = 0$. The complementary slackness conditions in (4) would then imply $\varsigma^{p1}_{r1} = 0$. The first-order condition in (4) would then reduce to $\beta_F \left[(1-q)\Delta u_{r,t+1} + q\Delta u_{r-1,t+1}\right] + \varsigma^{p0}_{r1} = 0$, a contradiction since $\Delta u^0_{r2} > 0$, $\Delta u^0_{r-1,2} > 0$ and $\varsigma^{p0}_{r1} \geq 0$. A similar proof establishes $q_{r1} > 0$.

of the activist's campaign (unless the activist's intensity is at the highest possible level). This is consistent with the empirical evidence in Lenox and Eesley (2009) that environmental activists tend to target firms with higher levels of greenhouse gas emissions.

The first-period equilibrium p_{r1}^*, q_{r1}^* solves the reaction functions simultaneously and is either a corner equilibrium given by $p_{r1}^* = \frac{\beta_F \Delta u_{r-1,2}^0}{c}$, $q_{r1}^* = 1$ or an interior equilibrium, $p_{r1}^* \in (0,1)$, $q_{r1}^* \in (0,1)$, given by:[21]

$$p_{r1}^* = \frac{\frac{\beta_F}{c}\left[\left(1 - \frac{\beta_A \psi \Delta w_{r-1,2}^0}{\gamma}\right)\Delta u_{r2}^0 + \frac{\beta_A \psi \Delta w_{r-1,2}^0}{\gamma}\Delta u_{r-1,2}^0\right]}{1 + \frac{\beta_F \beta_A \psi}{c\gamma}\left[\Delta u_{r-1,2}^0 - \Delta u_{r2}^0\right]\left[\Delta w_{r-1,2}^0 - \Delta w_{r2}^0\right]} \in (p_{r1}^0, p_{r-1,1}^0). \tag{20}$$

$$q_{r1}^* = \frac{\frac{\beta_A \psi}{\gamma}\left[\left(1 - \frac{\beta_F \Delta u_{r2}^0}{c}\right)\Delta w_{r-1,2}^0 + \frac{\beta_F \Delta u_{r2}^0}{c}\Delta w_{r2}^0\right]}{1 + \frac{\beta_F \beta_A \psi}{c\gamma}\left[\Delta u_{r-1,2}^0 - \Delta u_{r2}^0\right]\left[\Delta w_{r-1,2}^0 - \Delta w_{r2}^0\right]}. \tag{21}$$

The corner equilibrium arises if and only if:

$$\frac{\beta_A \psi}{\gamma}\left[\Delta w_{r-1,2}^0\left(1 - \frac{\beta_F \Delta u_{r-1,2}^0}{c}\right) + \Delta w_{r2}^0 \frac{\beta_F \Delta u_{r-1,2}^0}{c}\right] \geq 1. \tag{22}$$

Condition (22) will hold only if the activist is sufficiently patient, passionate, or cost efficient, or if the social benefit ω of private regulation is sufficiently large. Figure 1 shows the case of an interior equilibrium.

We now characterize the impact of the activist on private regulation in the first period and on expected private regulation in the second period.

Proposition 1 *For any initial reputation level r, the presence of the activist (a) increases private regulation in the first period, i.e., $p_{r1}^* > p_{r1}^0$, and (b) increases expected private regulation in the second period, i.e., $E_r^*(p_2) > E_r^0(p_2)$.*

Though the activist pushes private regulation in the same direction in periods 1 and 2, the channels through which these impacts operate are different. The activist increases private regulation *state by state* in period 1, a static effect. In period 2, the activist has no static effect; recall, $p_{r2}^* = p_{r2}^0$, i.e., second-period private regulation is the same in any given state with or without the activist. The increase in expected private regulation in period 2 (which, as we shall discuss, occurs *in spite of* the increase in private regulation in period one) arises because the activist campaign (and the associated change in first-period private regulation) changes the evolution of reputation between periods 1 and 2.

Let us first discuss the intuition for part (a) of Proposition 1. By launching a campaign, the activist makes it more difficult for the firm to improve its reputation between periods 1 and 2. In isolation, exogenous changes in factors that make it harder for the firm to improve its reputation would decrease incentives for private regulation. However, the activist also puts the firm at risk of a reputational loss, and given DDRR the firm gains more by avoiding this downside than it gains by increasing its reputation. Accordingly, the increase in private regulation can be interpreted as a form of risk mitigation in the face of possible reputation loss.

To explain part (b) of Proposition 1, it is useful to examine the expression for $E_r^*(p_2) - E_r^0(p_2)$:

$$E_r^*(p_2) - E_r^0(p_2) = h_{dr}^*\left(p_{r-1,2}^0 - p_{r2}^0\right) - \left(h_{ur}^* - h_{ur}^0\right)\left(p_{r2}^0 - p_{r+1,2}^0\right),$$

[21] The expressions in (20) and (21) were found by solving the reaction functions simultaneously assuming $q < 1$. The derivation of the necessary and sufficent condition for the corner solution in (22) is in the Appendix.

where $h^*_{ur} \equiv h_u(p^*_{r1}, q^*_{r1}) = p^*_{r1}(1-q^*_{r1})$ and $h^0_{ur} \equiv h_u(p^0_{r1}, 0) = p^0_{r1}$ are the probabilities the firm's reputation goes up when there is an activist and when there is not, and $h^*_{dr} \equiv h_d(p^*_{r1}, q^*_{r1}) = (1-p^*_{r1})q^*_{r1}$ is the probability that the firm's reputation declines when there is an activist. (Without an activist, this probability is 0.) Because $p^0_{r-1,2} > p^0_{r2}$ and $p^0_{r2} > p^0_{r+1,2}$, the sign of $E^*_r(p_2) - E^0_r(p_2)$ depends on how the activist shifts the distribution over reputational states in period 2. (Recall that when the initial state is r, the period 2 distribution would be over states $r-1$, r, and $r+1$.) The campaign creates a positive probability of reputation loss, which tends to "worsen" the distribution, and it could potentially interfere with the firm's efforts to improve its reputation, which would also worsen the distribution. However, the increase in private regulation in period 1 might offset this latter effect to such an extent that the probability that the firm's reputation goes up is actually higher when there is an activist than when there is not, i.e., $h^*_{ur} > h^0_{ur}$, a "rebound effect." If so the activist's campaign could conceivably result in a mean-increasing spread of distribution over reputational states, and with less private regulation associated with higher reputational states, this would imply $E^*_r(p_2) < E^0_r(p_2)$.

But in proving part (b) of Proposition 1, we show that the rebound effect does not arise, i.e., $h^*_{ur} < h^0_{ur}$.

The intuition for this as is follows. An increase in reputational risk (higher q) leads the firm to increase its private regulation in period 1 (the risk management motive). The rebound effect will not arise if the rate $\frac{dp^{\mathcal{R}}_{r1}(q)}{dq}$ at which the firm increases this private regulation is not "too large." (Essentially the proof of part (b) of Proposition 1 formalizes "not too large.") The risk management motive is weaker—and thus the rate of increase $\frac{dp^{\mathcal{R}}_{r1}(q)}{dq}$ smaller—to the extent that induced second-period risk aversion, measured by $\Delta u^0_{r-1,2} - \Delta u^0_{r2}$, the concavity of the second-period value function, is smaller. (Note from (18) that $\frac{dp^{\mathcal{R}}_{r1}(q)}{dq}$ is proportional to $\Delta u^0_{r-1,2} - \Delta u^0_{r2}$.) For parameter constellations satisfying Assumption 1, second-period private regulation is bounded away from one, reducing the likelihood that reputation and thus profit will increase between periods two and three. This flattens the second-period value function by muting the effect of the return to reputation, $\Delta\pi_r$, which ultimately limits the steepness of firm's first-period reaction function.[22]

The risk aversion of the firm with respect to its reputation that is implied by DSRR plays a key role in Proposition 1. Indeed, DSRR is the *pragmatic activist's lever*. It is what leads to the opposition of interest between the firm and the activist that is the fuel for an activist campaign, and it is what induces the firm to choose a higher level of private regulation in period 1 than it would have in the absence of the activist. We can make this point even more forcefully by considering the implications of constant or increasing returns to reputation, i.e., $\Delta\pi_r \geq \Delta\pi_{r-1}$ for all $r \in \mathcal{I}$. From (7), it follows that p^0_{r2} is constant in r with constant returns to reputation, and p^0_{r2} is increasing in r with increasing returns to reputation. From (16) and (17), we would have $\Delta w^0_{i2} \leq 0$ for $i = r-1, r$, from which it would follow from (5) that $q^{\mathcal{R}}_{r1}(p) = 0$. Indeed, when returns to reputation are increasing, the interests of the firm and the activist are fully aligned. Not only is there no need for the activist to launch a campaign to harm the firm's reputation, it would be counterproductive. If the activist could somehow help the firm improve its reputation, it would do so. This discussion can be summarized by the following proposition:

[22] This, of course, raises the question of what would happen if Assumption 1 did not hold. It should be noted that Assumption 1 is sufficient to prove part (b) of Proposition 1, but not necessary. However, if the assumption was significantly violated at adjacent reputation levels $r-1$ and r, such that $p^0_{r2} = p^0_{r-1,2} = 1$, then the equilibrium between the firm and the activist would be uninteresting if the initial reputation level was r, as the activist would have nothing to gain by launching a campaign. If, by contrast, 1 was strongly violated reputation level $r-1$ such that $p^0_{r-1,2} = 1$ but held at r, so that $p^0_{r2} < 1$, then we would have $q^*_{r1} > 0$ and $p^*_{r1} \in \left[p^0_{r1}, p^0_{r-1,1}\right]$. The rebound effect might or might not arise in this case, but we would not be able to use the sequence of logic in the proof of part (b) of Proposition 1 to rule it out.

Proposition 2 *When an activist is pragmatic, a necessary and sufficient condition for the activist to launch a campaign in period 1 is if the firm is risk averse in its reputation, i.e., DSRR holds. If the firm was not risk averse in its reputation, a campaign would be counterproductive for a pragmatic activist, and the activist would even prefer to help the firm build its reputation if static returns to reputation were increasing.*

It is conceivable that the relationship between π_r and r is s-shaped, with an initial region of increasing marginal returns followed by a region of diminishing marginal returns to reputation. A relationship of this type could arise if a firm must attain a certain critical reputation threshold in order to begin to benefit from further enhancements of reputation, followed by an eventual diminishment of the returns to reputation as reputation continues to rise. Proposition 2 implies that a pragmatic activist would avoid targeting a firm that is in the stage of increasing returns to reputation because as the firm's reputation grows over this range, it would increase its private regulation. To the extent that increasing returns to reputation is correlated with the maturity of the firm's business or where it is in the product life cycle, this insight could lead to testable implications for what types of firms are targeted by activists.

3.2.3 Comparative statics

As can be inferred from Figure 1, any parameter change that shifts the activist's reaction function rightward (leftward), but does not shift the firm's reaction function, results in an equilibrium with a higher (lower) level of private regulation and a more (less) intense campaign by the activist in period 1. From (19), it is straightforward to establish that $\frac{\partial q^{\mathcal{R}}_{r1}(p)}{\partial \psi} > 0$, $\frac{\partial q^{\mathcal{R}}_{r1}(p)}{\partial \beta_A} > 0$, $\frac{\partial q^{\mathcal{R}}_{r1}(p)}{\partial \omega} > 0$, and $\frac{\partial q^{\mathcal{R}}_{r1}(p)}{\partial \gamma} < 0$. Further, these parameters do not affect $p^{\mathcal{R}}_{r1}(q)$. It follows that the more passionate the activist (higher ψ), the more patient the activist (higher β_A), the more cost efficient the activist (lower γ), and the higher the social marginal benefit (higher ω), the greater the level of equilibrium private regulation and campaign intensity in period 1.[23]

Figure 1 also implies that a parameter change that simultaneously shifts the firm's and activist's reaction function rightward will increase equilibrium private regulation in the first period but have an ambiguous impact on equilibrium campaign intensity. This logic implies three additional comparative statics results: the more patient the firm (higher β_F), the flatter the marginal cost curve of abatement (lower c), and the weaker the firm's initial reputation, the greater the level of equilibrium private regulation in period 1.[24]

4 Do activist campaigns benefit society?

Are activist campaigns beneficial for society? Because we have assumed that public regulation is either absent or ineffective, in the absence of activist campaign, the only force pushing the firm to engage in abatement activity is reputation enhancement. Our benchmark, then, is the no-activist equilibrium, and we evaluate the extent to which the presence of the activist increases or decreases welfare relative to this equilibrium.

4.1 The impact of the activist on discounted social welfare

We evaluate the impact of the activist using expected discounted social welfare for an arbitrary starting reputation state r in period 1. Though this is not the only possible metric, it is both comprehensive and natural.

[23] These results can also be verified by straightforward (and tedious) differentiation of (20) and (21).
[24] The derivation of these results is presented in the Appendix.

Per-period social welfare equals the social benefit of abatement, ωp, plus the discounted value of firm profit, minus (if relevant) the cost of the activist campaign, $\frac{\gamma q^2}{2}$, plus any *additional* per period surplus σ_r of agents whose welfare varies with the firm's reputation and the firm does not internalize. For example, if consumer demand depended on the firm's reputation, σ_r would incorporate consumer surplus (taking into account the price the firm is able to charge because of its reputation). The increment to this non-internalized reputational spillover, $\Delta\sigma_r \equiv \sigma_{r+1} - \sigma_r$ could conceivably be positive or negative, but the more natural case is $\Delta\sigma_r \geq 0$, i.e., some of the social benefits of the firm's higher reputation are shared with parties that transact with the firm.[25] We let $\Delta\sigma_r = \eta\Delta\pi_r$ where $\eta \geq 0$ measures the degree of non-internalized reputational spillovers.[26] Throughout the rest of the analysis, we refer to σ_r as consumer surplus. To simplify notation without significant loss of insight, we assume that the social discount factor equals the firm's discount factor.

To begin the analysis, we note that because the firm does not internalize the social benefit of abatement or the benefits of reputation building on consumer surplus, the no-activist equilibrium does not maximize social welfare.[27]

Proposition 3 *Let p^F_{rt} be the first-best level of abatement chosen by a social welfare-maximizing planner. For $t = 2, 3$, $p^F_{rt} > p^0_{rt}$ i.e., the planner chooses strictly more abatement than the firm chooses in the absence of the activist. For $t = 1$, if non-internalized reputational spillovers are sufficiently small—specifically, $\eta < \frac{1}{\beta_F\left[p^0_{r2} - p^0_{r+1,2}\right]} - 1$— $p^F_{r1} > p^0_{rt}$, i.e., the planner chooses strictly more abatement in period 1 than arises in the absence of the activist. Otherwise, $p^F_{r1} \leq p^0_{rt}$.*

At first blush, Proposition 3 seems to reflect nothing more than the standard textbook intuition that a firm will underinvest in abatement effort if it does not internalize its social benefits. But the proposition reflects deeper trade-offs since there are circumstances under which the firm's abatement in the absence of the activist might be greater than the first-best level. This is because the dynamically optimal abatement levels for both the firm and the planner depend not only on the static social marginal benefit of abatement but also on the impact of abatement on the future path of reputation that is embedded in the firm's and planner's value functions. If reputational growth is more valuable to the firm than it is socially, it is conceivable that the firm would invest more in abatement in the first period (for reputation development purposes) than a planner would. If $\eta > \frac{1}{\beta_F\left[p^0_{r2} - p^0_{r+1,2}\right]} - 1$, the planner's value function is less elastic with respect to reputation than the firm's, leading to overinvestment in reputation building by the firm.

Proposition 3 opens the door to the possibility that the activist's presence could increase social welfare. However, this is not a "sure thing." For one thing, while the no-activist private regulation is less than the socially efficient level of abatement in periods 2 and 3 (in period 3, there is a classic textbook negative externality problem, which cascades back to period 2), the presence of the activist does not change private regulation in these periods. In period 1, the activist will (except when the non-internalized reputational spillover η is extremely large) push the firm in the right direction, but it does so by creating the risk that the firm's reputation will fall between period 1 and 2, something a benevolent planner would not do.[28] The upshot is that to discern the welfare impact of an activist campaign, we must compare one distorted equilibrium to another distorted equilibrium.

[25] For example, if σ_r is consumer surplus, then $\Delta\sigma_r \geq 0$ implies that a profit-maximizing firm does not capture—e.g., via higher prices—more than the entire benefit consumers enjoy due to the firm's enhanced reputation. In the On-line Appendix, we discuss the conditions on market demand that ensure this is the case for a single-product monopolist.

[26] In the On-line Appendix, we discuss how our results would change if $\eta < 0$.

[27] We characterize the first-best social welfare maximization problem in the On-line Appendix, and we prove Proposition 3 there.

[28] A planner would directly control abatement to maximize welfare and would not (as we show in the analysis of the first-best problem in the Appendix) choose positive levels of q.

4.1.1 Discounted social welfare

The expected discounted present value of social welfare in period 1, given that the firm's reputation is r, is given by

$$W_{r1} = \sum_{i\in\mathcal{I}}\sum_{t=1}^{3}\beta_F^{t-1}\left[\pi_i - \frac{c(p_{it})^2}{2} + \sigma_i + \omega p_{it} - \frac{\gamma(q_{it})^2}{2}\right]\mu_{\,it|r},$$

where $\mu_{\,it|r}$ is the probability that the firm's reputation is i at time t given it was initially r, and $\mu_{\,r1|r} = 1$. The impact of the activist on discounted welfare, $\Delta W_r^{*0} = W_{r1}^* - W_{r1}^0$ can be expressed as

$$\begin{aligned}\Delta W_r^{*0} \;=\; & \sum_{i\in\mathcal{I}}\sum_{t=1}^{3}\beta_F^{t-1}\left\{\left[\pi_i - \frac{c(p_{it}^*)^2}{2} + \sigma_i + \omega p_{it}^* - \frac{\gamma(q_{it}^*)^2}{2}\right] - \left[\pi_i - \frac{c(p_{it}^0)^2}{2} + \sigma_i + \omega p_{it}^0\right]\right\}\mu_{it|r}^* \\ & - \sum_{i\in\mathcal{I}}\sum_{t=1}^{3}\beta_F^{t-1}\left[\pi_i - \frac{c(p_{it}^0)^2}{2} + \sigma_i + \omega p_{it}^0\right]\left(\mu_{it|r}^0 - \mu_{it|r}^*\right), \qquad (23)\end{aligned}$$

where $\left\{\mu_{it|r}^0 | i\in\mathcal{I}, t\in\{1,2,3\}\right\}$ and $\left\{\mu_{it|r}^* | i\in\mathcal{I}, t\in\{1,2,3\}\right\}$ are the probability distributions over states induced by equilibrium play without and with the activist. The impact of the activist has both a static effect on discounted social welfare (the first term in (23)) and a dynamic effect (the second term in (23)). The static effect is how the activist influences private regulation and per-period welfare, *for a fixed reputational*. Because $p_{rt}^* = p_{rt}^0$ for all $r\in\mathcal{I}$, $t\in\{2,3\}$, the static effect operates only in the first period, and because $\mu_{r1|r}^* = 1$, it is given by the difference in the net benefit of abatement in period 1 minus campaign costs, i.e., $\left[\omega p_{r1}^* - \frac{c}{2}\left(p_{r1}^*\right)^2\right] - \left[\omega p_{r1}^0 - \frac{c}{2}\left(p_{r1}^0\right)^2\right] - \frac{\gamma(q_{r1}^*)^2}{2}$. The dynamic effect is the impact of the activist on *how the firm's reputational trajectory unfolds over time.* It is driven by how the activist changes the induced probability distributions over states. Because $\mu_{r1|r}^0 = \mu_{r1|r}^* = 1$ this effect operates on welfare in the second and third periods.

In a model with only three periods, the static effect and the dynamic effect operate in non-overlapping periods. However, in a T-period model, the static effect operates in periods $1, ..., T-2$, and the dynamic effect operates in periods $2, ..., T$. Thus, the two effects are more than just a relabeling of the impact of the activist in different periods. They reflect two conceptually distinct channels through which the presence of the activist can change discounted expected social welfare: by changing welfare statewise (by changing the level of private regulation in each reputation state) and by changing the evolution of states through time (by changing the transition probabilities, which depend on both private regulation and campaign intensity).

Utilizing the characterization of the equilibrium with and without the activist, we can rewrite ΔW_r^{*0} as[29]

$$\Delta W_r^{*0} = \left\{\begin{array}{c}\left[\omega p_{r1}^* - \frac{c}{2}\left(p_{r1}^*\right)^2\right] - \left[\omega p_{r1}^0 - \frac{c}{2}\left(p_{r1}^0\right)^2\right] - \frac{\gamma(q_{r1}^*)^2}{2} \\ + \left[h_{ur}^* - h_{ur}^0\right]\beta_F\Theta_{r2}^0 - h_{dr}^*\beta_F\Theta_{r-1,2}^0,\end{array}\right\}, \qquad (24)$$

where

$$\Theta_{i2}^0 \equiv \Delta u_{i2}^0 + \Delta s_{i2}^0 - \Delta w_{i2}^0,\ i = r, r-1,$$

and

$$s_{r2}^0 \equiv (1+\beta_F)\sigma_r + p_{r2}^0\beta_F\Delta\sigma_r \qquad (25)$$

is the discounted present value of consumer surplus in period 2.

[29] The On-line Appendix shows how we derive (24) from (23). A key insight in the derivation is that conditional on reputation in period 2, the probability distribution over period 3 states is the same with and without and activist. This is because, as established earlier, the activist does not launch a campaign in period 2.

Having established in proving part (b) of Proposition 1 that $h^*_{ur} - h^0_{ur} < 0$, the direction of the dynamic effect depends entirely on the terms Θ^0_{r2} and $\Theta^0_{r-1,2}$. They are the changes in second-period social welfare when the firm's reputation changes from r to $r+1$ and $r-1$ to r, respectively, or what we will call the *social returns to reputation*. The components of Θ^0_{i2}, $i = r, r-1$, are:

- $\Delta u^0_{i2} > 0$ and $\Delta s^0_{i2} > 0$: the increases in the expected discounted present value (over periods 2 and 3) of the firm's profits and consumer surplus, respectively, due to the firm having a stronger reputation in period 2.
- $\Delta w^0_{i2} = \omega\left[p^0_{i2} - p^0_{i+1,2}\right] > 0$: the social *cost* of an improvement in the firm's reputation due to the *lower* level of private regulation the firm undertakes in period 2 when its reputation increases.

Because $\Delta u^0_{i2}, \Delta s^0_{i2}$, and Δw^0_{i2} are each positive, Θ^0_{r2} and $\Theta^0_{r-1,2}$ can either be positive or negative depending on whether the direct benefits to the firm and consumers from a higher reputation exceed or fall short of the social cost that a stronger reputation creates because it induces the firm to cut back its private regulation.

4.1.2 Welfare comparison

Even though the presence of the activist increases private regulation in the first period and expected private regulation in the second period, discounted social welfare could either go up or down, depending on parameter values. (This can be seen in the numerical examples we present below.) Given that, we will identify parameter conditions that tend to make ΔW^{*0}_r positive or negative. To do this, we take a deeper look at the static and dynamic effects discussed above.

The dynamic effect: how the activist's presence affects expected second-period welfare At the core of the dynamic effect are the social returns to reputation which, using (31)[30], (16), and (25), can be expressed as:

$$\Theta^0_{r2} = \theta^{i-(r-1)}\Delta\pi_{r-1}\left\{(1+\beta_F)(1+\eta) - \frac{(\frac{1}{2}+\eta)\beta_F^2\theta^{i-(r-1)}\Delta\pi_{r-1}}{c}\left[1-\theta^2\right] - \frac{\omega\beta_F}{c}\left[1-\theta\right]\right\}, i = r, r-1, \tag{26}$$

It can be shown that if $\Theta^0_{r-1,2} > 0$, then $\Theta^0_{r,2} > 0$, and it can also be shown that if $\Theta^0_{r2} < 0$, then $\Theta^0_{r-1,2} < 0$**.** This implies that one of three possibilities can obtain: $\Theta^0_{r2} < 0$, $\Theta^0_{r-1,2} < 0$; $\Theta^0_{r2} > 0$, $\Theta^0_{r-1,2} > 0$; and $\Theta^0_{r2} > 0$, $\Theta^0_{r-1,2} < 0$. Under the first two conditions, the impact of the dynamic effect on social welfare is unambiguous, so we characterize the parameter conditions under which these cases obtain.

Rearranging (26) for $i = r$, a sufficient condition for $\Theta^0_{r2} < 0$ and $\Theta^0_{r-1,2} < 0$—which implies that the presence of the activist increases second-period expected welfare—is that the social marginal cost of abatement is sufficiently large, i.e.:

$$\omega > \frac{(1+\beta_F)(1+\eta)c}{\beta_F(1-\theta)} - (\frac{1}{2}+\eta)(1+\theta)\beta_F\theta\Delta\pi_{r-1} \equiv \Gamma(c, \beta_F, \eta, \theta, \Delta\pi_{r-1}). \tag{27}$$

This condition makes sense: if the social marginal benefit of abatement is sufficiently large, the social cost Δw^0_{i2} of the reduced abatement due an improvement in a firm's reputation will be large and will dominate the social benefit, $\Delta u^0_{i2} + \Delta s^0_{i2}$, making the social returns to reputation negative. It is straightforward to

[30] This equation is in the Appendix.

show that $\frac{\partial \Gamma}{\partial c} > 0$, $\frac{\partial \Gamma}{\partial \eta} > 0$, $\frac{\partial \Gamma}{\partial \Delta \pi_{r-1}} < 0$, $\frac{\partial \Gamma}{\partial \beta_F} < 0$, and $\frac{\partial \Gamma}{\partial \theta} > 0$.[31] In light of (27), it follows that the presence of the activist is more likely to increase second-period expected welfare: (a) the flatter the marginal abatement cost function (i.e., the lower is c); (b) the smaller the non-internalized reputational spillover (i.e., the smaller is η); (c) the more reputationally risk averse the firm (i.e., the smaller is θ); (d) the larger the return to reputation, $\Delta \pi_{r-1}$, and (e) the more patient the firm (i.e., the higher is β_F).

Analogously, using (26) for $i = r-1$, a sufficient condition for $\Theta^0_{r2} > 0$ and $\Theta^0_{r-1,2} > 0$—which implies that the presence of the activist decreases second-period expected welfare—is that the social marginal benefit of abatement is sufficiently small:

$$\omega < \frac{(1+\beta_F)(1+\eta)c}{\beta_F (1-\theta)} - (\frac{1}{2}+\eta)(1+\theta)\beta_F \Delta \pi_{r-1} \equiv \Lambda(c, \beta_F, \eta, \theta, \Delta \pi_{r-1}). \tag{28}$$

It can be shown that $\frac{\partial \Lambda}{\partial c} > 0$, $\frac{\partial \Lambda}{\partial \eta} > 0$, $\frac{\partial \Lambda}{\partial \Delta \pi_{r-1}} < 0$, $\frac{\partial \Lambda}{\partial \beta_F} < 0$, and $\frac{\partial \Lambda}{\partial \theta} > 0$. In light of (28), the presence of the activist is more likely to decrease second-period expected welfare: (a) the steeper the marginal abatement cost function (i.e., the larger is c); (b) the larger the non-internalized reputational spillover (i.e., the larger is η); (c) the less reputationally risk averse the firm (i.e., the larger is θ); (d) the smaller the return to reputation, $\Delta \pi_{r-1}$, and (e) the less patient the firm (i.e., the lower is β_F).

The static effect: how the activist affects first-period welfare If private regulation without an activist exceeds the static abatement optimum p^S, the presence of the activist reduces the net benefit of abatement in period 1 by exaggerating the over-delivery of abatement effort. Thus, $p^0_{r1} > \frac{\omega}{c}$ is a sufficient condition for a campaign to decrease the net benefit of abatement in period 1 and first-period social welfare as well (because the activist's campaign is costly). Using (8) and (31), we can rewrite this sufficient condition as

$$\omega < \beta_F \theta \Delta \pi_{r-1} \left\{ (1+\beta_F) - \frac{\theta\left(1-\theta^2\right)\beta_F^2 \Delta \pi_{r-1}}{2c} \right\} \equiv \Xi(c, \beta_F, \theta, \Delta \pi_{r-1}). \tag{29}$$

Thus, if the social marginal benefit of abatement ω is small enough, the activist's presence decreases the net benefit of abatement in period 1. It is straightforward to show that $\frac{\partial \Xi}{\partial c} > 0$, $\frac{\partial \Xi}{\partial \Delta \pi_{r-1}} > 0$, and $\frac{\partial \Xi}{\partial \beta_F} > 0$. Given (29), the presence of the activist is more likely to reduce first-period welfare: (a) the steeper the marginal abatement cost function; (b) the greater the return to reputation, and (c) the higher the discount factor. The impact of θ on Ξ is ambiguous.

Analogously, because $p^0_{r1} < p^*_{r1} < p^0_{r-1,1}$, if $p^0_{r-1,1}$ is less than the static abatement optimum p^S, the presence of the activist increases the net benefit of abatement in period 1 by stimulating under-supplied abatement effort. If q^*_{r1} is sufficiently small (which occurs if β_A or ψ is small), the cost of the campaign will be small, and the activist's presence would decease first-period welfare. We can rewrite the condition $p^0_{r-1,1} < p^S$as:

$$\omega > \beta_F \Delta \pi_{r-1} \left\{ (1+\beta_F) - \frac{\left(1-\theta^2\right)\beta_F^2 \Delta \pi_{r-1}}{2c} \right\} \equiv \Upsilon(c, \beta_F, \theta, \Delta \pi_{r-1}), \tag{30}$$

which implies that if ω is large enough, the activist's presence increases the net benefit of abatement in period 1. It is straightforward to show that $\frac{\partial \Upsilon}{\partial c} > 0$, $\frac{\partial \Upsilon}{\partial \Delta \pi r_{-1}} > 0$, $\frac{\partial \Upsilon}{\partial \beta_F} > 0$, and $\frac{\partial \Upsilon}{\partial \theta} > 0$. Thus, the presence of the activist is more likely to increase first-period welfare: (a) the flatter the marginal abatement cost function; (b) the smaller the return to reputation, and (c) the smaller the discount factor.

[31] The results for $\frac{\partial \Gamma}{\partial \eta}$ and $\frac{\partial \Gamma}{\partial \theta}$ are not immediately obvious and are derived in the On-line Appendix to this paper.

Overall welfare comparison Figure 2 catalogs the possible welfare scenarios, depending on the impact of the activist on first-period welfare and expected second-period welfare, the latter depending on the nature of the social return to reputation. Using the figure and the results just presented, Table 1 summarizes how changes in parameters affect the impact of the activist on social welfare. in each period and overall. A "+" ("−") indicates that the parameter change makes it more likely (less likely) that the activist's presence increases welfare:

Parameter change	**First-period welfare***	**Second-period expected welfare**	**Overall welfare**
Larger social marginal benefit of abatement (ω)	+	+	+
Steeper marginal abatement cost curve (c)	+	+	+
Smaller non-internalized reputational spillover (η)	no effect	+	+
More reputationally risk aversion (lower θ)	ambiguous	+	ambiguous
Larger return to reputation ($\Delta\pi_{r-1}$)	−	+	ambiguous
More patient firm (β_F)	−	+	ambiguous

Table 1

*Provided β_A or ψ is sufficiently small.

Table 1 gives us three sufficient conditions for the presence of the activist to increase (decrease) discounted expected welfare:

- A sufficiently large (small) social marginal benefit of abatement;
- A sufficiently flat (steep) marginal cost of abatement curve;
- A sufficiently small (large) reputational spillover to consumers.

The intuition for these first two conditions will be clearer after we discuss Proposition 4. To understand the impact of η,we note from (29) and (30) that changes in reputational spillover have no impact on how the activist affects first-period welfare. However, a smaller reputational spillover lowers the social returns to reputation, making it more likely that they will be negative, thus making it more likely that the activist's presence increases second-period social welfare.

The discussion so far has focused on sufficient conditions for the activist to increase or decrease social welfare. We now offer a necessary condition.

Proposition 4 *The lowest level of the social marginal benefit of abatement ω consistent with $\Delta W_r^{*0} > 0$ is bounded away from 0. Thus, for the activist's presence to increase expected discounted social welfare, the social marginal benefit of abatement must be sufficiently large.*

Expressed equivalently, for $\omega > 0$ but extremely small, the presence of the activist reduces welfare.

This result is not obvious. When $\omega = 0$, there is no activist campaign, and there is no apparent intuition that would suggest that an infinitessimal increase in the social marginal benefit of abatement ω would reduce welfare. Indeed, one might expect that a slight increase in ω from zero could increase social welfare: equilibrium campaign intensity q_{r1}^* would become positive and that would be associated with an increase in private regulation that would increase the *gross* social benefit of abatement, ωp.

But this intuition is incomplete. First, if first-period private regulation without an activist is already distorted above the static abatement optimum p^S—which must necessarily be the case if ω is positive but

sufficiently small[32]—the static effect of the activist (i.e., the increase in private regulation in the first period) will have a first-order negative impact on first-period welfare. Second, when ω is positive but sufficiently small, the social return to regulation must be positive (intuitively because the social cost of the drop in abatement due to an improvement in the firm's reputation is low).[33] Thus, the dynamic effect of the activist's presence is to reduce second-period expected welfare by stochastically shifting the distribution of second-period reputation in an unfavorable way.[34] As a concluding point, we note that this intuition is robust to more general function forms: if the social marginal marginal benefit of abatement is sufficiently low, the firm is likely to over-supply private regulation relative to the static abatement optimum, and the social returns to reputation in the second period are likely to be positive, irrespective of functional forms.[35]

A possible reaction to Proposition 4 is to question whether cases in which an activist campaign increases social welfare are likely to be relatively rare? As a way to gain partial insight into this question, we consider a case in which it might seem unlikely that the activist could increase social welfare: when the activist-specific parameters ψ and β_A are large enough so that condition (22) holds, and $q^*_{r1} = 1$. We refer to this as the case of a *strong activist.* A strong activist will either reduce the firm's reputation from r to $r-1$ or leave it unchanged, and the direct cost of a campaign is as high as it can be. One might expect, then, that a strong activist is potentially unfavorable to social welfare. However, as the next proposition indicates, even a strong activist can increase social welfare under certain circumstances.

Proposition 5 *Suppose the activist is strong (i.e., ψ and β_A are sufficiently large so that $q^*_{r1} \approx 1$), and condition (27) holds so there are negative social returns to reputation. If γ is positive but sufficiently small, the presence of the activist increases social welfare.*

A possible scenario under which the conditions in Proposition 5 might prevail would be where the issue on which the firm's reputation depends has already been well reported in the media, so the additional cost to the activist of drawing public attention to it is low, and the marginal net benefit of even a little more abatement activity by the firm is very high, possibly because very little is currently being done to mitigate the social harm. A campaign launched against a high-profile firm in the wake of a well-publicized accident in which the firm or its suppliers are implicated as having done too little to prevent the accident might fit this set of circumstances.

By contrast, if the activist is strong but there are positive social returns to reputation, then the impact of the activist on social welfare is ambiguous: the strong activist would increase social welfare in the first period but decrease it in the second period. For the strong activist to *decrease* first-period welfare overall discounted welfare in the face of positive social returns to reputation, the firm would need to do significantly more private regulation in the first period than the static abatement optimum. A possible example when this set of circumstances might arise is when the activist is extremely passionate and private regulation takes the form of "greenwashing"—costly activities such relabeling products or PR campaigns aimed at improving the firm's reputation for being environmentally responsible without significantly abating an underlying harm.

[32] In the proof of Proposition 4, this is established by showing that $\Xi(c, \beta_F, \eta, \theta, \Delta\pi_{r-1})|_{c=\beta_F(1+\beta_F)\Delta\pi_{r-1}} > 0$.

[33] In the proof of Proposition 4, this is established by showing that $\Lambda(c, \beta_F, \eta, \theta, \Delta\pi_{r-1})|_{c=\beta_F(1+\beta_F)\Delta\pi_{r-1}} > 0$.

[34] A technical point is in order here. Proposition 4 relies Assumption 1. If Assumption 1 were strongly violated and did not hold for states r and $r-1$, we would have $p^0_{r-1,1} = p^0_{r1} = 1$, in which case the equilibrium would involve no campaign, and the activist trivially has no welfare effect. arise. On the other hand, if Assumption 1 was violated for state $r-1$ but not state r (so that $p^0_{r1} < 1$), then the result in the proposition would continue to hold.

[35] However, a complicating factor is that the rebound effect cannot be ruled out when analyzing the model with more general function forms. If so, the activist's impact on second- period expected welfare does not depend singularly on the sign of the social return to reputation, but rather on the sign of the entire change in second-period welfare, $\left[h^*_{ur} - h^0_{ur}\right]\beta_F\Theta^0_{r2} - h^*_{dr}\beta_F\Theta^0_{r-1,2}$. In this case, the result in Proposition 4 might be weaker in that it could be conditional on whatever parameter conditions in the more general case preclude the rebound effect from arising or at least make it sufficiently weak.

Under these conditions, the social marginal benefit of the firm's effort will be small (so p^S is low), but it is conceivable that the firm will increase the odds of a reputation boost from its efforts.

Finally, consider the distributional effects of the activist's presence. We can rewrite (24) as the sum of the change in the expected present value of the *gross* benefit of abatement, plus the change $\left[u^*_{r1} - u^0_{r1}\right]$ in the present value of the target firm's profits, plus the change $\left[s^*_{r1} - s^0_{r1}\right]$ in the present value of consumer surplus, minus the cost of the campaign:

$$\Delta W^{*0}_r = \omega \left\{ \left[p^*_{r1} - p^0_{r1}\right] + \beta_F \left[E^*_r(p_2) - E^0_r(p_2)\right] \right\} + \left[u^*_{r1} - u^0_{r1}\right] + \left[s^*_{r1} - s^0_{r1}\right] - \frac{\gamma \left(q^*_{r1}\right)^2}{2}.$$

From Proposition 1, the change in the expected present value of the gross benefit of abatement is positive. Further, it is straightforward to show that $u^*_{r1} - u^0_{r1} < 0$ and $s^*_{r1} - s^0_{r1} < 0$. Thus, the activist's presence helps constituencies that benefit from increased abatement at the expense of the target firm and possibly its consumers. When $\Delta W^{*0}_r < 0$, the activist's presence is akin to an imperfect corrective tax whose excess burden exceeds its externality reducing benefit.

Summing up, this section has identified sufficient conditions for the activist's presence to have a positive effect on social welfare—social marginal benefit of abatement sufficiently large, the marginal cost of abatement sufficiently flat, and non-internalized reputational spillovers small. The other parameters $\beta_F, \eta, \theta, \Delta\pi_{r-1}$ have an ambiguous effect on how the activist impacts social welfare. This section also showed that a necessary condition for the activist to have a beneficial impact on social welfare is that ω is bounded away from 0. Thus, the abatement activity the activist wants to stimulate must have a sufficiently large benefit-cost ratio for the activist's campaign to be welfare-improving. Finally, the activist's presence redistributes surplus away from the target firm and possibly its consumers, toward constituencies that benefit from the firm's abatement activity.

Numerical analysis To further illustrate how the activist's presence impacts social welfare, we illustrate numerical calculations over portions of parameter space. Figures 3-6 illustrate combinations of ω and θ for which the presence of the activist increases or decreases discounted social welfare relative to the no-activist equilibrium. Each figure corresponds to a particular combination of $\Delta\pi_{r-1}$ and ψ.[36] The figures suggest the following regularities:

- An activist campaign increases discounted social welfare only if the social marginal benefit of abatement ω is sufficiently large.
- Conditional on ω being large enough, the activist campaign is more likely to increase discounted social welfare the more risk averse the firm.
- A passionate activist shrinks the range of (ω, θ) for which a campaign is socially beneficial.
- Higher returns to reputation $\Delta\pi_{r-1}$ expand the range of (ω, θ) for which a campaign is socially beneficial for a highly passionate activist but not for a less passionate activist.

The numerical analysis is consistent with the results above, but it also adds additional insight. An activist is more likely to increase social welfare the more risk averse is the target firm, provided that the marginal cost of abatement is sufficiently large. This suggests that society should welcome campaigns over high-stakes issues against high-profile companies that care a lot about protecting their reputations. A key

[36] The other parameter values whose values remained fixed in these calculations are: $\beta_F = 0.95$, $\beta_A = 0.95$, $c = 80$, $\gamma = 175$, $\pi_{r-1} = 0$, $\eta = 0.25$, and $\sigma_r = 1$.

caveat, however, is that a highly passionate activist may undertake such an intense campaign that it ends up being counterproductive even when the social marginal benefit of abatement is reasonably high. In such cases, the social benefit of the additional private regulation engendered by the activist is offset by the costs of reputation loss to the firm and consumers, and possibly, too, by the direct costs of the campaign.

To conserve space, we do not show the results of calculations that vary β_F. However, a representative plot in (ω, β_F) space is presented in the On-line Appendix, and it shows that a higher discount factor for the firm modestly expands the range over which $\Delta W_r^{*0} > 0$. And as in the figures above, increases in ω also expand that range.

4.1.3 Private regulation that is redistributive or has negative net benefits

In the analysis so far we have assumed that the social marginal benefit ω is strictly positive. However, it is possible to imagine circumstances under which $\omega \leq 0$. This would occur, for example, if private regulation results in redistributive profit sharing between firms and workers ($\omega = 0$), as Harrison and Scorse (2010) suggest was the case in Indonesia with the codes of conduct adopted by major footwear brands such as Nike and Reebok in the wake of anti-sweatshop campaigns in the 1990s. It would also be the case if private regulation gives rise to a deadweight loss ($\omega < 0$), as could be the case if a code of conduct for upstream suppliers establishing a living wage caused significant job loss relative to any increase in welfare from the redistribution of surplus from firms to employed workers. Our assumption that $\omega > 0$ rules out both of these possibilities, and thus "stack the deck" against finding that activist campaigns could reduce social welfare. As we showed above, activist campaigns *can* reduce social welfare even when $\omega > 0$, indicating that there are additional forces for why an activist campaign might reduce social welfare beyond the direct possibility that private regulation itself may socially undesirable. Still, the case of $\omega \leq 0$ is interesting when viewed through the lens of our model, so it is useful to discuss it here.

To incorporate this case, while at the same time giving a pragmatic activist an actionable motive (which it would not have given our current specification), we assume that there is a wedge $\varepsilon > 0$ between the true social marginal benefit of abatement ω and the marginal benefit of abatement $\widehat{\omega} = \omega + \varepsilon$ perceived by the activist. Thus, if the wedge is large enough, we can have $\widehat{\omega} > 0$, while $\omega \leq 0$. The wedge could represent the fact that the activist has explicitly redistributive motives that would not be captured in the social welfare function or that it is unaware of the invisible costs of private regulation.

The key role of the wedge is to provide a rationale for a pragmatic activist to mount a campaign. (With $\omega \leq 0$ and no other changes to our model, we would have $q_{r1}^* = 0$.) Specifically, under this formulation, the activist's reaction function becomes

$$\widehat{q}_{r1}^{\mathcal{R}}(p) = \min\left\{\frac{\beta_A \psi}{\gamma}\left[\Delta\widehat{w}_{r-1,2}^0 - \left(\Delta\widehat{w}_{r-1,2}^0 - \Delta\widehat{w}_{r2}^0\right)p\right], 1\right\}$$

where

$$\begin{aligned}\Delta\widehat{w}_{r2}^0 &= \Delta w_{r2}^0 + \frac{\varepsilon\beta_F \Delta\pi_{r-1}}{c}\theta(1-\theta)\\ \Delta\widehat{w}_{r-1,2}^0 &= \Delta w_{r-1,2}^0 + \frac{\varepsilon\beta_F \Delta\pi_{r-1}}{c}(1-\theta).\end{aligned}$$

If $\omega \leq 0$, $\Delta w_{r2}^0 \leq 0$ and $\Delta w_{r-1,2}^0 \leq 0$, but if ε is large enough, $\widehat{q}_{r1}^{\mathcal{R}}(p) > 0$, for some p, and we will have an equilibrium in which $q_{r1}^* > 0$. Under these circumstances, all our results pertaining to the equilibrium hold. Moreover, the welfare comparison is exactly as it was before since ε does not enter the social welfare

function. Because ω must be sufficiently large for the activist's presence to enhance welfare, we conclude that campaigns that result in greater amounts of private regulation that are essentially redistributive ($\omega = 0$) or involves direct social costs that outweigh social benefits ($\omega < 0$) will necessarily decrease social welfare.

Proposition 6 *If the activist perceives a marginal benefit of abatement $\widehat{\omega} = \omega + \varepsilon$ that exceeds the true social marginal benefit of abatement ω by a sufficiently large amount so that $q_{r1}^* > 0$, then if $\omega \leq 0$, the activist's campaign decreases social welfare.*

This result provides a welfare benchmark that could be used to think about Harrison and Scorse's (2010) findings of large wage increases for workers of Indonesian suppliers that followed the adoption of codes of conduct by leading U.S. and European footwear brands. It seems plausible that the actions needed to implement the codes of conduct involved at least some incremental administrative costs to the firms involved ($c > 0$). But if the wage increases did not lead to a detectable loss of employment, then as Harrison and Scorse suggest, they probably represented redistributive profit sharing that transferred producer surplus from Indonesian suppliers to Indonesian workers ($\omega = 0$). Our model highlights that while the campaign that led to the codes of conduct arguably resulted in a more just outcome, they probably reduced social welfare. This raises the question of whether, in a world of imperfect governance where public regulation and redistribution are difficult to affect, there are better mechanisms than campaigns to achieve redistributions of surplus that are desired on grounds of justice and fairness.

However, an important qualification is in order. Our model assumes a welfare function that gives equal weight to a dollar of surplus for the firm, consumers, and beneficiaries of abatement. This is the standard Kaldor-Hicks welfare formulation that is widely used in the regulatory economics and industrial organization literatures and in cost-benefit analysis. It essentially uses potential Pareto improvements as the welfare metric. If we were to consider more general welfare formulations—e.g., a utilitarian social welfare function in which the marginal utilities of income differed across parties—then private regulation that affects a pure transfer between the firm and a group of stakeholders would not be welfare neutral, i.e., $\omega \neq 0$. Still, it is straightforward to reinterpret Proportion 6 to take into account social welfare functions that embody different notions of distributive justice, such as utilitarianism. For example, a surplus-neutral transfer from those with lower marginal utilities of income (e.g., a firm's owners) to those with higher marginal utilities of income could increase welfare (e.g., employees of an upstream supplier in a developing country), would increase the utilitarian social welfare function. If the transfer had no deadweight loss, then it would be as if $\omega > 0$, and if ω were large enough, our model would imply that an activist campaign (e.g., for living wages) that induced a greater amount of purely redistributive private regulation would increase social welfare. On the other hand, if the redistribution was associated with a deadweight loss that was significant relative to the differential in marginal utilities in income, then the redistribution could still reduce social welfare even with a utilitarian welfare formulation, and it would be as if $\omega < 0$. In such a case, our model implies that an activist campaign inducing additional private regulation of this kind would reduce social welfare.

4.2 Is passion socially valuable?

We have modeled the activist's passion as the weight ψ it gives to the gross benefits of reducing a social harm. As is evident from the analysis of the first period equilibrium, this weight is isomorphic to the steepness of a "perceived" marginal cost of campaign curve, with higher values of ψ corresponding to a perception by the activist that the marginal cost of increasing campaign intensity is lower for any given level of intensity. As far as we know, the passion of actual activist groups has not been measured, but within given movements, activist groups do differ in the extent to which they rely on confrontational tactics to advance their aims,

and these differences could arguably be mapped into differences in passion and/or the perceived opportunity cost of confrontational activity. For example, U.S. companies tend to view The Nature Conservancy and the Environmental Defense Fund as cooperative partners with whom they can easily work on sustainability projects, while they view RAN, Dogwood Alliance, and Earth First as more confrontational, with the latter two groups especially prone to rely on "name and shame" tactics (Davies 2014).[37]

This raises the question of whether activist passion is a valuable social asset. Figures 3-6 seem to suggest that, broadly speaking, the answer to this question is no. As noted, a comparison of Figures 3 and 4 reveals that all other things equal, the activist's presence is more likely to increase social welfare when it is less passionate ($\psi = 1$) than more ($\psi = 5$). But because these figures only show ranges over which the effect on welfare due to the presence of the activist changes, they do not allow us to ascertain whether, on the margin, increasing the passion of the activist increases the welfare gain from the activist.

To explore this issue, recall that increasing the activist's passion has the effect of shifting the activist's reaction function rightward and positioning the first period equilibrium toward the northeast along the firm's reaction function. Imagine, then, starting at an equilibrium $(p^*_{r1}(1), q^*_{r1}(1))$ corresponding to $\psi = 1$, and then perturbing ψ upward, resulting in a change in the equilibrium $dq^*_{r1} > 0$ and $dp^*_{r1} = \frac{dp^{\mathcal{R}}_{r1}(q)}{dq} dq^*_{r1} = \left[p^0_{r-1,1} - p^0_{r1}\right] dq^*_{r1} > 0$. The social marginal benefit of increasing the activist's passion is given by

$$d\Delta W^{*0}_r = \left[\omega - cp^*_{r1}(1)\right] dp^*_{r1} - \gamma q^*_{r1}(1) dq^*_{r1} + dh^*_{ur}\beta_F \Theta^0_{r2} - dh^*_{dr}\beta_F \Theta^0_{r-1,2},$$

where dh^*_{ur} and dh^*_{dr} are the changes in the probabilities that the firm's reputation increases and decreases, respectively. The former is positive, while the latter is ambiguous.

The sign of $d\Delta W^{*0}_r$ is ambiguous, but there are circumstances under which it is positive. For example, if the social returns to reputation are negative, first period private regulation is less than the static abatement optimum (i.e., $\omega - cp^*_{r1} > 0$), and increased activist passion does not decrease the probability of reduction in reputation ($dh^*_{dr} \geq 0$), then $d\Delta W^{*0}_r > 0$. Figure 7 illustrates that these conditions are not vacuous. It shows the relationship between the change in welfare ΔW^{*0}_r and the activist's passion for four different levels of the social marginal benefit of abatement and for a very risk-averse target ($\theta = 0.1$).[38] If the social marginal benefit of abatement is sufficiently high ($\omega = 150$ or 200), there is a range of passion parameters over which the social marginal benefit of passion is positive. In these cases, society achieves a bigger welfare gain from a campaign led by an activist that overweights the welfare gains from abatement ($\psi > 1$) than by one whose objective function includes the true social marginal benefit of abatement. On the other hand, if the social marginal benefit of abatement is low enough, the activist's presence reduces social welfare, no matter what the activist's passion is, and furthermore, the social marginal benefit of passion is everywhere negative.

Our analysis thus suggests that activist passion can be a valuable social asset. A campaign that targets a firm that benefits significantly from protecting its reputation may generate a larger welfare gain with a more passionate activist, provided that the activist is not too passionate. (Even with $\omega = 200$, beyond a certain point, increases in activist passion reduce the welfare gain.)

5 Robustness to changes in assumptions and model specification

In this section, we discuss the robustness of the insights of the model by considering a number of natural extensions of the simple modeling structure we employ.

[37] Other groups, such as the Sierra Club and the National Audubon Society, fall somewhere in between.

[38] The other parameter values are those used in the figures above: $\beta_F = 0.95$, $\beta_A = 0.95$, $c = 80$, $\gamma = 175$, $\pi_{r-1} = 0$, $\eta = 0.25$, and $\sigma_r = 1$.

5.1 More general functional forms and infinite horizon dynamics

Our model could be analyzed with more general functional forms than the linear-quadratic specification used here. We have done this analysis, and we obtain the same key result we obtain here: the presence of the activist increases first period private regulation. However, we cannot rule out the possibility of the rebound effect as we can with the linear-quadratic formulation. Still, the welfare trade-offs identified in the linear-quadratic specification hold more generally.

In Abito, Besanko, and Diermeier (2012), we analyze an infinite-horizon game between an activist and a firm aimed at understanding how different activist tactics —criticism and confrontation —affect reputational and private regulation dynamics. As in this paper, the activist's presence has both a static effect (changes in private regulation statewise) and a dynamic effect (changes in transition probabilities and thus the path of reputation and private regulation over time). In the infinite-horizon model, as in our three-period model, the activist tries to damage the firm's reputation in order to motivate the firm to rebuild its reputation by increasing its private regulation. Criticism —the analogue of campaign intensity in our three-period model—increases private regulation on the margin, just as we find here. However, confrontation —which can cause a catastrophic collapse in reputation that the firm is powerless to counteract and has no analogue in our three-period model—may work in the opposite direction. This is because a collapse in reputation can reduce the firm's expected returns to private regulation by so much that it more than offsets the risk management motive highlighted by our three-period analysis. The possibility of reputational collapse could be added to our three-period model, and we conjecture that it would operate the same way it operates in the infinite-horizon model. If so, it would increase the set of circumstances under which the activist's presence reduces first-period welfare and decrease the set of conditions under which the activist increases expected second-period welfare, with the overall impact on discounted social welfare being ambiguous. Summing up, then, while the infinite-horizon model allows for richer reputational dynamics, we believe that as long as the activist's opportunities for affecting reputation are modeled in a consistent way, the welfare trade-offs in an infinite horizon model will parallel those in a three-period model.

5.2 Exogenous probability of reputation loss

Firms face reputational crises for reasons other than activist campaigns. This possibility could be captured by adding an exogenous probability of reduction in reputation to the model. Under this specification, the firm has a risk management motive even in the absence of the activist, and thus for any given reputation level, the firm will engage in more private regulation than in the no-activist equilibrium without the exogenous probability of a reduction in reputation. When the activist is added to the mix under this specification, the exogenous probability of a reputation drop tends to make the activist's campaign intensity less than it would otherwise be because the activist can "free ride" on the exogenous probability of a reduction in reputation. This would marginally improve the welfare properties of the equilibrium with the activist because it would, in effect, reduce the cost of the activist's campaign.

5.3 Reputation enhancement based on solving a specific problem or eliminating a specific practice

Our model assumes that abatement activity is valuable on an ongoing basis and the firm's reputation improves over time (stochastically) to the extent that it engages in a stream of private regulation over time. In practice, however, some activist campaigns are aimed at inducing a firm to solve a specific problem or end a specific practice, and the firm's reputational improvement is not a function of how much abatement

activity it engages in but whether it solves the problem. It is not difficult to adapt our model to this case. Let p be the probability that the firm actually solves the problem, conditional on not having solved it in the previous period. In the absence of the activist, if the firm is perceived as having solved the problem, its reputation improves; otherwise, its reputation remains the same. The no-activist equilibrium for this formulation is similar to the no-activist equilibrium in our model, but the second period continuation values would differ to reflect the possibility that the game could end after the first period. Suppose that there is an activist that endeavors to present evidence that the firm has made the problem worse, and suppose there is a probability q that this evidence becomes salient to the public. This introduces three possible states: the firm solves the problem and the activist's evidence does not become salient, which occurs with probability $p(1-q)$, resulting in an increase in the firm's reputation from r to $r+1$; the firm does not solve the problem and the activist's evidence becomes salient, which occurs with probability $(1-p)q$, resulting in a decrease in the firm's reputation from r to $r-1$, and the firm solves the problem and the activist's evidence becomes salient, which occurs with probability $1-p(1-q)-p(1-q)$. Though other assumptions are possible, a plausible assumption is that in the face of this "reputational noise" the firm's reputation remains unchanged. Except for the changes in the continuation values just noted, this model has the same structure as the model we have analyzed. Changing the way in which abatement activity affects reputation does not have a consequential impact on the insights we derive from our model.

5.4 Multiple activists and/or multiple firms

Finally, our model could be extended to include either multiple activists and/or multiple firms. With multiple activists, any one activist group could free ride on the efforts of other activists. This suggests that overall campaign intensity would be less than it would be with a single activist. In circumstances in which a single activist's presence enhances welfare, the welfare gain under multiple activists (if there is one) would be less, and in circumstances in which a single activist's presence reduces welfare, the welfare loss would be less with multiple activists. The fragmentation of activists under circumstances in which campaigns are more likely to be socially costly—low social marginal benefit of abatement, highly passionate activists, and a target for which building or repairing reputation does not matter that much—would be a good thing from society's perspective.

With multiple firms the issues are more subtle. If the activist has a "multi-tasking" cost function, meaning that its total cost depends on sum of the campaign intensities across firms, in any period the activist will allocate its intensity to a single firm. The firm that a pragmatic activist would be target would be the one for which a drop in reputation would elicit the greatest increase in private regulation. That would depend on firm-specific characteristics (e.g., the marginal cost of abatement and the degree of reputational risk aversion), but it would also depend on firms' returns to reputation $\Delta\pi$, which could depend on how changes in reputation change the strategic interactions among firms.

6 Conclusions

This paper presents a three-period model in which private regulation is motivated by the firm's desire to improve its reputation. Reputational concerns are present without an activist, but the activist induces an additional motive for the firm to engage in private regulation: risk management to protect reputation. By launching a campaign in the first period, the activist creates a risk that the firm's reputation might decline between the first and second periods. This has two distinct effects. The static effect—derived from the risk management motive—is that whatever the firm's first period reputation is, it engages in more private

regulation in the first period than it would have in the absence of an activist. The dynamic effect is that the activist changes the firm's reputational trajectory, making it more likely that the firm will have a lower reputation in the second period, which in turn leads it to undertake more private regulation in period 2 than it would have if its reputation had been higher. Comparing state by state, the firm's private regulation in period 2 when there is an activist is identical to what it would be without the activist. (Indeed, the activist does not launch a campaign in period 2 in any state.) The activist's impact on second period outcomes lies in its ability to "alter the course of history" by putting the firm on a reputational path in which it is hungrier to invest in reputation building. Overall, the activist induces an increase in the expected value of private regulation in period 2.

An activist campaign potentially creates social value by inducing more abatement activity. We show that a firm that faces neither public regulation nor pressure from activists will choose less than the first-best level of abatement in periods 2 and 3, and will do so in period 1 if reputational spillovers are sufficiently small. This opens the door to the possibility that the activist's presence, by increasing abatement activity in period 1 and (in expectation) period 2, can enhance social welfare relative to the no-activist equilibrium.

However, the activist campaign is not a costless vehicle for increasing private regulation. Not only is there a direct cost to an activist campaign, but the static and dynamic effects of the activist's campaign themselves have social costs. The static effect can reduce social welfare if, in the absence of the activist, the firm would have chosen a level of abatement exceeding the static abatement optimum—in which case the activist worsens the first-period distortion and reduces the net benefit from abatement in the first period—or if the firm's abatement was less than the static abatement optimum, but the activist-induced increase in private regulation is so large that the net benefit from abatement in the first period goes down.

The dynamic effect has the benefit of making it more likely that the firm ends up in states in period 2 where it is highly motivated to engage in private regulation. But creating the incentive to keep the firm "hungry" has the cost of reducing the firm's profits and consumer surplus. The net impact of the dynamic effect is summarized by the social returns to reputation.

We show that these static and dynamic welfare effects are complementary. Both first period and second period welfare gains are more likely to arise when the social marginal benefit of abatement is high and when the marginal cost curve of abatement is flat. First period and second period welfare losses from the activist campaign are more likely to occur when the social marginal benefit of abatement is low and the marginal cost curve of abatement is steep. Indeed, we show that the activist's presence can increase discounted social welfare only if the social marginal benefit of abatement is bounded away from 0.

With a slight tweak to our model—allowing the activist to be motivated by a perceived social marginal benefit of abatement that exceeds the actual social marginal benefit, a wedge that might represent distributive concerns or a lack of awareness of certain costs of private regulation—we can show an activist will launch a campaign even though the social marginal benefit of abatement is non-positive, and when it does, the campaign must necessarily reduce discounted social welfare. This insight resonates with recent criticisms of activist campaigns by some scholars, such as Bhagwati and Narlikar (2013). Our model suggests that when a campaign leads a firm to adopt a code of conduct that raise wages either in a distributionally neutral way (i.e., the wage increases are essentially a profit-sharing transfer between the firm and the upstream workers) or in a socially costly way (i.e., the wage increases create a "Harberger triangle" by reducing employment opportunities for other workers), the campaign must necessarily reduce discounted social welfare. This reduction in welfare can be seen as the price that society pays for the redistribution that the activist campaign aims to affect.

However, our model goes beyond this point to suggest that even when activists are not blind to the invis-

ible costs of private regulation and care about the true social marginal costs abatements, activist campaigns may still reduce social welfare. In addition to settings in which the social marginal cost of abatement is low and the activist passion is high, campaigns may also be counterproductive when the firm is not especially risk averse in its reputation and when the gain in profit from an increase in reputation is low. In these latter two cases, the additional motivation to invest in private regulation provided by reputation protection is not especially strong. The activist creates a cost (a reduction in the firm's reputation) with little corresponding benefit (increases in private regulation in the first period and little incremental gain in the second period from placing the firm in a position in which it undertakes more private regulation). This suggests the canonical case in which campaigns would serve the public interest: high-stakes externalities for which the social marginal benefit of abatement is big; the firm is likely to be doing little voluntary abatement, but is highly risk averse and thus cares intensely about protecting its current reputation. An example might be a campaign aimed at inducing a high-profile consumer goods firm or retailer to cut back on the use of an input that has been discovered to pose a serious but heretofore unregulated environmental, safety, or health risk. Efforts on the part of activist groups such as the Safer Chemicals, Healthy Family Coalition in the late 2000s and early 2010s to pressure retailers such as Walmart to not sell baby bottles made with Bisphenol A (BPA)—a plastics additive whose health risks hit the U.S. public consciousness in 2008—is a possible candidate to fit this profile. The model also suggests the canonical case in which campaigns would not serve the public interest: settings in which the firm's activities create low-stakes harm; the firm faces highly passionate activists and is not particularly risk averse about its reputation. An example that fits this profile might be campaigns aimed at preventing animal testing.

7 Appendix

Derivation of the no-activist benchmark:

Period 3 is the terminal period, so there is no gain from undertaking private regulation, establishing (6) and (9). And since $\Delta\pi_r > 0$ by assumption, (13) holds for $t = 3$.

Conditions (7) and (10) follow from solving the Kuhn-Tucker conditions in period 2 with $q = 0$. The fact that $p^0_{r2} \in (0, 1)$, is an implication of Assumption 1: $c > \beta_F(1 + \beta_F)\Delta\pi_r > \beta_F\Delta\pi_r$. Condition (12) follows directly from DSRR.

To establish (13) for $t = 2$, note from (10)

$$\begin{aligned}\Delta u^0_{r2} &= u^0_{r+1,2} - u^0_{r2} = \left\{(1+\beta_F)\pi_{r+1} + \frac{\beta_F^2}{2c}\left(\Delta\pi_{r+1}\right)^2\right\} - \left\{(1+\beta_F)\pi_r + \frac{\beta_F^2}{2c}\left(\Delta\pi_r\right)^2\right\} \\ &= (1+\beta_F)\Delta\pi_r + \frac{\beta_F^2}{2c}\left[\left(\Delta\pi_{r+1}\right)^2 - \left(\Delta\pi_r\right)^2\right]. \qquad (31)\end{aligned}$$

Noting that $\left(\Delta\pi_{r+1}\right)^2 - \left(\Delta\pi_r\right)^2 = \left[\Delta\pi_{r+1} - \Delta\pi_r\right]\left[\Delta\pi_{r+1} + \Delta\pi_r\right]$, with some slight algebraic rearrangement, (31) can be written as

$$\Delta u^0_{r2} = (1 + \beta_F - \frac{\beta_F^2\left[\Delta\pi_{r+1} + \Delta\pi_r\right]}{2c})\Delta\pi_r + \frac{\beta_F^2\left[\Delta\pi_{r+1} + \Delta\pi_r\right]}{2c}\Delta\pi_{r+1} > 0,$$

where the inequality follows because, $\frac{\beta_F}{c}\frac{\left[\Delta\pi_{r+1}+\Delta\pi_r\right]}{2} < 1$ from Assumption 1 and thus $1+\beta_F-\frac{\beta_F^2\left[\Delta\pi_{r+1}+\Delta\pi_r\right]}{2c} > 0$.

Conditions (8) and (11) also follow from solving the Kuhn-Tucker conditions in period 1 with $q = 0$. The

fact that $p^0_{r1} \in (0,1)$ follows because from (31)

$$\begin{aligned}\frac{\beta_F \Delta u^0_{r2}}{c} &= \frac{\beta_F(1+\beta_F)\Delta\pi_r}{c} + \frac{\beta_F^3}{2c^2}\left[(\Delta\pi_{r+1})^2 - (\Delta\pi_r)^2\right] \\ &< \frac{\beta_F(1+\beta_F)\Delta\pi_r}{c} < 1,\end{aligned}$$

where the first inequality follows because $(\Delta\pi_{r+1})^2 < (\Delta\pi_r)^2$ by DSRR and the second follows from Assumption 1.

To prove that $\Delta u^0_{r1} > 0$, we note that p^0_{r1} is feasible, though not necessarily optimal in reputation state $r+1$. Thus,

$$u^0_{r+1,1} \geq \pi_{r+1} - \frac{c\left(p^0_{r1}\right)^2}{2} + \beta_F u^0_{r+1,2} + \beta_F \Delta u^0_{r+1,2} p^0_{r1}.$$

Also

$$u^0_{r1} = \pi_r - \frac{c\left(p^0_{r1}\right)^2}{2} + \beta_F u^0_{r2} + \beta_F \Delta u^0_{r2} p^0_{r1}.$$

Thus,

$$\Delta u^0_{r1} = u^0_{r+1,1} - u^0_{r1} \geq \Delta\pi_r + \beta_F \Delta u^0_{r2}\left(1 - p^0_{r1}\right) + \beta_F \Delta u^0_{r+1,2} p^0_{r1}.$$

Having just established $\Delta u^0_{r2} > 0$ and $\Delta u^0_{r+1,2} > 0$, and since $\Delta\pi_r > 0$, it follows that $\Delta u^0_{r1} > 0$, establishing (13) for $t = 1$.

Using (31) and (7), we have $\frac{\partial \Delta u^0_{r2}}{\partial \beta_F} = \Delta\pi_r\left(1 - p^0_{r2}\right) + \frac{\beta_F}{2c}\left(\Delta\pi_{r+1}\right)^2 > 0$.
To prove that the firm's second period value function exhibits diminishing marginal returns to reputation, we use (31) to get

$$\begin{aligned}\Delta u^0_{r-1,2} - \Delta u^0_{r2} &= (1+\beta_F)\left[\Delta\pi_{r-1} - \Delta\pi_r\right] + \frac{\beta_F^2}{2c}\left\{\left[(\Delta\pi_r)^2 - (\Delta\pi_{r-1})^2\right] - \left[(\Delta\pi_{r+1})^2 - (\Delta\pi_r)^2\right]\right\} \\ &= (1+\beta_F)\left[\Delta\pi_{r-1} - \Delta\pi_r\right] + \frac{\beta_F^2}{2c}\left\{\begin{array}{c}\left[\Delta\pi_r - \Delta\pi_{r-1}\right]\left[\Delta\pi_r + \Delta\pi_{r-1}\right] \\ -\left[\Delta\pi_{r+1} - \Delta\pi_r\right]\left[\Delta\pi_{r+1} + \Delta\pi_r\right]\end{array}\right\} \\ &= \left\{\begin{array}{c}\left\{1 + \beta_F\left[1 - \frac{\beta_F\left(\frac{\Delta\pi_r + \Delta\pi_{r-1}}{2}\right)}{c}\right]\right\}\left[\Delta\pi_{r-1} - \Delta\pi_r\right] \\ +\frac{\beta_F^2}{2c}\left[\Delta\pi_r - \Delta\pi_{r+1}\right]\left[\Delta\pi_{r+1} + \Delta\pi_r\right]\end{array}\right\} > 0, \qquad (32)\end{aligned}$$

where the last inequality follows because Assumption 1 implies $\frac{\beta_F \Delta\pi_r}{c} < 1$ and $\frac{\beta_F \Delta\pi_{r-1}}{c} < 1$, making $\left[1 - \frac{\beta_F\left(\frac{\Delta\pi_r + \Delta\pi_{r-1}}{2}\right)}{c}\right]$ positive.[39] The result that $p^0_{r+1,1} < p^0_{r1}$ follows immediately from (8).

Let's now establish that we have diminishing private regulation over time. Because $p^0_{r3} = 0$, it suffices to show that $p^0_{r1} \geq p^0_{r2}$. From (7) and (8) that inequality turns on the comparison between Δu^0_{r2} and $\Delta\pi_r$. From (31)

$$\Delta u^0_{r2} = \Delta\pi_r + \beta_F \Delta\pi_r + \beta_F\left[\Delta\pi_{r+1} - \Delta\pi_r\right]\frac{\beta_F}{c}\frac{\left[\Delta\pi_{r+1} + \Delta\pi_r\right]}{2}.$$

To prove the result, it suffices to prove that $\beta_F \Delta\pi_r + \beta_F\left[\Delta\pi_{r+1} - \Delta\pi_r\right]\frac{\beta_F}{c}\frac{\left[\Delta\pi_{r+1} + \Delta\pi_r\right]}{2} > 0$. This can be rewritten as

$$\left\{1 - \frac{\beta_F}{c}\frac{\left[\Delta\pi_{r+1} + \Delta\pi_r\right]}{2}\right\}\beta_F \Delta\pi_r + \left\{\frac{\beta_F}{c}\frac{\left[\Delta\pi_{r+1} + \Delta\pi_r\right]}{2}\right\}\beta_F \Delta\pi_{r+1}.$$

Now, by Assumption 1, $\frac{\beta_F}{c}\Delta\pi_r < 1$ and $\frac{\beta_F}{c}\Delta\pi_{r+1} < 1$, so $\frac{\beta_F}{c}\frac{\left[\Delta\pi_{r+1} + \Delta\pi_r\right]}{2} < 1$. Thus

$$\left\{1 - \frac{\beta_F}{c}\frac{\left[\Delta\pi_{r+1} + \Delta\pi_r\right]}{2}\right\}\beta_F \Delta\pi_r + \left\{\frac{\beta_F}{c}\frac{\left[\Delta\pi_{r+1} + \Delta\pi_r\right]}{2}\right\}\beta_F \Delta\pi_{r+1} > \beta_F \Delta\pi_{r+1} > 0.$$

[39] Assumption 1 is not needed to prove this result. If we allow for corner solutions on p_{r2}, we can still establish that $\Delta u^0_{r-1,2} - \Delta u^0_{r2} > 0$.

Hence $\Delta u^0_{r2} > \Delta\pi_r$.

Finally, to establish coasting, we note that given $\Delta u^0_{r2} > \Delta\pi_r$, (8) and (7) imply $p^0_{r2} < p^0_{r1}$. Moreover, as established above, $p^0_{r+1,2} < p^0_{r2}$. Thus $p^0_{r+1,2} < p^0_{r1}$.

Proofs of preliminary results for the equilibrium with an activist:

The result that $p^*_{r3} = q^*_{r3}$ follows immediately from the fact that the third period is the terminal period of the game. Thus, $u^*_{r3} = \pi_r$, and (from 3) $v^*_{r3} = 0$. Now, using (3) again, we have $v_{r2} = \max_{q_{rt}\in[0,1]} \psi w(p^*_{r2}) - \frac{\gamma q^2_{r2}}{2}$, which implies $q^*_{r2} = 0$ and $v^*_{r2} = w(p^*_{r2})$. The result that $p^*_{r2} = p^0_{r2}$ follows immediately from $q^*_{r2} = 0$. The characterization of equilibrium values follows immediately from backward induction. The result that $\Delta v^*_{r2} < 0$ arises because $\Delta v^*_{r2} = v^*_{r+1,2} - v^*_{r2} = \psi\left[w(p^0_{r+1,2}) - w(p^0_{r2})\right]$ and because $p^0_{r+1,2} < p^0_{r2}$ as noted earlier in our discussion of the no-activist benchmark.

Necessary and sufficient condition for a corner equilibrium in the first period:

If

$$\frac{\beta_A\psi}{\gamma}\left[\Delta w^0_{r-1,2}\left(1 - \frac{\beta_F\Delta u^0_{r-1,2}}{c}\right) + \Delta w^0_{r2}\frac{\beta_F\Delta u^0_{r-1,2}}{c}\right] \geq 1,$$

then $q^{\mathcal{R}}_{r1}(\frac{\beta_F\Delta u^0_{r-1,2}}{c}) = 1$. From (18), $\frac{\beta_F\Delta u^0_{r-1,2}}{c}$ is the highest possible equilibrium value for p, and thus, $p^*_{r1} \leq \frac{\beta_F\Delta u^0_{r-1,2}}{c}$. Moreover, from (16) and (17), $\Delta w^0_{r-1,2} > \Delta w^0_{r2}$, so the reaction function $q^{\mathcal{R}}_{r1}(p)$ is non-increasing in p. Thus, $q^*_{r1} = q^{\mathcal{R}}_{r1}(p^*_{r1}) \geq q^{\mathcal{R}}_{r1}(\frac{\beta_F\Delta u^0_{r-1,2}}{c}) = 1$, but since $q^{\mathcal{R}}_{r1}(p)$ cannot be greater than 1, we must have $q^*_{r1} = 1$.

Now, if $q^*_{r1} = 1$, then from (18), $p^*_{r1} = \frac{\beta_F\Delta u^0_{r-1,2}}{c}$. Also, since $1 = q^*_{r1} = q^{\mathcal{R}}_{r1}(p^*_{r1})$, it must be the case that $\min\left\{\frac{\beta_A\psi}{\gamma}\left[\Delta w^0_{r-1,2} - \left(\Delta w^0_{r-1,2} - \Delta w^0_{r2}\right)\left(\frac{\beta_F\Delta u^0_{r-1,2}}{c}\right)\right], 1\right\} = 1$, which implies that

$$\frac{\beta_A\psi}{\gamma}\left[\Delta w^0_{r-1,2}\left(1 - \frac{\beta_F\Delta u^0_{r-1,2}}{c}\right) + \Delta w^0_{r2}\frac{\beta_F\Delta u^0_{r-1,2}}{c}\right] \geq 1.$$

■

Proof of Proposition 1:

Part (a)

Condition (18) in conjunction with (8) implies

$$p^*_{r1} = p^0_{r1} + q^*_{r1}\left(p^0_{r-1,1} - p^0_{r1}\right) > p^0_{r1},$$

because from (19) $q^*_{r1} > 0$ and as pointed out in the characterization of the no-activist equilibrium, $p^0_{r-1,1} > p^0_{r1}$.■

Part (b):

Note that the expression for $E^*_r(p_2) - E^0_r(p_2)$ is given by

$$\begin{aligned} E^*_r(p_2) - E^0_r(p_2) &= \left\{h^*_{dr}p^0_{r-1,2} + (1 - h^*_{ur} - h^*_{dr})p^0_{r,2} + h^*_{ur}p^0_{r+1,2}\right\} - \left\{(1 - h^0_{ur})p^0_{r2} + h^0_{ur}p^0_{r+1,2}\right\} \\ &= h^*_{dr}\left(p^0_{r-1,2} - p^0_{r2}\right) - \left(h^*_{ur} - h^0_{ur}\right)(p^0_{r2} - p^0_{r+1,2}). \end{aligned}$$

where $h^*_{ur} \equiv h_u(p^*_{r1}, q^*_{r1}) = p^*_{r1}(1 - q^*_{r1})$ and $h^0_{ur} \equiv h_u(p^0_{r1}, 0) = p^0_{r1}$ are the probabilities the firm's reputation goes up when there is an activist and when there is not, and $h^*_{dr} \equiv h_d(p^*_{r1}, q^*_{r1}) = (1 - p^*_{r1})q^*_{r1}$ is the probability that the firm's reputation declines when there is an activist. (Without an activist, this probability is 0.) Since $p^0_{r2} > p^0_{r+1,2}$ and $p^0_{r-1,2} > p^0_{r2}$, a sufficient condition for $E^*_r(p_2) - E^0_r(p_2) > 0$ is $h^*_{ur} - h^0_{ur} < 0$. We will now prove that $h^*_{ur} - h^0_{ur} < 0$.

First, note that for a corner equilibrium, $q^*_{r1} = 1$, so $h^*_{ur} = 0$.Thus, $h^*_{ur} - h^0_{ur} < 0$, and $E^*_r(p_2) - E^0_r(p_2) > 0$. Now, for an interior equilibrium, a sufficient condition for the activist's presence to increase expected private regulation is the absence of the rebound effect, i.e., $h^*_{ur} - h^0_{ur} < 0$. We first establish that the rebound effect will not arise if and only if $\frac{\Delta u^0_{r2}}{\Delta u^0_{r-1,2}} > \frac{1 - q^*_{r1}}{2 - q^*_{r1}}$. To do so, define a function $\Delta h_{ur}(q) \equiv p^{\mathcal{R}}_{r1}(q)(1 - q) - p^0_{r1}$. This is the difference in the transition probability from r to $r + 1$ when an activist chooses an arbitrary q and the

firm reacts optimally and when there is no activist. Using (18) and (19), we note that:

$$\begin{aligned}
\Delta h_{ur}(0) &= 0. \\
\Delta h_{ur}(q^*_{r1}) &= h^*_{ur} - h^0_{ur}. \\
\frac{d\Delta h_{ur}(q)}{dq} &= \frac{dp^{\mathcal{R}}_{r1}(q)}{dq}(1-q) - p^{\mathcal{R}}_{r1}(q) = \left[p^0_{r-1,1} - p^0_{r1}\right](1-2q) - p^0_{r1}. \\
\frac{d\Delta h_{ur}(0)}{dq} &= \left[p^0_{r-1,1} - 2p^0_{r1}\right]. \\
\frac{d^2\Delta h_{ur}(q)}{dq^2} &= -2\left[p^0_{r-1,1} - p^0_{r1}\right] < 0.
\end{aligned}$$

There are two possibilities for this function. If $\frac{d\Delta h_{ur}(0)}{dq} < 0$, the function takes its maximum value for $q < 0$, and thus $\Delta h_{ur}(q) < 0$ for $q \in [0,1]$, and in particular, $\Delta h_{ur}(q^*_{r1}) = h^*_{ur} - h^0_{ur} < 0$. From above, this case arises if $p^0_{r-1,1} - 2p^0_{r1} < 0$ or equivalently, $\Delta u^0_{r-1,2} - 2\Delta u^0_{r2} < 0$.

If $\frac{d\Delta h_{ur}(0)}{dq} \geq 0$, then $\Delta h_{ur}(q)$ takes on its maximum value at some $q > 0$ and that maximum value will exceed 0. For this case, the rebound effect does not arise only if the larger root q^+_r of $\Delta h_{ur}(q)$ is less than q^*_{r1} (because $\Delta h_{ur}(q)$ is decreasing for $q > q^+_r$, and thus we would have $0 = \Delta h_{ur}(q^+_r) > \Delta h_{ur}(q^*_{r1}) = h^*_{ur} - h^0_{ur}$). Using the expression for $p^{\mathcal{R}}_{r1}(q)$ in (18) we have

$$\Delta h_{ur}(q) = \frac{\beta_F}{c} q\left[\left(\Delta u^0_{r-1,2} - 2\Delta u^0_{r2}\right) - \left(\Delta u^0_{r-1,2} - \Delta u^0_{r2}\right) q\right].$$

The larger root is $\frac{\Delta u^0_{r-1,2} - 2\Delta u^0_{r2}}{\Delta u^0_{r-1,2} - \Delta u^0_{r2}}$.

Pulling these threads together, we have three cases:

$$\begin{aligned}
\Delta u^0_{r-1,2} - 2\Delta u^0_{r2} &< 0 \Rightarrow h^*_{ur} - h^0_{ur} < 0. \\
0 &\leq \Delta u^0_{r-1,2} - 2\Delta u^0_{r2} < \left[\Delta u^0_{r-1,2} - \Delta u^0_{r2}\right] q^*_{r1} \Rightarrow h^*_{ur} - h^0_{ur} < 0. \\
\Delta u^0_{r-1,2} - 2\Delta u^0_{r2} &\geq \left[\Delta u^0_{r-1,2} - \Delta u^0_{r2}\right] q^*_{r1} \Rightarrow h^*_{ur} - h^0_{ur} \geq 0.
\end{aligned}$$

Thus, the necessary and sufficient condition for the rebound effect not to arise is

$$\Delta u^0_{r-1,2} - 2\Delta u^0_{r2} < \left[\Delta u^0_{r-1,2} - \Delta u^0_{r2}\right] q^*_{r1},$$

or

$$\frac{\Delta u^0_{r2}}{\Delta u^0_{r-1,2}} > \frac{1 - q^*_{r1}}{2 - q^*_{r1}}.$$

Note that the right-hand side of this expression is strictly decreasing in q^*_{r1} and thus $\frac{1}{2} > \frac{1-q^*_{r1}}{2-q^*_{r1}}$.

We now show that given Assumption 1, $\frac{\Delta u^0_{r2}}{\Delta u^0_{r-1,2}} > \frac{1}{2}$, which then implies $\frac{\Delta u^0_{r2}}{\Delta u^0_{r-1,2}} > \frac{1-q^*_{r1}}{2-q^*_{r1}}$ whatever q^*_{r1} is. To establish this, note that (14) implies, $\frac{\Delta u^0_{r2}}{\Delta u^0_{r-1,2}} > \frac{1}{2}$ if and only if

$$\begin{aligned}
& 2\theta\Delta\pi_{r-1}\left\{(1+\beta_F) - \frac{\theta(1-\theta^2)\beta^2_F\Delta\pi_{r-1}}{2c}\right\} \\
> \; & \Delta\pi_{r-1}\left\{(1+\beta_F) - \frac{(1-\theta^2)\beta^2_F\Delta\pi_{r-1}}{2c}\right\},
\end{aligned}$$

which is equivalent to

$$c > \frac{(2\theta^2-1)(1-\theta^2)\beta^2_F\Delta\pi_{r-1}}{2(2\theta-1)(1+\beta_F)}.$$

But by Assumption 1, $c > (1+\beta_F)\beta_F\Delta\pi_{r-1} > \beta_F\Delta\pi_{r-1}$, which in turn is greater than $\frac{(2\theta^2-1)(1-\theta^2)\beta^2_F\Delta\pi_{r-1}}{(2\theta-1)(1+\beta_F)}$ since $\frac{(2\theta^2-1)}{(2\theta-1)} < 1$, $1-\theta^2 < 1$, and $\frac{\beta_F}{1+\beta_F} < 1$. Thus $\frac{\Delta u^0_{r2}}{\Delta u^0_{r-1,2}} > \frac{1}{2} \geq \frac{1-q^*_{r1}}{2-q^*_{r1}}$ which, as just established, implies

$h^*_{ur} - h^0_{ur} < 0$, so the rebound effect does not arise. It follows that $E^*_r(p_2) - E^0_r(p_2) > 0$.■

Proof of comparative statics results:

To prove that higher β_F results in greater first-period private regulation, recall that in the discussion of the no-activist benchmark, $\frac{\partial \Delta u^0_{i2}}{\partial \beta_F} > 0$ for $i = r, r-1$, which from (18) implies $\frac{\partial p^{\mathcal{R}}_{r1}(q)}{\partial \beta_F} > 0$. From (16) and (17), $\frac{\partial \Delta w^0_{i2}}{\partial \beta_F} > 0$ for $i = r, r-1$,which implies $\frac{\partial q^{\mathcal{R}}_{r1}(p)}{\partial \beta_F} > 0$. Thus, a higher β_F simultaneously shifts the firm's reaction function rightward and the activist's reaction function rightward, implying that equilibrium private regulation in the first period must increase.

To prove that a flatter marginal abatement cost curve results in greater first-period private regulation, with straightforward algebra it can be shown that $\frac{\partial\left(\frac{\beta_F \Delta u^0_{i2}(c)}{c}\right)}{\partial c} < 0$ for $i = r, r-1$, and so $\frac{\partial p^{\mathcal{R}}_{r1}(q)}{\partial c} < 0$.[40] From (16) and (17), we see that that $\frac{\partial \Delta w^0_{i2}}{\partial c} < 0$ for $i = r, r-1$. Thus, a lower c simultaneously shifts the firm's reaction function rightward and the activist's reaction function rightward, implying that equilibrium private regulation in the first period must increase.

Finally, to show that in the first-period equilibrium, the weaker the firm's reputation, the greater its private regulation, i.e., $p^*_{r-1,1} > p^*_{r1}$, we note that $p^{\mathcal{R}}_{r1}(q) = \frac{\beta_F}{c}\left[(1-q)\Delta u^0_{r2} + q\Delta u^0_{r-1,2}\right]$ and $p^{\mathcal{R}}_{r-1,1}(q) = \frac{\beta_F}{c}\left[(1-q)\Delta u^0_{r-1,2} + q\Delta u^0_{r-2,2}\right]$. From DDRR, $\Delta u^0_{r-2,2} > \Delta u^0_{r-1,2} > \Delta u^0_{r-1,2}$, so $p^{\mathcal{R}}_{r-1,1}(q) > p^{\mathcal{R}}_{r1}(q)$ for any $q \in [0,1]$. Thus, a decrease in r shifts the firm's reaction function rightward. We also note that $q^{\mathcal{R}}_{r1}(p) = \min\left\{\frac{\beta_A \psi}{\gamma}\left[(1-p)\Delta w^0_{r-1,2} + p\Delta w^0_{r2}\right], 1\right\}$ and $q^{\mathcal{R}}_{r-1,1}(p) = \min\left\{\frac{\beta_A \psi}{\gamma}\left[(1-p)\Delta w^0_{r-2,2} + p\Delta w^0_{r-1,2}\right], 1\right\}$. From (16) and (17), $\Delta w^0_{r-1,2} > \Delta w^0_{r2}$, and it is straightforward to show that $\Delta w^0_{r-2,2} > \Delta w^0_{r-1,2}$. Thus, $q^{\mathcal{R}}_{r-1,1}(p) \geq q^{\mathcal{R}}_{r1}(p)$ for any $p \in [0,1]$. Thus, a decrease in r shifts the activist's reaction function rightward or not at all. As a result, $p^*_{r-1,1} > p^*_{r1}$. ■

Proof of Proposition 4:

Because $\omega < \Lambda(c, \beta_F, \eta, \theta, \Delta\pi_{r-1})$ and $\omega < \Xi(c, \beta_F, \theta, \Delta\pi_{r-1})$ imply $\Delta W^{*0}_r < 0$, it follows that $\omega \geq \Lambda(c, \beta_F, \eta, \theta, \Delta\pi_{r-1})$ or $\omega \geq \Xi(c, \beta_F, \theta, \Delta\pi_{r-1})$ is a necessary condition for $\Delta W^{*0}_r \geq 0$. Because both $\Lambda(c, \beta_F, \eta, \theta, \Delta\pi_{r-1})$ and $\Xi(c, \beta_F, \theta, \Delta\pi_{r-1})$ are increasing in c whose smallest possible value consistent with Assumption 1 is $\beta_F(1+\beta_F)\Delta\pi_{r-1}$, it follows that the lowest possible value of ω for the activist to have a positive impact on social welfare must be greater than $\min\{\Lambda(c, \beta_F, \eta, \theta, \Delta\pi_{r-1})|_{c=\beta_F(1+\beta_F)\Delta\pi_{r-1}}, \Xi(c, \beta_F, \theta, \Delta\pi_{r-1})|_{c=\beta_F(1+\beta_F)\Delta\pi_{r-1}}, 0\}$. We establish that this minimum is strictly positive by showing that $\Lambda(c, \beta_F, \eta, \theta, \Delta\pi_{r-1})|_{c=\beta_F(1+\beta_F)\Delta\pi_{r-1}} > 0$ and $\Xi(c, \beta_F, \theta, \Delta\pi_{r-1})|_{c=\beta_F(1+\beta_F)\Delta\pi_{r-1}} > 0$. Substituting $\beta_F(1+\beta_F)\Delta\pi_{r-1}$ for c in (28) and (29), and rearranging terms gives us,

$$\Lambda(c, \beta_F, \eta, \theta, \Delta\pi_{r-1})|_{c=\beta_F(1+\beta_F)\Delta\pi_{r-1}} = \frac{(1+\eta)\Delta\pi_{r-1}}{(1-\theta)}\left\{(1+\beta_F)^2 - \left(\frac{\frac{1}{2}+\eta}{1+\eta}\right)\beta_F(1-\theta^2)\right\} > 0.$$

$$\Xi(c, \beta_F, \theta, \Delta\pi_{r-1})|_{c=\beta_F(1+\beta_F)\Delta\pi_{r-1}} = \frac{\beta_F\theta\Delta\pi_{r-1}}{1+\beta_F}\left\{(1+\beta_F)^2 - \frac{\theta(1-\theta)}{2}\right\} > 0,$$

since $\theta \in (0,1)$, $\beta_F \in (0,1)$, and $1 + \eta \geq 0$.■

Proof of Proposition 5:

When the activist is strong, we have $q^*_{r1} \approx 1$, and from the discussion above, $h^*_{ur} - h^0_{ur} < 0$. Moreover, $h^*_{dr} = (1 - p^*_{r1}) > 0$ because from (18), $p^*_{r1} \in (p^0_{r1}, p^0_{r-1,1}) < 1$. Because (27) holds, we have $\Theta^0_{r2} < 0$ and $\Theta^0_{r-1,2} < 0$. The second line of (24) is thus positive.

Now, note that (27) implies

$$\frac{\omega}{c} > \frac{(1+\beta_F)(1+\eta)}{\beta_F(1-\theta)} - \left(\frac{1}{2}+\eta\right)(1+\theta)\frac{\beta_F\theta\Delta\pi_{r-1}}{c}$$

[40] Using (31),

$$\frac{\partial\left(\frac{\beta_F \Delta u^0_{s2}}{c}\right)}{\partial c} = -\frac{\beta_F}{c^2}\Delta\pi_r - \frac{\beta_F^3}{c^3}(\Delta\pi_{r+1})^2 - \frac{\beta_F^2}{c^2}\left[1 - \frac{\beta_F\Delta\pi_r}{c}\right]\Delta\pi_r < 0$$

because $1 - \frac{\beta_F\Delta\pi_r}{c} < 0$ due to Assumption 1.

which can be rewritten as

$$\frac{\omega}{c} > \frac{(1+\eta)}{(1-\theta)}\left[\frac{1}{\beta_F} + 1 - \frac{\left(\frac{1}{2}+\eta\right)\left(1-\theta^2\right)}{(1+\eta)}\frac{\beta_F \theta \Delta\pi_{r-1}}{c}\right]$$

Since $\frac{\beta_F \theta \Delta\pi_{r-1}}{c} < 1$, the term in brackets is greater than one and so $\frac{\omega}{c} > \frac{(1+\eta)}{(1-\theta)} > 1$. The implication that $\omega > c$ implies that $\omega p - \frac{c}{2}p^2$ is strictly increasing for all $p \in [0,1]$, and because $p^*_{r1} > p^0_{r1}$, it follows that $\omega p^*_{r1} - \frac{c}{2}\left(p^*_{r1}\right)^2 > \omega p^0_{r1} - \frac{c}{2}\left(p^0_{r1}\right)^2$, which establishes that the first line of ΔW^{*0}_r is positive. If γ is positive but sufficiently close to zero, then $\Delta W^{*0}_r > 0$.■

8 References

1. Abito, J.M., D. Besanko, and D. Diermeier (2012) "Corporate Reputational Dynamics, Private Regulation, and Activist Pressure," Working Paper, Northwestern University.

2. Baron, D. P. (2001) "Private Politics, Corporate Social Responsibility, and Integrated Strategy," *Journal of Economics and Management Strategy*, Vol. 10, No. 1, pp. 7–45.

3. Baron, D. P. (2003) "Private Politics," *Journal of Economics and Management Strategy,*Vol. 12, No. 1, 31–66.

4. Baron, D. P. (2009) "A Positive Theory of Moral Management, Social Pressure, and Corporate Social Performance," *Journal of Economics and Management Strategy*,Vol. 18, No. 1, pp. 7-43.

5. Baron, D. P. and E. Yurday (2004) "Strategic Activism: The Rainforest Action Network," Case P-44, Graduate School of Business, Stanford University, Stanford, CA.

6. Baron, D. P. and D. Diermeier (2007) "Strategic Activism and Nonmarket Strategy," *Journal of Economics and Management Strategy*, Vol. 16, No. 3, pp. 599-634.

7. Bhagwati, J., and A. Narlikar (2013) "Signing Up to Safety Laws Does Workers More Harm than Good," *Financial Times*, July 18.

8. Bermiss, Y. S., E. J. Zajac, and B.G. King (2013) "Under Construction: How Commensuration and Management Fashion Affect Corporate Reputation Rankings," *Organizational Science* (forthcoming).

9. Board, S. and M. Meyer-Ter-Vehn (2013) "Reputation for Quality," *Econometrica*, Vol. 81, No. 6, pp. 2381-2462.

10. Bond, J., and Kirshenbaum, R. (1998) *Under the Radar: Talking to Today's Cynical Consumer*, (New York: John Wiley & Sons).

11. Clardy, A. (2012) "Organizational Reputation: Issues in Conceptualization and Measurement," *Corporate Reputation Review*, Vol. 15, No. 4, pp. 285-303.

12. Davies, J. (2014) "The 2014 GreenBiz NGO Report: How Companies Rate Activists," GreenBiz Group, http://www.greenbiz.com/research/report/2014/02/18/greenbiz-ngo-report.

13. Diermeier, D. (2011) *Reputation Rules: Strategies for Building Your Company's Most Valuable Asset,* (New York: McGraw-Hill).

14. Eesley, C. and M. Lenox (2005) "Firm Responses to Secondary Stakeholder Action." *Best Paper Proceedings of the 2005 Academy of Management Conference,* (Honolulu, HI).

15. Eesley, C. and M. Lenox (2006) "Secondary Stakeholder Actions and the Selection of Firm Targets," *Best Paper Proceedings of the 2006 Academy of Management Conference,* (Atlanta, GA).

16. Feddersen, T. and T. Gilligan (2001) "Saints and Markets: Activists and the Supply of Credence Goods," *Journal of Economics and Management Strategy*, Vol. 10, No. 1, pp. 149-171.

17. Fombrun, C. J. (1996) *Reputation: Realizing Value from the Corporate Image*, (Boston: Harvard Business School Press).

18. Harrison, A. and Scorse, J. (2010) "Multinationals and Anti-sweatshop Activism". *American Economic Review*, Vol. 100, No. 1, pp. 247-273.

19. Haufler, V. (2001) *A Public Role for the Private Sector: Industry Self Regulation in a Global Economy* (Washington, D.C.: Carnegie Endowment for International Peace).

20. Hendel, I., Lach, S. and Y. Spiegel (2013). "Social Media and Buyers' Power: The Cottage Cheese Boycott," Working Paper, Northwestern University.

21. Kaeb, C. (2008) "Emerging Issues of Human Rights Responsibility in the Extractive and Manufacturing Industries: Patterns and Liability Risks," *Northwestern Journal of International Human Rights,* Vol. 6, pp. 327-353.

22. King, B. and M-H. McDonnell (2012) "Good Firms, Good Targets: The Relationship between Corporate Social Responsibility, Reputation, and Activist Targeting," Working Paper, Northwestern University.

23. King, B. and N. Pearce (2010) "The Contentiousness of Markets: Politics, Social Movements and Institutional Change in Markets." *Annual Review of Sociology* 36, pp. 249-67.

24. King, B. and S. Soule (2007) "Social Movements as Extra-Institutional Entrepreneurs: The Effects of Protests on Stock Price Returns." *Administrative Science Quarterly* Vol. 52, pp. 413-442.

25. Lenox, M. and C. Eesley (2009) "Private Environmental Activism and the Selection and Response of Firm Targets," *Journal of Economics & Management Strategy,* Vol. 18, No. 1, pp. 45-73.

26. Lyon, T.P. and J. W. Maxwell (2002) "'Voluntary' Approaches to Environmental Regulation: A Survey, in In M. Franzini & A. Nicita (eds.), *Economic Institutions and Environmental Policy* (Aldershot, Hampshire, UK: Ashgate Publishing Ltd.).

27. Maxwell, J. W., T. P. Lyon, and S. C. Hackett (2000), "Self-Regulation and Social Welfare: The Political Economy of Corporate Environmentalism," *Journal of Law and Economics*, Vol. 43, No. 2, pp. 583-618.

28. Minor, D. and J. Morgan. (2012) "CSR as Reputation Insurance: Primum Non Nocere," *California Management Review*, Vol. 53, No. 3, pp. 40-59.

29. Vandenbergh, M.P. (2007) "The New Wal-Mart Effect: The Role of Private Contracting in Global Governance," *UCLA Law Review*, Vol. 54, pp. 913-970.

30. Vandenbergh, M.P. (2013) "Private Environmental Governance," *Cornell Law Review*,Vol. 99, pp. 101-166.

31. Vogel, D. (2010) "The Private Regulation of Global Corporate Conduct: Achievements and Limitations," *Business and Society*, Vol. 49, No. 1 (March 2010).

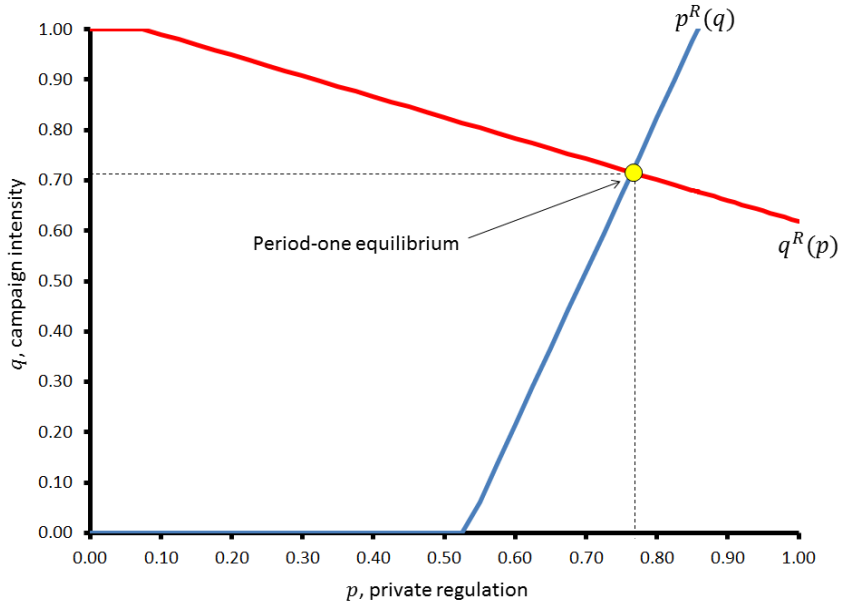

Figure 1: Firm and activist reaction functions and first-period equilibrium

	Social returns to reputation positive in states r and $r-1$	**Social returns to reputation negative in states r and $r-1$**	**Social returns to reputation positive in state r but negative in state $r-1$**
Activist increases period 1 welfare	• Activist decreases period 2 welfare • Activist has ambiguous effect on total welfare	• Activist increases period 2 welfare • Activist increases total welfare	• Activist has ambiguous effect on period 2 welfare • Activist has ambiguous effect on total welfare
Activist decreases period 1 welfare	• Activist decreases period 2 welfare • Activist decreases total welfare	• Activist increases period 2 welfare • Activist has ambiguous effect on total welfare	• Activist has ambiguous effect on period 2 welfare • Activist has ambiguous effect on total welfare

Figure 2: Possible scenarios for first-period, second-period, and total welfare

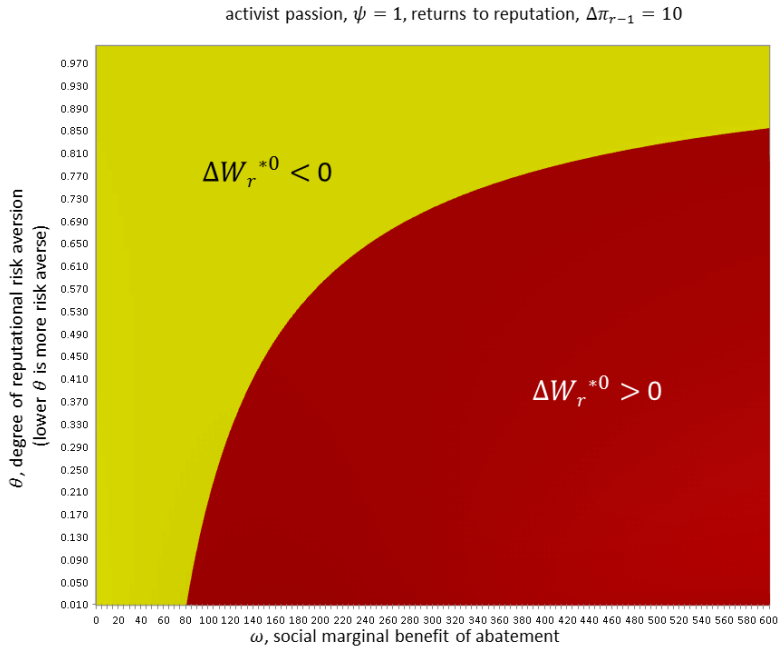

Figure 3: Difference in welfare in (ω, θ) space for $\psi = 1, \Delta\pi_{r-1} = 10$.

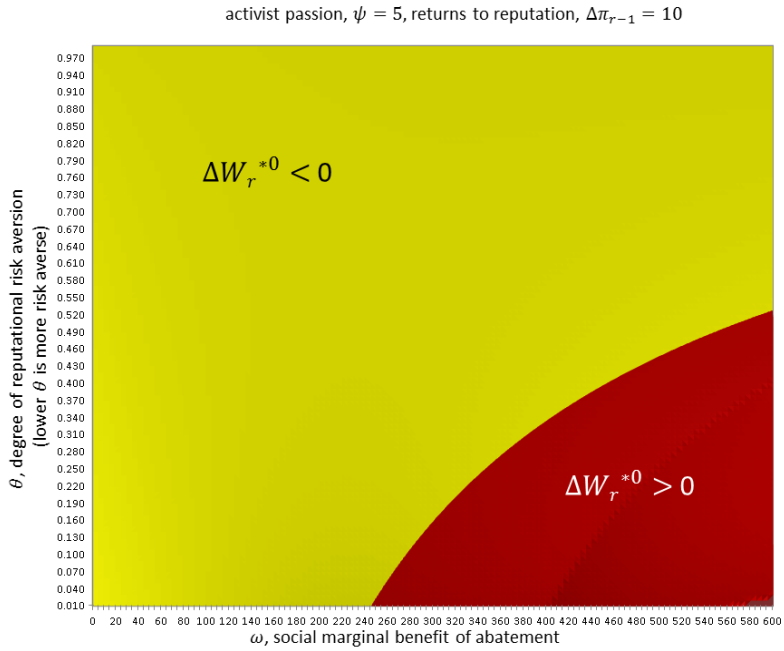

Figure 4: Difference in welfare in (ω, θ) space for $\psi = 5, \Delta\pi_{r-1} = 10$.

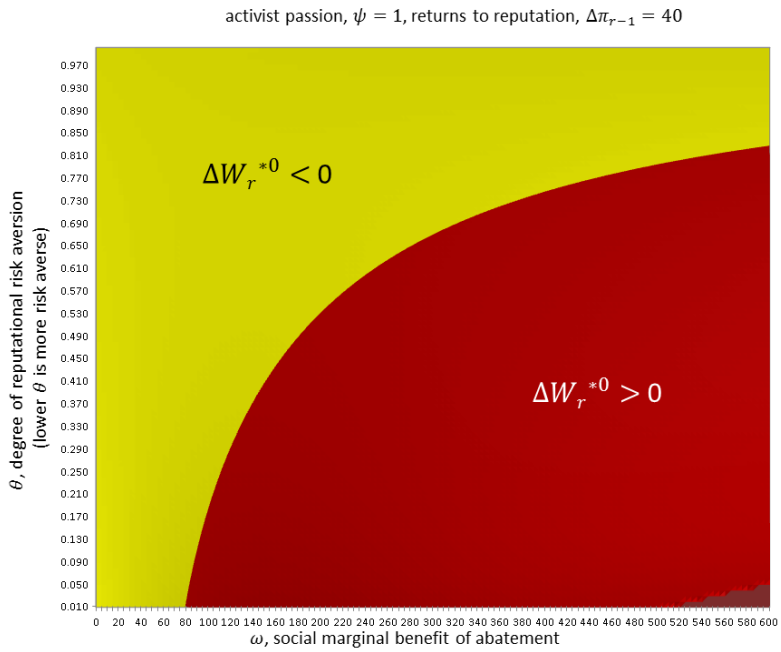

Figure 5: Difference in welfare in (ω, θ) space for $\psi = 1, \Delta\pi_{r-1} = 40$.

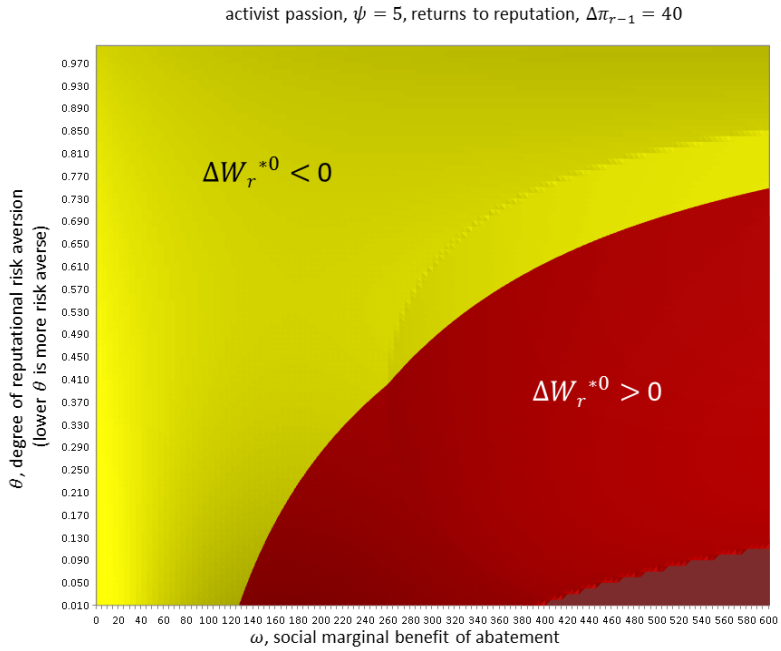

Figure 6: Difference in welfare in (ω, θ) space for $\psi = 5, \Delta\pi_{r-1} = 40$.

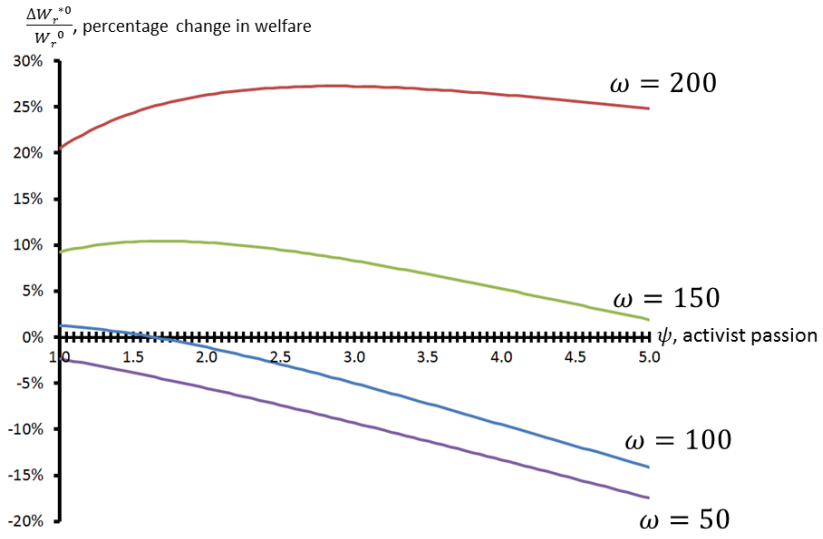

Figure 7: Change in welfare as a function of activist passion