Radical Activism and Self-regulation: An Optimal Campaign Mechanism

Mireille Chiroleu-Assouline, Ariane Lambert-Mogiliansky

PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

WORKING PAPER N° 2022 – 07

Radical Activism and Self-regulation: An Optimal Campaign Mechanism

Mireille Chiroleu-Assouline
Ariane Lambert Mogiliansky

JEL Codes: D44, D82, H23, L31, Q59
Keywords: Activist campaign, Mechanism design, Self-regulation, Repression

agence nationale de la recherche
AU SERVICE DE LA SCIENCE
INVESTIR L'AVENIR

Radical Activism and Self-regulation: An Optimal Campaign Mechanism *

Mireille Chiroleu-Assouline † Ariane Lambert-Mogiliansky ‡

February 23, 2022 §

Abstract

We study the problem faced by activists who want to maximize firms' compliance with high environmental standards. Our focus is on radical activism which relies on non-violent civil disobedience. The threat of disruptive actions is used to force firms to concede i.e., to engage in self-regulation. We adopt a mechanism design approach to characterize an optimal campaign. The analysis informs that the least vulnerable and most polluting firms should be targeted with disruptive actions while the others are granted a guarantee not to be targeted in exchange for a concession. This characterization allows studying the determinants of the activist's strength and how it is affected by repression, a central feature in civil disobedience. We find that an optimal campaign is relatively resilient to repression and that it creates incentives to free ride in prosecution for individual firms. Next, we consider heterogeneity in firms' abatement cost to find that an optimal campaign optimizes the allocation of abatment efforts and creates incentives for innovation. We discuss some other welfare properties of optimal campaign.

Keywords: Activist campaign, Mechanism design, Self-regulation, Repression.

JEL: D44, D82, H23, L31, Q59

---

*We would like to thank Bård Harstad, David Martimort, Thierry Verdier and Philippe Jehiel as well as PSE seminar participants for very useful comments. We are also grateful to participants in EAERE, IIPF and SURED conferences held in 2020.

† *Corresponding author.* Paris School of Economics, University Paris 1 Panthéon-Sorbonne, Mireille.Chiroleu-Assouline@univ-paris1.fr

‡ Paris School of Economics, alambert@pse.ens.fr

§ This research was funded by the French National Research Agency (ANR-15-CE05-0008-01) and the French government's Investissements d'Avenir program (ANR-17-EURE-001).

1 Introduction

The inadequacy of public action around the world to address the global environmental crisis, despite the scientific consensus on its urgency, (see e.g., the Intergovernmental Panel on Climate Change (IPCC) 2014 report or Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) 2019 report) is driving ordinary citizens to engage in various forms of activism to pressure firms and government institutions to act.

In this paper we are interested in the phenomenon of radical environmental activism defined as non-violent civil disobedience movements that engage in *disruptive* actions in order to pressure firms and institutional players to take action in response to the climate emergency (Gunningham, 2019). Recent examples include Ende Gaelende (EG) in Germany and Extinction Rebellion (XR) which is a global movement. Greenpeace is the best known precursor of these fast developing movements. Some of the most famous campaigns that Greenpeace has organized include the Kitkat campaign against Nestlé to raise awareness about deforestation due to the production of palm oil.[1] The significance of the threat to business of NGO's activism aimed at defending the environment, labour and human rights has been documented in Gans and Kaplan (2017) and Culling (2016).

Civil disobedience has been explored in the philosophical and political science literature (see e.g., Thoreau, 1849; Rawls, 2009; Lefkowitz, 2007, Rawls, 2013). A main issue has been its legitimacy and how it combines with the institutions of a democracy. Interestingly, the famous article 35 of the Declaration of Human Rights (1793) states that " *When the government violates the rights of the people, insurrection is for the people and for each portion of the people the most sacred of rights and the most indispensable of duties.*" For radical environmentalists, governments around the world fail to protect their citizens from the disastrous consequences of the current patterns of exploitation of natural resources. In view of the numerous failed attempts to, within the frame of laws, convince governments to take action, they argue that the situation is a one that legitimates civil disobedience.

In this paper we investigate the interaction between radical activists and firms (or government agencies). We focus on how activists can use their disruptive power to achieve the largest

[1] The campaign involved a number of radical actions against Nestlé including mass interventions of activists dressed as orangutans in Nestlé headquarters, in Amsterdan, Frankfort and London as well as in 7 of its factories in Germany in March 2010. On April 2010 they also disrupted the annual shareholders meeting in Lausanne. Most well-known is the social media campaign with the diffusion of a short video in breach of copyrights that induced a large mouvement against Nestlé (IBS, 2011).

improvement in firms' behavior with respect to the preservation of the environment. Firms' change in practice are instances of self-regulation and we refer to them as concessions (to the activists' demands).[2] The class of situations that we consider is characterized by three basic features. The first is that firms are unwilling to self-regulate because that involves costs. Therefore, in the absence of adequate regulation, our activists decide to act directly and confront firms.[3] In this respect, we are close to Galbraith (1952) and more recently Daubanes and Rochet (2019) who view activists as a countervailing force in face of regulators captured by industrial and financial interest groups. The second feature is that any single firm's vulnerability to disruptive actions (their cost of enduring an attack) typically depends on the complex interplay between technical and economic factors. Therefore, as in Baron and Diermeier (2007), we assume that the activist only has incomplete information about firms' vulnerability. The third ingredient is that we assume that actions of non-violent civil disobedience contribute to bringing public support to the activist i.e., it generates a symbolic or political gain. It has been shown that the gain is all the stronger when the actions, even illegal ones, are non-violent and the cause defended is perceived as legitimate (Chenoweth and Stephan, 2011; Feinberg et al., 2020).[4] In situations characterized by those three features, we are interested in the following questions: how should the activists' finite disruptive resources be used to maximize the value of concessions made by firms? What are the criteria for targeting with a disruptive action any particular firm? What are the determinants of the activists' strength? What is the impact of police and judicial repression? What are the welfare properties of an optimal activist campaign?

To answer these questions we develop a campaign mechanism where an activist (group) threatens with disruptive actions privately informed firms with the aim of extracting concessions in terms of self-regulation measures that reduce their environmental damage. While all the firms in the campaign are threatened only a smaller number are actually targeted, the others end up self-regulating. To achieve that outcome the activist effectively creates competition between firms for avoiding being targeted. This allows capturing an empirically established regularity.

[2] As a result of the Greenpeace Palm Oil campaign and its actions against Nestlé, many other firms chose to disassociate themselves from Sinar Mas (the Indonesian producer engaged in unsustainable exploitation) and committed to double-check the reputation of supply firms (see IBS (2011) p. 15).

[3] We acknowledge that some environmental NGOs are opting for more cooperative ways to push companies to change their behavior (Baron, 2012; Lyon, 2012), but we do not consider the potential trade-offs they face between different strategies and focus on disruptive strategies.

[4] Feinberg et al. (2020) show that extreme protest actions generally reduce popular support for social movements, but that this is no longer the case when the cause does not appear immoral and when there is an emotional connection and social identification of the public with the activists.

According to a number of studies (e.g. McDonnell and King, 2013; Hiatt et al., 2015; Briscoe and Gupta, 2016), environmental activism affects significantly more firms than those that could actually be harmed by activist actions. In addition, putting firms in competition with each other (to avoid being harmed) allows achieving the largest environmental gains from a given disruptive power in a situation of incomplete information. Formally, we adopt a mechanism design approach closely related to the developments in Koessler and Lambert-Mogiliansky (2014). The model covers quite general situations in which firms may be asymmetric with respect to their expected vulnerability and their efficiency in damage reduction and where the political gain (or cost) from taking action for the activist that may either be decreasing or increasing in firm's vulnerability.

The optimal campaign mechanism that we derive is characterized by the thresholds that determine whether a firm is targeted or not. The thresholds depend on firms' vulnerability and the political gain associated with taking action. Little vulnerable big polluters are the first to be targeted. The intuition is that the more vulnerable a firm the more it is willing to concede to avoid an attack. The optimal campaign mechanism also determines the magnitude of the concessions paid in equilibrium by the firms that are not targeted. For the case firms are ex-ante symmetric, a simple auction-concession game implements the optimal mechanism. Next, we investigate the determinants of the strength of the activist measured by three indicators: the probability for action, the size of concessions and the global gain from the campaign (concessions and political gains). Our results are quite intuitive and allow addressing the impact of repression. Ghandi and Martin Luther King viewed the acceptance of physical and legal risks as a core principle of civil disobedience. This is because repression gives the opportunity to expose to a larger public the illegitimacy of a policy and the immorality of repression (DeNardo, 1985; Opp and Roehl, 1990; Chenoweth and Stephan, 2011; Ayanian and Tausch, 2016; Aytaç et al., 2018). Davenport (2007) defines the "punishment puzzle" as the empirically established fact that the impact of repression can go either way, i.e., to embolden the activists or reduce their strength (e.g., Francisco, 1996; Gupta et al., 1993; Young, 2019; Edwards, 2021). We take these two contrasting effects into consideration and we find that in our model the class of situations in which repression is counter-productive is non empty: optimal non-violent activism is quite resilient. Increased repression may lead to more actions, larger concessions and a larger global gain. We note that our results imply that the government faces rather extreme choices: either it effectively concedes to disobedience (implementing a strategy of 'negotiated accommodation' as advocated by Smith, 2012) or it applies harsh repression. As we consider firm specific repression (e.g., prosecution), the optimal campaign mechanism is shown to create incentives for firms to free ride on prosecution. The intuition is that a firm known for prosecuting activists becomes a more attractive target for actions relative to other firms. Competition to avoid being targeted benefits the activists yet in another way. We next focus on heterogeneity in firms' abatement efficiency. We show that our model can be adapted to this instance of non-transferability in utility and we find that our optimal campaign mechanism delivers an efficient allocation of the global abatement cost between firms. The firms most efficient with respect to pollution abatement are more prone to concede under the threat because of lower compliance costs. In addition, an optimal campaign induces interesting cross effects that stimulate innovation in damage reduction. Thereafter, we consider some extensions to address the case when attacked firms concede, how the activist selects the population of firms, the case when the potential targets are government agencies and an alternative formulation of the political gain. Finally, we address some social welfare issues to note that optimal activism is not Pareto efficient by construction but is endowed with interesting efficiency properties. We conclude with an extensive discussion of the empirical literature.

Related Literature. There exists as for today a limited literature on activism in economics. Baron and Diermeier (2007) study strategic activism in a model where an activist confronts a firm with a campaign including a demand, a reward and a disruptive action. In their article the authors touch upon a number of questions and make conjectures. In particular, they note that activism shows similarity with extortion and that the activist can benefit from competition between firms. The present work investigates these conjectures. Baron (2012, 2016) is also closely related to the present work. He analyses the choices of activists who use the threat of campaign to induce firms to self-regulate. Our approach is however quite different. Baron's model is a one of market for activists where activists monitor firms in bilateral relationships while we develop a model where firms compete in concessions under threat. Daubanes and Rochet (2019) build on Baron's model to explain the growing role of NGOs as a response to the capture of regulators by powerful industry lobby. In a different setting, Egorov and Harstad (2017) study the dynamics of conflict between NGOs and firms and investigate when a benevolent regulator can resolve the situation. Closely related is also Heyes, Lyon and Martin (2018) who study a salience game where an industry and NGO compete in influence over limited public attention to social or environmental impacts.

Our paper is related to another strand of literature that addresses firms' incentive to self-regulate. A central contribution is in Maxwell et al. (2000). They investigate preemptive environmental self-regulation in a model where there is a threat of regulation determined by a costly influence game between firms and consumers.

Our work also contributes to the limited literature on coercion in economics. Konrad and Skapergas (1997, 1998) adopt a game theoretical approach to study racketeering by criminal gangs with focus on the issue of the credibility of the threat. Koessler et al. (2014) adopt a mechanism design approach to study bureaucratic corruption and its response to political risk insurance. In the same vein, we use the tools of mechanism design to characterize an optimal activist campaign and its resilience to repression. The ambiguous effect of repression of collective action was addressed by Siegel (2011), who adopts a network theoretical argument to investigate how participation is affected by repression. In Shadmehr and Boleslavsky (2021), it is the public's uncertainty about the merits of activists' demands and the intentions of state repression that can lead to the stifling or, conversely, the amplification of the movement.

A main contribution of this paper is to provide a formal analysis of the power of the threat of disruptive actions on firms's behavior with respect to the harm their activities bring to the environment. For that purpose, we have characterized the properties of an optimal campaign mechanism aimed at maximizing firms' concessions to environmental damage reduction relying on limited disruptive resources. The mechanism design approach features both competition between firms and the similarity with extortion suggested by Baron and Diermeier (2007). We show how it allows optimally exploiting the threat of disruptive actions by creating competition between firms while accommodating imperfect transferability reflecting heterogeneity with respect to firms' abatement efficiency. Our analysis characterizes conditions for repression to be counter-productive when directed against non-violent activism. A free riding effect in prosecution arising from competition between firms to avoid disruptive actions further reveals the power of the mechanism. The analysis provides a contribution to the burgeoning litterature on the economics of coercion.

The remaining of the article is structured as follows. In the next section we present the framework and characterize the optimal campaign mechanism. In Section 3, we focus on the determinants of the strength of activism. In Section 4, we examine the impact of police and judicial repression. Section 5 develops the campaign mechanism when firms are asymmetric with respect to their damage abatement efficiency. Several extensions of the basic model are addressed in Section 6. Section 7 gathers some elements of welfare analysis and Section 8 provides a discussion of our results in light of the existing literature.

2 Optimal Campaign Mechanism

2.1 Basic Model

The firms

We have a population $N = \{1, \ldots, n\}$ of (risk neutral) firms whose activity generates a negative externality on the environment. The externality can be reduced at cost for the firms. In absence of any effort to contain environmental damages, all firms earn the same profit π. We let each individual firm i be characterized by a few parameters (h_i, D_i, α_i). The most central to the analysis is h_i, the firm's vulnerability $h_i \in H_i \equiv [a_i, b_i]$, where $0 \leq a_i < b_i < \pi$. It captures the losses it incurs when targeted by a disruptive action (see below). $D_i \in R^+$ captures the value of the environmental damage resulting from its activity (h_i and D_i are assumed to be independent) and α_i is abatement efficiency. The vulnerability parameter reflects a variety of features. For instance, the existence of close substitutes tends to make a firm vulnerable as consumers can more easily express their support to the activist by switching to a substitute in response to action. Similarly, a visible brand helps consumers identify the firm which increases its vulnerability as long as brand loyalty is not too strong. Another source of vulnerability is when a firm's activity includes key processes in e.g., transportation or production that are easy to disrupt at high cost to the firm. The true value of the harm resulting from an action against firm i is assumed to be unknown to the activist (and the public). It depends on the combination of a number of factors including the firm's technical and economic ability to mitigate some of the harm.[5] For instance when considering blocking Amazon, the activist does not know whether Amazon has alternative means of storage to face delays and/or capacity to mobilize extra resources from other locations when delivery can resume.

Only firm i knows the true value of its vulnerability or "type" h_i. Public information about vulnerability is given by $f_i : H_i \rightarrow \mathbb{R}_+$ the continuous density function for i's type, and F_i the corresponding cumulative distribution function. For simplicity we assume that firms' types are independently distributed. When it comes to the environmental damage D_i, we assume that it is publicly known e.g., CO_2 emissions.

When firm i is targeted by an action its payoff is

$$\pi - h_i.$$

[5] We do not exclude that the activist has received some signals about the vulnerability of the firms. The distribution can be interpreted as capturing the residual uncertainty.

Firm i can make concessions $x_i(h) \in \mathbb{R}$ to avoid being targeted so it earns

$$U_i = \pi - x_i(h)$$

where h is the vector of firms' announced vulnerabilities (see below). Concessions are self-regulating measures aimed at reducing the firm's detrimental impact on the environment. In the basic model, firms have the same average cost for abatement $\alpha_i = \alpha$ for all i, and it is normalized *to* $\alpha = 1$. We relax this assumption in Section 5.

The activist

There is one (risk neutral) activist organization, referred to as "she", the activist. Her objective is to minimize firms' total harm to the environment by means of self-regulating measures (concessions) to reduce the detrimental impact of their activity on the environment e.g., to abate their CO_2 emissions. The instrument she uses to achieve that goal is the threat of disruptive actions.[6] Her capacity to effectively carry out disruptive actions, $A_i \in \{0,1\}$, $i = 0, ..., n$, is limited: $\sum A_i \leq k$, for some $k < n$. We are dealing with civil disobedience, therefore actions are connected with legal risks (costs) to the activist but also with political gains arising from the sympathy of the public. When targeting a firm i with an action, the activist receives $w_i(h_i, D_i; Z) \in \mathbb{R}$. It represents the activist's net gain(loss) from implementing the action including impact on public awareness (support for the action and the cause), her own reputation and credibility less the cost of carrying out the action. The functions $w_i(.)$ are parametrized by Z, the (exogenous) level of repression. For most part of the paper we take Z as given so we write $w_i(h_i, D_i; Z) = w_i(h_i, D_i)$. In Section 4, where we investigate the impact of repression on the equilibrium solution by considering its respective effects on $w_i(.)$ and k, we also briefly consider firm specific levels of repression ($Z_i \neq Z_j$) e.g., a firm may have a reputation for filing complaints and prosecuting activists or on the contrary for keeping a low profile.

The action

The term action is used to refer to a collection of measures that reduces the firms's ability to earn profit. Actions include symbolic disruptions like staging a die-in or a happening but also organizing a boycott, a blockade of the activity or performing some sabotage.[7] They are

[6] Here, we depart from Baron and Diermeier (2007) and Baron (2012, 2016) in viewing disruptive actions as a mean to build up a threat rather than using them to obtain concession of targeted firms. In Section 6 we show how our model can accomodate concession from attacked firms.

[7] In some respects, radical activism of the kind investigated here reminds of actions used in labour conflicts like strikes and blocage. Both strikers and activists aim at inflicting a cost on the firm as a mean to pressure its managers into concession. While the right to strike is recognized, the right for citizens to temporarily block an

aimed at affecting the firm's reputation, the demand for its products or directly disrupting its production activity. While actions are binary variables $A_i \in \{0, 1\}$, the impact of $A_i = 1$ on a firm i and on the activist's objective depends on the type of the firm (h_i). The cost to the firm is a loss equal to h_i and the political value to the activist is captured by $w_i(h_i, D_i) \in \mathbb{R}$. For most part of the paper, we assume that $\frac{\partial w_i(h_i,D_i)}{\partial h_i} \leq 0$ and $\frac{\partial w_i(h_i,D_i)}{\partial D_i} \geq 0$ i.e., the political gain is decreasing in the vulnerability of the firm and increasing in its environmental damage.[8] This is aimed at capturing the idea that confronting "a big villain", with low vulnerability and high detrimental impact, is more visible and tend to gain more public support than targeting a conciliatory firm.[9] In an extension (6.4), we consider the case when $\frac{\partial w_i(h_i,D_i)}{\partial h_i} \geq 0$.

The mechanism

The mechanism relies on the activist's exploitation of her disruptive power to achieve maximal improvement in firms' production practice. It determines the probability for each firm to be targeted by a disruptive action and how much each firm must "pay" to avoid being targeted. The payment is in terms of concessions to environmental demands (abatements). We shall be also be interested in the global gain to the activist (the sum of the value of conceded abatements plus the political gain from action). The activist does not know the true value that firms attach to avoiding actions (their vulnerability) or equivalently the maximal concession that each firm is willing to make to avoid being targeted. If the activist perfectly knew each firm's vulnerability, she would be able to obtain the maximal concession from each of them by threatening any non-obedient firm with action. We adopt a mechanism design approach.[10] Since each firm privately knows its own vulnerability (i.e., the loss of profit induced by the action), the activist has a role similar to the designer of an auction mechanism with private values. The activist is "selling" *guarantee not to target a specific firm* in exchange for a concession from that firm. The activist aims at maximizing her revenue, here the sum of environmental concessions (and political gains). When adopting this approach we give the activist commitment power: she can commit to target and *not* to target a firm.[11]

environmentally harmful activity is not.

[8] In our model, the political gain does not include possible change in pratice following the action.

[9] This political gain depends on the characteristics of the targeted firm (den Hond and de Bakker, 2007).

[10] This approach implies that firms' bargaining power is minimized. Clearly firms can protect themselves against activists using government action, public relations or misinformation. In our approach this is (implicitely) accounted for in the realized vulnerability which determines the outcome.

[11] This is equivalent to assuming that activist groups respect each other's campaigns. To the extent that they compete for public support, they do so by staging actions against different targets or at different time so as not to interfere in each other's campaign.

The optimal mechanism that we characterize is obtained by slightly adapting the design of an optimal auction in Myerson (1981). The differences with Myerson's original setting are similar to those in Koessler et Lambert-Mogiliansky (2014, hereafter KLM). In addition, in our setting utility is not perfectly transferable (the value of a concession for the activist may not be the same as for the firm). Despite those differences, as in KLM the formal analysis of optimal auction mechanism carries over to our setting.

A (direct revelation) mechanism is given by outcome functions $p : H \to [0,1]^n$ and $x : H \to \mathbb{R}^n_+$. Given a profile of announced types $h = (h_1, \ldots h_n)$, $p_i(h)$ is the probability of *not* targeting firm i and $x_i(h)$ is the expected magnitude of the concession made by firm i. Although concessions to the activist demands do not take the form of money transfers, this formulation measures concession in terms of their monetary cost to the firm (in the basic model $\alpha_i = \alpha = 1, \forall i \in N$). In addition, it assumes that the activist utility is linear in that monetary cost.

Given a mechanism (p, x) the (interim) expected utility of firm i when its type is $h_i \in H_i$ is given by

$$U_i(p, x; h_i) = \int_{H_{-i}} (h_i p_i(h) - x_i(h))\, f_{-i}(h_{-i}) dh_{-i}, \tag{1}$$

and the (ex ante) expected utility of the activist is

$$U_0(p, x) = \int_H \left(\sum_{i \in N} (1 - p_i(h)) w_i(h_i, D_i) + x_i(h) \right) f(h) dh. \tag{2}$$

A mechanism is feasible if it satisfies the individual rationality (IR) constraint

$$U_i(p, x; h_i) \geq 0 \text{ which is equivalent to } h_i - x_i(h) \geq 0, \text{for all } i \in N, \tag{3}$$

and the incentive-compatibility (IC) constraint has standard form

$$U_i(p, x; h_i) \geq \int_{H_{-i}} (h_i p_i(s_i, h_{-i}) - x_i(s_i, h_{-i}))\, f_{-i}(h_{-i}) dh_{-i}, \quad \text{for all } i \in N,\ s_i, h_i \in H_i. \tag{4}$$

Condition (3) means that firms must get an expected payoff which is at least as large as the expected payoff they obtain when they are targeted with probability one. Any firm that refuses to participate is targeted.[12] This is the coercive feature of the mechanism which we have in common with the extortion set-up in KLM.[13] Condition (4) means that firms have no incentive

[12] Note that in equilibrium the threat is credible in spite of the activist resource constraint. In face of a deviation i.e., a (single) firm that refuses to concede, the activist replaces one of the targeted firms with the deviating one.

[13] By definition of Bayes Nash Equilibium, we consider deviation by single firms. Allowing for the possibility that some firms collude against the activist is a significant departure which requires an analysis of its own. Obviously if no firm ever concedes the mechanism loses its purpose but there may be a multiplicity of intermediary cases.

to misreport their types to the activist when they expect that all other firms truthfully report their types.

In addition to these standard constraints, we have a resource constraint (RC): the activist can target at most $k \in \{1, \ldots, n\}$ firms, so that the probabilities for getting the guarantee not to be targeted must satisfy

$$\sum_{i \in N} p_i(h) \geq n - k, \qquad \text{for all } h \in H. \tag{5}$$

The magnitude k is exogenous to the mechanism, it captures the fact that actions are both resource and time-consuming. In particular action requires the participation of people.

2.2 Feasible and Optimal Mechanisms

The activist selects the mechanism (p, x) that maximizes her expected payoff $U_0(p, x)$ under the above IR constraint (3), IC constraint (4) and RC (5). Following exactly the characterization in Myerson (1981), the optimal mechanism is given by

$$x_i(h) = p_i(h)h_i - \int_{a_i}^{h_i} p_i(s_i, h_{-i})ds_i, \tag{6}$$

and $p : H \to [0, 1]^n$ that maximizes

$$\int_H \sum_{i \in N} \underbrace{\left(h_i - w_i(h_i, D_i) - \frac{1 - F_i(h_i)}{f_i(h_i)} \right)}_{c_i(h_i)} p_i(h)f(h)dh, \tag{7}$$

subject to the RC (5) and the monotonicity constraint of the interim probability $\int_{H_{-i}} p_i(h)f_{-i}(h_{-i})dh_{-i}$ that firm i of type h_i is not targeted. The expression $c_i(h_i) = h_i - w_i(h_i, D_i) - \frac{1-F_i(h_i)}{f_i(h_i)}$ is referred to as the virtual type of firm i. It includes the true type h_i minus a term related to the firm's information rents $\frac{1-F_i(h_i)}{f_i(h_i)}$ minus the political gain from *not* selling the guarantee, i.e., from targeting that firm, $w_i(h_i, D_i)$.

The function $w_i(h_i, D_i)$ captures the political gain(loss) to the activist. As earlier mentioned, we assume that it is decreasing in the vulnerability of the firm i.e., $\partial w_i(h_i, D_i)/\partial h_i \leq 0$. This secures that $c_i(h_i) = h_i - w_i(h_i, D_i) - \frac{1-F_i(h_i)}{f_i(h_i)}$ is strictly increasing in h_i which in turn secures that the problem is regular. Note that this formulation assumes that information about h_i becomes public as a result of the action so the political gain can be realized.[14] We immediately

[14] This formulation simplifies the presentation but we could as well assume that the political gain depends on some informative signal of vulnerability available ex-post. For instance, the public may learn about a targeted firm's vulnerability from the market's reaction to the action. The function $w_i(h_i, D_i)$ would then be interpreted as expected political gains.

get the following characterization of the optimal mechanism:

Proposition 1 (Optimal Activism) *Under regularity, the optimal extortion mechanism* (p, x) *is such that* $p : H \to [0, 1]^n$ *maximizes*

$$\sum_{i \in N} c_i(h_i) p_i(h) \quad \textit{subject to} \quad n - k \leq \sum_{i \in N} p_i(h) \leq n \quad \textit{for all } h \in H,$$

where the virtual type $c_i(h)$ *of firm* i *is given in (7). That is,* $p_i(h) = 0$ *for the firms with the (up to)* k *lowest virtual types below* 0*, and* $p_i(h) = 1$ *for the others. Firm* i *(*$p_i(h) = 1$*) pays a concession to the activist given by* $x_i(h) = \min\{c_i^{-1}(0), \min_{j \neq i}^k c_j(h_j)\}$*. The other firms (*$p_i(h) = 0$*) pay nothing.*

For each firm i the optimal mechanism involves a (possibly firm specific) threshold value $c_i^{-1}(0)$ for non-targeting which is determined so that the virtual type of firm i is equal to zero. The threshold value plays a role similar to the reserve price in optimal auction mechanisms, and is chosen by the activist in order to maximize the expected concessions. If the activist has no resource constraint ($k = n$) she never grants a guarantee not to target a firm for a concession below that threshold value. She grants to each firm i a guarantee not to be targeted in exchange for a concession of value $c_i^{-1}(0)$, and if firm i does not concede accordingly, it is targeted. When the activist cannot target as many firms as she wishes, i.e., when she is forced to grant a guarantee of non-targeting to at least $n-k$ firms, she cannot obtain the threshold values from all non-targeted firms. Instead, she must decrease the concession for a non-targeted firm i to $\min\{0, {\min_{j \neq i}}^k c_j(h_j)\}$,[15] i.e., the highest concession acceptable to the lowest non-targeted firm i's type. As a result the role of the threshold value in the activist mechanism is somehow more limited the tighter the RC (the smaller k) and the larger the total number of firms.

When firms are ex-ante symmetric, i.e., $w_i(\cdot) = w_j(\cdot)$ and $f_i(\cdot) = f_j(\cdot)$ for every $i, j \in N$, we denote by $h_0 = c_i^{-1}(0)$ the optimal and common threshold value for non-targeting. In that case, the optimal mechanism is much simpler. Any firm i whose type h_i is above h_0 is never targeted ($p_i(h) = 1$). When the RC is not binding (i.e., $|\{i \in N : h_i < h_0\}| < k$), every firm i whose type h_i is below h_0 is targeted ($p_i(h) = 0$) and concedes nothing, and the others are not targeted and make a concession corresponding to the threshold value h_0. When the RC is binding (i.e., $|\{i \in N : h_i < h_0\}| \geq k$), then only the k firms whose types are the k lowest types

[15] For any finite set $\{x_1, x_2, \ldots\}$ of real numbers, denote by $\min_i^k x_i$ the $k-$th smallest element of this set. That is, if $x_1 < x_2 < \cdots < x_k < \cdots$, then $\min_i^k x_i = x_k$.

below h_0 are targeted and concede nothing, and the others are not targeted and make the same concession of value: $\min_{j\in N}^{k} h_j$, the $k-$th lowest type in $\{h_1, \ldots, h_n\}$. Notice that contrary to standard auctions, when the RC is binding the effective concession ($\min_{j\in N}^{k} h_j$) may be strictly lower than the activist's "reserve price" (h_0).[16]

In the symmetric case, the optimal mechanism can be implemented through a simple concession game similar to a second price auction with a reserve price: each firm $i \in N$ simultaneously and voluntarily submits a concession offer $o_i(h_i) \geq 0$ as a function of its type $h_i \in H_i$; then, up to k firms with the lowest bid below h_0 are targeted with action, and the others are not targeted and concede $x = \min\{h_0, \min_{j\in N}^{k} o_j(h_j)\}$. Observe that, like in second-price auctions, it is a weakly dominant strategy for each firm i to bid its value: $o_i(h_i) = h_i$ for every $h_i \in H_i$. Like in auction mechanisms, if firms are not ex-ante symmetric, then the optimal mechanism takes into account the heterogeneity of firms' observable characteristics e.g., firm specific damage.

As an illustration of our results, we consider a simple example that we use throughout the paper.

Example 1 Assume that for every $i \in N$ the vulnerability h_i of firm i is uniformly distributed on $[a, b]$ with $0 \leq a < b$, and the action value for the activist of an action against firm i is $w_i(h_i, D_i) = (\gamma D - h_i)$, where $D \geq 0$ is the common damage and $\gamma < 1$ reflects the public's awareness of the damage. Then, the virtual type of firm i is given by

$$c(h_i) = h_i - (\gamma D - h_i) - \frac{1 - (h_i - a)/(b - a)}{1/(b - a)} = 3h_i - (\gamma D + b)\,.$$

The (common) threshold value for non-targeting is $c^{-1}(0) = h_0 = \frac{\gamma D + b}{3}$. The capacity constraint is binding only when the $k-th$ lowest type is below h_0. The optimal mechanism characterized in Proposition 1 is such that up to k firms with the k lowest types below $h_0 = \frac{\gamma D + b}{3}$, are targeted, and the others concede and pay $\min\{h_0, \min_{j\in N}^{k} h_j\}$.

With heterogenous damages D_i, we obtain firm specific threshold values: $c_i^{-1}(0) = h_{i0} = \frac{\gamma D_i + b}{3}$. The activist grants to each firm i a guarantee not to be targeted in exchange for a concession of value $\frac{\gamma D_i + b}{3}$ and targets with action all firms that do not concede accordingly. When the activist cannot target as many firms as she wishes, the concession for a non-targeted firm i

[16] Note that firm $i's$ equilibrium concession could lie above its damage i.e., $x_i^*(h) > D_i$. The interpretation is that such a firm commits not only to eliminate its own dammage but also to contribute the restauration of damaged environment.

reduces to $\min\{\frac{\gamma D_i+b}{3}, \min^k_{j\neq i, j\in N}\frac{\gamma D_j+b}{3}\}$, where $\min^k_{j\neq i, j\in N}\frac{\gamma D_j+b}{3}$ is the highest concession acceptable to the lowest non-targeted firm i's type.

Optimal mechanism with asymmetric firms

When firms are not ex-ante symmetric with respect to their expected vulnerability as in the example above, the optimal mechanism discriminates among different firms depending on their vulnerability distributions. To see this, consider two different firms i and j with the same vulnerability (type) y and such that $w_i(y, D_i) = w_j(y, D_j)$, we notice that

$$c_j(y) \geq c_i(y) \iff \frac{1-F_j(y)}{f_j(y)} \leq \frac{1-F_i(y)}{f_i(y)}.$$

Hence, firm j, associated with a higher hazard rate $\frac{f_j(y)}{1-F_j(y)}$, will be targeted less often and will have to make smaller concession than firm i with the same vulnerability y as firm i.

In the equilibrium of Proposition 1 attacked firms never concede.[17] This result underlines the distinction with Baron and Diermer (2007) and Baron (2016). Our approach provides the solution for the optimal use of disruptive resources to achieve concessions. It therefore focuses on the power of the threat of actions to induce concessions from non-targeted firms. In Section 6.1 we suggest how our model can accommodate concessions from firms that have been attacked.

Finally, note that the main results in Proposition 1 do not hinge on the political gain function and thus apply to more traditional forms of activism that exert costly pressure on firms. In the remaining of the paper, we provide results that directly address the specificities of radical activism.

3 The Strength of Optimal Campaigns

In this subsection we are interested in the determinants of the strength of radical activism in optimal campaigns. We measure that strength with three indicators, the probability that a firm is targeted, the magnitude of its concessions when not targeted and the global gains from the campaign. Our first result addresses the contribution of the political gain to the strength of the activist campaign.

[17] Our analysis echoes the work of Heyes et al. (2018) who also consider an activist facing a group of companies. In their context allowing some of firms not to concede and causing some disaster could be exploited by the activist to gain support for the cause.

Proposition 2 *For each firm i, the probability of being targeted of this firm, the concession made by this firm when it is not targeted, and the global gain of activism are increasing with the political gain from action $w_i(\cdot)$.*

Proof. Consider a political gain function $\widetilde{w}_i(\cdot)$ of some firm $i \in N$ such that $\widetilde{w}_i(\cdot) > w_i(h_i, D_i)$ for every $h_i \in H_i,\ D_i \in R^+$. Then, the virtual type of firm i is given by $\tilde{c}_i(h_i) < c_i(h_i)$ for every $h_i \in H_i$, which implies that the threshold value for non-targeting is $\tilde{c}_i^{-1}(0) > c_i^{-1}(0)$. Hence, the probability of being targeted of firm i and the concession made by firm i when it is not targeted are higher with $\widetilde{w}_i(\cdot)$ than with $w_i(\cdot)$. To show that the global gain of the activist is also higher with $\widetilde{w}_i(\cdot)$ than with $w_i(\cdot)$ it suffices to notice that the optimal mechanism with $w_i(\cdot)$ is also feasible with $\widetilde{w}_i(\cdot)$ because the value of action does not enter into firms' utilities, and yields the same concession but a higher political value of actions. Therefore, the optimal mechanism with $\widetilde{w}_i(\cdot)$ necessarily yields a higher total expected global gain for the activist. ■

Proposition 2 establishes that the larger the political gain from attacking a firm, the stronger the activist along all three indicators. Note that this is a general comparative statics property of the optimal mechanism for arbitrary distributions of types and for values of action that are not necessarily symmetric and linear in firms' types. By assumption, the political gain is an increasing function of the damages D_i, so the more polluting the firm the "better" for the activist. This suggests that to the extent that the activist chooses the population of firms, she should include the biggest polluters (i.e., with large D_i). We return to that issue in section 6.2.

Our next result addresses the contribution of the resource available to the activist. We consider the resource constraint of radical activism to be primarily a reflection of the population's commitment to environmental protection as expressed in the size and readiness of its most radical fringe. This includes the amount of time and energy activists are willing to devote to implementing actions or the amount of physical and legal risk they are willing to take. It may also capture donors' financial support to the activist organization.

The resource constraint limits the activist's ability to target firms with actions. When she can target as many firms as she wants i.e., $k \geq n$, the optimal mechanism calls for targeting all the firms with $c_i(h_i) < 0$ and for each firm i the types that are not targeted make a fixed concession $c_i^{-1}(0)$, which is independent of other firms' types. When the activist can target at most k firms, $k < n$, the k firms with the k-lowest virtual types below 0 are targeted and the others make concession equivalent to the smallest possible types allowing them not to be targeted given others' types. Hence, the weaker the RC (the larger k) the larger the probability for being targeted and the concession made by each firm. Since k only appears as a constraint in the activist's optimization program (through Equation (5)), her revenue (i.e., the sum of concessions) is also increasing in k.

Proposition 3 *The risk of action, the concession made by the firms when they are not targeted, and the global gain of activism are increasing with the number k of firms the activist has the resource to target with actions.*

Proof. It follows directly from Proposition 1 and the observations above. ■

Propositions 2 and Proposition 3 above set the stage for addressing the impact of police and judicial repression on optimal campaigns.

4 Optimal Campaign and Repression

The rationale for civil disobedient environmental activism is the conviction that if the public understood the true reasons for the environmental situation, it would support limiting firms' destructive impact on global warming and biodiversity. On the one hand, the disruptive impact of an action creates an opportunity to bring publicity on a concrete example of bad practice and raise awareness more broadly. On the other hand, the associated physical and legal risks serve the activist's objective in several ways. The risks signal the sincerity and seriousness of the engagement (Brown, 2021). And the prosecution of activists provides a forum that allows to elaborate further on the legitimacy of the cause. In addition, a trial is used by activists to demonstrate the immorality of laws that treats as ordinary offenders people who take high personal risks to promote better environmental protection while serious deeds from profit seeking firms and/or captured government agencies are left unpunished.[18]

The disruptive actions that we consider may violate law and thus give rise to police and judicial repression aiming at securing the rule of law and maintaining public order (Smith, 2012). Repression intimidates activists (reduces available resources) by the prospect of facing costs in terms of physical harm, fines and jail terms. But in line with the rationale for civil disobedience, we account for a positive impact of repression in terms of political gains.

[18] One example is a recent mass action in France. On October 3rd 2020, 2000 people participated blocking several airports around the country. Around 150 people were arrested with some spending more than 70 hours in custody. The trial in October 2021 saw climate specialists witnessing of the state of necessity while the government granted unconditional financial aid to Air France in the Covid pandemic. The tribunal decided in favor of the activists.

In our model the public support impact produced by repression is explicitely captured by the function $w_i(h_i, D_i) = w_i(h_i, D_i; Z)$ with $\frac{\partial w_i}{\partial Z} \geq 0$ i.e., increasing the harshness of repression increases the political gain from action against firm i for any level of vulnerability, h_i and damage D_i. We assume that $\frac{\partial w_i}{\partial Z} \geq 0$ in the logic of civil disobedience: the sympathy from the public is rising in the extent and harshness of the police and judicial repression faced by the activist. Repression also affects the resource constraint (k) though intimidation. The impact in this respect is likely to be ambiguous. On the one hand repression increases the costs of taking action so fewer the people dare participating which reduces the number of feasible actions. On the other hand repression can rally people to join as outrage stimulates sacrificial spirit which loosens the resource constraint. Below we adopt a conservative approach and assume repression tightens the resource constraint.

Some aspects of repression can be firm specific ($\frac{\partial w_i(h_i,D_i;Z_i)}{\partial Z_i} \geq 0$). In particular this is true for judicial prosecution. Following an action e.g., the occupation of Amazon's offices in Paris on July 2$^{\text{nd}}$ 2019, the corporation chose not to file a complaint. A trial would have provided activists with a forum for explaining their action by exposing evidence of the firm's social and environmental records which would have increased their political payoff at the expense of the firm's reputation. Proposition 4 below provides a rationale for Amazon not to prosecute.

Relying on our results in Propositions 2 and 3 above, Proposition 4 characterizes the impact of repression on optimal environmental activism.

Proposition 4 *(i). The class of situations where repression is counter productive is strictly non empty:*

- *(i.a). In situations where the resource constraint is not binding, increased repression leads to more actions, larger concessions and a larger global gain for the activist.*
- *(i.b). In situations where the resource constraint is binding, repression decreases the probability for firms of being targeted and the magnitude of concessions. The impact on the global gain is ambiguous.*

(ii). For any single firm i, an increase in firm specific repression is unambiguously detrimental to that firm. It may however benefit other firms.

The proof of $(i.a)$ and $(i.b)$ follows immediately from Propositions 2 and 3 above. To establish (ii) we use the proof of Proposition 3. We note that since $\frac{\partial w_i(h_i,D_i)}{\partial Z_i} \geq 0$, if $\widetilde{Z}_i > Z$ where Z is e.g., standard level of repression that applies, we have $\widetilde{w}_i(\cdot) > w_i(h_i, D_i)$ for every $h_i \in H_i,\ D_i \in R^+$.

Because $\widetilde{Z}_i > Z$ means a higher political gain from action against firm i, the threshold for not targeting firm i will be higher for $\widetilde{Z}_i$. A higher threshold implies a larger probability for action against firm i and a higher concession to avoid action. An increase in the political value of the action against firm i may be beneficial to another firm j. Indeed larger values of $w_i(\cdot)$ imply lower values of i's virtual types $c_i(\cdot)$, and hence the virtual types of any other firm j becomes larger relative to i's virtual type. The resulting concession and risk of being targeted of firm j could therefore decrease as $x_j(h)$ in Proposition 1 1 is increasing in $c_i(\cdot)$.$\blacksquare$

Result (i) in Proposition 4 establishes that optimal campaigns are by construction quite resilient to repression, *in spite of the fact that repression tightens the resource constraint.* This is because an optimal campaign does not always exhaust the activists' resources - it depends on the realization of the firms' type. Therefore, there is always a strict positive probability that the resource constraint is not binding. In that case, the impact of harsher repression on the equilibrium outcome may be limited to its impact channeled through the increase in the political gain from actions. By Proposition 2 above we know that this strengthens activism along all three indicators.

In a situation when the resource constraint is binding $(i.b)$ - because a large number of firms turn out to be little vulnerable - Proposition 4 establishes that increased repression contains activism by reducing the probability of action and the magnitude of the concessions. The political gains e.g., in terms of a public backlash against harsh repression and associated support for activists may still overweigh the loss. Therefore the impact on the global gain is ambiguous. Clearly, when repression brings down k to 0 (e.g., because no activist can bear the expected long jail terms) it effectively prevents activism.

Result (ii) establishes that an increase in firm specific repression (Z_i) is unambiguously detrimental to that firm. Notably, that result obtains without appealing to additional harm to the firm's reputation presumably associated with the publicity of a trial (increased vulnerability, h_i). Result (ii) reveals an unexpected property of optimal campaigns. Under conditions $(i.b)$, firms collectively benefit from the intimidation impact of the prospect of facing judicial prosecution. However, individual firms face disincentives to prosecute activists. This is because more resistant targets are relatively more attractive for disruptive actions. We thus find that competition between firms (to avoid being targeted) creates a situation where individual firms have incentives to free ride on prosecution which benefits the activists.

We revisit our lead example to illustrate our results in Proposition 4:

Example 2 contd. Let $w_i(h_i, D_i; Z) = Z + \gamma D - h_i,\ D \geq 0$ is the common damage and Z the level of repression. Then, the virtual type of firm i is given by

$$c(h_i) = h_i - (Z + \gamma D - h_i) - \frac{1 - (h_i - a)/(b - a)}{1/(b - a)} = 3h_i - (Z + \gamma D + b)\,.$$

The (common) threshold value for non-targeting is $c^{-1}(0) = h_0 = \frac{Z+\gamma D+b}{3}$, it is increasing in Z, which means that unless the resource constraint is binding, a higher value of Z induces more actions against firms and higher concessions. When $\min^k_{j\in N} h_j < \frac{Z+\gamma D+b}{3}$ an increase in Z that tightens the constraint from k to k' with $k' < k$, implying $\min^{k'}_{j\in N} h_j < \min^k_{j\in N} h_j$ reduces the probability for action and the magnitude of the concessions.

We end this section by noting that our model's predictions are consistent with the logic of the sacrificial gamble that defines civil disobedience movements i.e., the cause may benefit from (costly to the activist) repression. The effect is obtained as the equilibrium response of the optimal campaign mechanism where the only source of uncertainty is firms' cost of facing a disruptive action. Importantly, it does not require that the direct political gain from repression looms particularly large. Clearly, if the public always disapproves of civil disobedience, which is less likely to occur for non-violent movements than for sabotage or violent actions that we do not consider in this paper, repression always and unambiguously reduces equilibrium radical activism.

Our findings in Proposition 4 suggests that the choice faced by a government that aims at containing radical activism with repression tends to be corner solutions i.e., rather extreme and risky options. Either the government responds with harsh repression or it effectively concedes to disobedience. Both options are risky as harsh repression may backfire in terms of general public support for the government and conceding to disobedience may encourage radical actions.[19]

5 Optimal Campaign with Asymmetric Concessions

In the basic model we assume that the activist maximizes concessions and political gains. While we did allow for asymmetric political gain function and distributions, we assumed that concessions were equal to their monetary equivalent and thus symmetric. But firms often differ in their abatement cost, i.e. in their efficiency in the use of the concession money in terms of environmental damage reduction. So the value of a concession x from a firm i that only at high

[19] As an example, the strategy of the French government in October 2019 was to concede to Extinction Rebellion - letting activists occupy central Paris for 5 days. This significantly reduced the impact of the action.

cost decreases its emission of CO_2, is smaller than that of the same concession x from firm j that is more efficient at reducing its emission of CO_2. This means that the utility is not fully transferable, the activist is not indifferent to the identity of the firms: selling the guarantee (not to target) to firm i at price x is not the same as selling it to firm j at the same price x.

We next show how this feature can be accommodated in the basic model with only minor modifications. We shall consider a variant of the basic model that allows for asymmetry in damage control efficiency as captured by the parameters α_i which we assume are common knowledge.[20] The higher the average abatement cost of firm i, the lower its efficiency α_i. We note that there exists a natural link between the level of initial damages caused by firm i, D_i, and its average abatement cost $1/\alpha_i$: in the absence of environmental regulation, firms select the level of damage that maximizes profit. Consider two similar firms i and j, if we have $D_j > D_i$, we expect both marginal and average abatement costs to be higher for firm j than for i, and therefore $\alpha_j < \alpha_i$.

As before, given a mechanism (p, x) the (interim) expected utility of firm i when its type is $h_i \in H_i$ is given by

$$U_i(p, x; h_i) = \int_{H_{-i}} (h_i p_i(h) - x_i(h)) f_{-i}(h_{-i}) dh_{-i}, \tag{8}$$

The new feature enters into the (ex ante) expected utility of the activist

$$U_0(p, x) = \int_H \left(\sum_{i \in N} (1 - p_i(h)) w_i(h_i, D_i) + \alpha_i x_i(h) \right) f(h) dh. \tag{9}$$

so the concessions are weighted by the coefficients $\alpha_i \in R^+, i = 1, ..., n$.

A first important thing to note is that the firms' utility is the same as before, the value of the harm to the firm is unchanged so are the IR and IC constraint. Therefore, the optimal mechanism is, as before, characterized by

$$x_i(h) = p_i(h) h_i - \int_{a_i}^{h_i} p_i(s_i, h_{-i}) ds_i, \tag{10}$$

However since the activist's objective function has been modified, we now need $p : H \to [0, 1]^n$ that maximizes

$$\int_H \sum_{i \in N} \underbrace{\left(\alpha_i h_i - w_i(h_i, D_i) - \alpha_i \frac{1 - F_i(h_i)}{f_i(h_i)} \right)}_{c_i(h_i)} p_i(h) f(h) dh, \tag{11}$$

[20] Is is common to consider that abatement costs are privately known by the firms. However in our model since initial damages are common knowledge and firms are symmetric the activist can infer their abatement cost efficiency.

where the α_i enter the virtual types (see the proof in the Appendix 9) and modify the probabilities for being targeted as compared with the basic model. As usual, we need to assume regularity:

Assumption 2

For every $i \in N$, the virtual type

$$c_i(h_i) = \alpha_i h_i - w_i(h_i, D_i) - \alpha_i \frac{1 - F_i(h_i)}{f_i(h_i)} \tag{12}$$

is strictly increasing in h_i.

Equation (12) implies firm specific thresholds for targeting. The optimal mechanism discriminates among firms depending on their efficiency in damage reduction. Next, we note that $\frac{\partial c_i(h_i)}{\partial \alpha_i} \geq 0$ provided that $h_i - \frac{1-F(h_i)}{f(h_i)} > 0$ i.e., the virtual type is increasing in the efficiency of the firm in a region ranging from negative values (not smaller than the political gain) to positive ones. When it comes to the impact of the α_i on the threshold, we first note that in the absence of political gains the threshold h_{i0} given by $c_i(h_{i0}) = \alpha_i h_{i0} - \alpha_i \frac{1-F_i(h_{i0})}{f_i(h_{i0})} = 0$ does not depend on α_i. However, in the presence of political gains $w_i(h_i, D_i)$, the threshold value is affected by α_i. To see how let us first return to our leading example.

Example 3 In the example with differentiated abatement costs, the virtual types write

$$c_i(h) = \alpha_i h_i - (\gamma D - h_i) - \alpha_i(b - h_i).$$

We first note that regularity is not an issue $\frac{\partial c_i(h)}{\partial h_i} = 2\alpha_i + 1 > 0$. Next, the virtual type is increasing in α_i for $h_i > \frac{b}{2}$ and decreasing otherwise. The firm specific threshold value is

$$h_{i0} = \frac{\gamma D + \alpha_i b}{2\alpha_i + 1}$$

with $\frac{\partial h_{i0}}{\partial \alpha_i} = b - \frac{(\gamma D + \alpha_i b)}{\alpha_i + 1/2} < 0$ for $b < 2\gamma D$. So the more efficient firm i, the less often will it be targeted (the threshold is lower) and it will have to make smaller concession equal to h_{i0}. This holds when the RC is not binding. When the RC is binding there is no impact of α_i on the magnitude of the concession. For the case there would be no political gain the threshold value would be the same for all firms but the firms with higher α_i and higher virtual type would be targeted less often.

The next proposition establishes that these results hold in general.

Proposition 5 *The larger firm i's abatement efficiency α_i, (i) the lower the threshold for action against firm i, implying a lower probability that firm i will be targeted and (ii) the lower the concession to be paid to avoid being targeted.*

Proof: The threshold value for no targeting is $c_i^{-1}(0)$ defined by $\alpha_i \left(h_{i0} - \frac{1-F_i(h_{i0})}{f_i(h_{i0})} \right) - w_i(h_{i0}, D_{i0}) = 0$. Since $w_i(h_{i0}, D_{i0}) > 0$, we must have $h_{i0} - \frac{1-F_i(h_{i0})}{f_i(h_{i0})} > 0$ so the lhs of this equation increases in α_i. This in turn implies that increasing α_i reduces the value of $c_i^{-1}(0)$. A lower threshold implies a lower probability for targeting and a lower concession when not targeted. When the RC is binding there is no effect on the effective threshold. ■

As in the basic model *(ii)* is weak because in the absence of resource constraint each firm pays its threshold concession. But in case the constraint is binding, the common price is determined by the highest concession acceptable to the lowest non-targeted firm. As a consequence the concessions actually paid do not always reflect the abatement costs of the firms.

Proposition 5 shows that the standard technics we use are able to account for the fact that the activist cares about concessions in real terms e.g., reduction of CO_2 emissions while the firms care about their cost for the corresponding abatement. The result is intuitive : if the activist wants to maximize emissions abatement, she should grant the guarantee not to be targeted to the most efficient firms in exchange for a concession while targeting those firms that would not give very valuable concessions anyway. We also note that, when accounting for the correlation between D_i and α_i(see above), the result in Proposition 5 reinforces the selection bias in targeting against big polluters. They are more often targeted both because the political gain of harming increases with the damage D_i and because they are less efficient at damage reduction.

Competition in innovation for damage reduction

The mechanism induces interesting cross effects of a technical innovation that increases the abatement efficiency of a firm. Consider in particular the case when the RC is binding and the $k-$lowest firm (i) improves its efficiency to $\alpha_i' > \alpha_i$. Firm i may not be the k-lowest virtual type anymore. Another firm, j, (previously $k+1$), becomes the most vulnerable firm to be targeted while prior to the change in firm i's efficiency, firm j would have paid its concession. Another consequence is that the magnitude of the concession is now determined by firm l (previously $k+1$ lowest virtual type) i.e., it increases for everyone. Thus in this scenario, the innovation in firm $j's$ efficiency has an impact on all conceding firms and on firm l's status.

The reasoning above suggests that in a dynamic perspective optimal activism provides firms with incentives to innovate in abatement technologies in order to avoid being targeted. In addition because the mechanism relies on competition between firms to avoid being targeted, it also induces competition in innovations. This dynamic property of optimal activism benefits the activist and the environmental cause. Innovations increase the value of concessions in terms damage reductions. It also increases the threshold value defining the magnitude of concession for the case the RC is binding.

6 Extensions

In this section we briefly address some extensions.

6.1 Obtaining concessions from targeted firms

The purpose of this paper is to analyze the power of the threat of disruptive action and in equilibrium attacked firms never concede. In practice, numerous attacks are indeed not followed by concessions. One example is Monsanto that has been targeted by numerous campaigns but did not concede in any significant manner. One of the messages of the present paper is that this should not be understood as a failure of the campaign. Other more vulnerable firms do concede to avoid being attacked. However in practice, we do also see examples of campaigns that result in the targeted firm eventually conceding. In addition, disruptive actions are often undertaken with the declared objective of achieving a change in the targeted firm's behavior. This was for instance the case with the Nestlé example in the Introduction.

We below suggest how an extension of the model to two periods may suffice to obtain that some firms that have been attacked in the first period concede in the next following period. The only new ingredient that we introduce is that a firm that has been the target of an attack comes out weakened in the sense that its average vulnerability increases. This assumption seems most realistic and is e.g., suited to the Nestlé example.[21]

Assume that the first period mechanism is played out as described in section 2 resulting in a number of firms having conceded and some having been subject to a disruptive action. In the next following period, a new population of firms is selected which may include some resistant firms i.e., that have previously refused to concede and been subjected to a disruptive

[21] The first actions undertaken against Nestlé negatively impacted on its public image. When activists came back with new threats, Nestlé eventually ended up conceding.

action.[22] Each period is characterized by its own conditions which determine firms' vulnerability reflecting changing technological and market conditions. Among the factors that determines vulnerability, we include past attacks. Therefore, while firms' behavior in the previous period is not fully revealing, it carries some information. Compared with similar firms, a resistant firm is represented by a distribution shifted to the right (more vulnerable on average).

In the general model each firm has its own threshold defining the probability to be targeted (the probability that it chooses to resist).

$$c_i(h_i) = h_i - w_i(h_i, D_i) - \frac{1 - F_i(h_i)}{f_i(h_i)}. \tag{13}$$

As pointed out in section 2.2, this implies that a firm with a higher hazard rate will be targeted less often. To illustrate the point lets go back to our example and assume that after an attack at time t firm i's vulnerability is uniformly distributed over $[a, b_{t+1}]$ with $b_{t+1} > b_t$, which captures the assumption that the firm has become more vulnerable. We compare the threshold in period $t+1$ with that of period t

$$c_{i,t+1}(h_i) = 3h_i - (\gamma D + b_{t+1}) < c_{i,t}(h_i) = 3h_i - (\gamma D + b_t) .$$

We readily see that the threshold in period $t+1$ is lower than in period t implying a higher probability that the realized $h_{i,t+1}$ is above the threshold so that firm i chooses to concede.

A complete solution of the two-period model requires a proper dynamic analysis.[23] We can however identify some conditions under which the anticipation of a future campaign does not affect firms' incentives in the first period. In particular, if the expected second period concession is the same as the expected first period concession, then incurring the cost of the action in the first period remains optimal for a little vulnerable firm ($h_i < Ex_i(h)$) even if it anticipates that it may concede in the second period.[24] We note that this behavior is consistent with the observation that a vast majority of firms' tend to delay adaptation to more stringent environmental standards.

The argument above suggests that for the case an attack leaves a firm more vulnerable, there are two channels in our model that can explain why an attacked firm may concede in a next coming campaign. On the one hand, the average vulnerability increases making it more

[22] In the next sub-section, we suggest that this is indeed consistent with the activist's interest when selecting the population of firms.

[23] Because the purpose of this paper is to analyze the power of the threat of disruptive action, we chose to focus on the more tractable one period model.

[24] The cost of concession is a reduction of profit from the period the firm concede onward, i.e., it is repeated.

likely that the realized vulnerability is above the threshold so the firm chooses to concede. On the other hand the change in the distribution reduces the firm specific threshold so that the probability that the firms chooses to resist decreases for that reason too.

6.2 Selecting the population of firms

In the basic model, we assume that the population of firms is exogenous. However and in contrast with the standard auction context where firms choose whether or not to participate in the auction, no firm would ever choose to be part of an activist campaign since they only suffer losses. This choice is made by the activist. How should she choose among populations of firms?

In section 3 we noted that the result in Proposition 1 implied that the activist has an incentive to include big polluters in her campaign as it enhances the political value of actions. This result was further strengthened by Proposition 5. In this section, we are interested in the activist incentives with respect to the distribution of vulnerability of the firms.

Let us return to our lead example and consider two sectors: oil industry and banking (concession correspond to disvestment in polluting sectors). Assume that the oil producing firms are less vulnerable e.g., because when a consumer needs to tank, he must do so in the closest station. In contrast banks may be quite vulnerable because it is easy to switch to a close substitute. Let vulnerability in the oil producing sector (A population) be represented by $h_i^A \in [a, \underline{b}]$ and in the banking sector (B population) by $h_i^B \in [a, \bar{b}]$ with $\bar{b} > \underline{b}$. Computing the threshold values in the basic version of the example we get $c^{-1}(0) = h_0^A = \frac{\gamma D + \underline{b}}{3} < h_0^B = \frac{\gamma D + \bar{b}}{3}$. The magnitude of the reserve concession is higher for the more vulnerable population (B). But if banks are willing to pay h_0^B to avoid being targeted, they certainly would pay $h_0^A < h_0^B$ for the same purpose i.e., the mechanism optimal for population A is feasible for population B. So it must be that the activist is better off with population B from which she extracts higher concessions.

This suggests that since the objective of the mechanism is to maximize the sum of expected concessions the activist should (among equally polluting populations) select for her campaign, the population with the higher average expected vulnerability. We note that this is consistent with the argument of the previous sub-section i.e., that attacked firms do sometime concede albeit not in the first period. Indeed, firms that have been targeted in the past are more vulnerable on average and thus should be selected by the activist to be part of the population for her (new) campaign.

While this finding is not surprising, the question has, to the best of our knowledge, never been raised. It deserves further investigation to characterize the statistical properties of biggest interest for the selection of the population. Indeed, the study of optimal coercive mechanisms based on the threat of a harmful action has broad relevance in a variety economic contexts e.g., corruption (see KLM) but also in regulation or in taxation.[25]

Finally, but importantly, this preliminary insight i.e., that the activist should focus on a population of firms that are on average more vulnerable rather than less, is consistent with some of the empirical literature that finds that activists tend to campaign against vulnerable firms, see Discussion below (section 8). Note that this should not be confused with one of our main results (Proposition 1) that establishes that the optimal mechanism selects the least vulnerable firms for action within a given population.

6.3 Government agencies as targets

We have developed the analysis for a campaign against private firms. But public entities have also been the targets of activists. Indeed the first of Extinction Rebellion constitutive demands is the declaration of a state of climatic emergency by public entities at all levels. The model can easily be adapted to address a situation in which the activist campaigns against public entities like municipalities, regional and central government agencies.[26]

We have a population of $N = \{1, \dots, n\}$ (risk neutral) government agencies or elected public entities (in the following we refer to them as government agencies or GA for short) managing public affairs in a way that has detrimental impact on the environment. This impact can be reduced at cost for the GA's budget. In absence of any environmental effort, all agencies allocate the same budget B to secure the provision of services and infrastructures and carry out policies to maximize re-election probability. Thus, we let B play the same role for GA as π for firms. This is a short cut which views the magnitude of the available budget as the agency's resource to secure public support and realize politicians' or civil servants' private benefits.

We let each individual GA i be characterized by its 'vulnerability' i.e., the magnitude of the harm when targeted by an action. As before it is denoted $h_i \in H_i \equiv [a_i, b_i]$, where $0 \leq a_i < b_i < +\infty$. The vulnerability of a GA reflects the costs to the incumbent administration when subjected to an attack by the activist. It is measured in terms of a loss of budget (losses in private benefits and public support). The vulnerability of a GA depends on several factors.

[25] The threat of government regulation has been shown to induce self-regulation (Maxwell et al., 2000; Lyon and Maxwell, 2004: Carveras, 2007).

[26] The first of Extinction Rebellion constitutive demands is the declaration of a state of climatic emergency by public entities at all levels.

Among them we wish to mention the links between the GA and industrial interest groups. Close links raise the suspicion that it favors the interest of those groups.[27] Another source is broken promises: when a GA has been elected on a environmental program but is failing to deliver it. Yet another source of vulnerability is a short majority for the incumbent or the coalition in power. Actions against a GA can take various forms. One is the occupation of territories to prevent the realization of a new project with expected detrimental impact on the environment.[28] Another form includes symbolic actions in front of public buildings e.g., to demand the declaration of environmental emergency, etc. When a GA i is targeted by an action, detrimental information is revealed, its activity is disrupted so it cannot deliver as promised, all of which weakens its public support and reduces private benefits from power.

We assume that only GA i knows the true value of its vulnerability or "type" h_i. Public information about vulnerability is given as before by a density and a cumulative distribution function. When an agency is targeted with an action it suffers a loss so its available budget is reduced to $B - h_i$. Government agencies can make concessions $x_i(h) \in \mathbb{R}$ to avoid being targeted in which case the budget reduces to $B - x_i(h)$. The concessions corresponds to public measures and commitments to improve the ecological records of the GA at the expense of other policy measures. Clearly, in many countries there are constraints to local governments' concessions imposed by national policies. However, local leaders may distinguish themselves by promoting stricter measures.[29]

The analysis proceeds in a way similar to the one developed in the corporate context. There will be thresholds determining which GA to attack and which GA to accept concession from. In equilibrium the less vulnerable GAs will be targeted and the other make concessions.

A distinction with the corporate context is that we expect virtuous dynamic effects through another channel. Environmental improvements in one GA can inspire people in neighboring GA to request similar measures thus encouraging politicians to realize such measures to preserve or win public support. Such effects have been evidenced by Billard (2020).

[27] This is because the administration will appear to put private interests before public ones.

[28] This form of action has received a name in France "zone à défendre" or ZAD. There are plenty of examples as in Bures against the nuclear industry or in Notre Dame des Landes against the construction of a new airport.

[29] One recent example is in France where a coalition of local leaders united in 2020 to promote stricter security distances to houses for the use of pesticide in agriculture than the legally enforceable ones. (https://www.journaldesmaires.com/fr/environnement/des-maires-en-guerre-contre-les-pesticides)

6.4 When the political gain grows with vulnerability

We next demonstrate that our central result that the activist chooses to attack little vulnerable firms while receiving concession from the vulnerable ones, does not depend on the assumption that political gains decreases with vulnerability. Consider a situation where in contrast with what we assumed so far, the political gain are an increasing function of firms' vulnerability because e.g., the public is more impressed by an action that knocks down a firm. This makes the regularity condition more demanding. We need that for every $i \in N$, the virtual type

$$c_i(h_i) = h_i - w_i(h_i, D_i) - \frac{1 - F_i(h_i)}{f_i(h_i)}$$

is strictly increasing in h_i. When the political gain(loss) to the activist, $w_i(h_i, D_i)$, is increasing in h_i, regularity may not hold for all parameter combinations. Recall however that the first term corresponds to the value of the concession for the activist. When accounting for abatement efficiency, regularity is secured for large enough α_i. The α_i coefficients could also capture the activist's valuation of damage reduction.

All the formal results in the paper hold with this alternative assumption provided the regularity condition is met. Letting the political gain grow with vulnerability nevertheless induces some changes as we illustrate in our lead example.

Example 4 Assume that $w_i(h_i, D_i) = \gamma D + \beta h_i$ and $\beta < 2$, then, the virtual type of firm i is given by

$$c(h_i) = h_i - (\gamma D + \beta h_i) - \frac{1 - (h_i - a)/(b - a)}{1/(b - a)} = (2 - \beta)h_i - (\gamma D + b)\,.$$

The (common) threshold value for non-targeting is $c^{-1}(0) = h_0 = \frac{\gamma D + b}{2 - \beta}$. The optimal mechanism characterized in Proposition 1 is such that up to k firms with the k lowest types below $h_0 = \frac{\gamma D + b}{2 - \beta}$, are targeted, and the others concede and pay $\min\{h_0, \min^k_{j \in N} h_j\}$.

So compared with the original formulation of the example, the threshold is higher when the political gain grows in the vulnerability of firms. This means that more firms are expected to fall under the threshold and thus be targeted. This is not surprising because taking action is more valuable for the activist. As a consequence the resource constraint is likely to be more often binding.

7 Elements of Welfare Analysis

A first and obvious point is that the optimal mechanism studied in this paper is not Pareto improving.[30] The firms targeted by the activist's action lose an amount corresponding to their collected vulnerability $\sum_{i\in\{\text{targeted}\}} h_i$. In addition, transfers to compensate harmed firms are precluded because the threat of being harmed is at the heart of the mechanism.[31] What concessions concerns, we are, in the basic model, dealing with pure redistribution from the firms to the activist. But the value of damage reduction for the activist could also reflect progress in preventing the expected climate catastrophe which of course is much larger than the cost to the firms. Clearly, it would be preferable to achieve the same damage reduction with regulation i.e., without the cost brought forth by disruptive actions. But the very rationale for radical activism is in environmentalists' assessment that governments have effectively shown incapable (unwilling) to enact and enforce needed regulations. Radical activism is a response to that de facto constraint on available instruments for collective action in face of the ecological crisis. In that context actions are not only a necessary cost, but also a technology to 'activate' awareness and make possible welfare improving environmental policies.

As is the case with an optimal auction our optimal campaign mechanism exhibits inefficiencies. There are firms that would be willing to make concessions and the activist would prefer concessions as well (when the political gain does not outweighed the value of concession). However their types lie below the threshold for non-targeting and they will suffer the harm from a disruptive action. The mechanism is however endowed with some remarkable efficiency properties:

- The least vulnerable firm are targeted which minimizes the collected cost of actions;
- When firms differ in abatement efficiency, the mechanism secures an efficient allocation of damage reduction efforts among conceding firms.
- Competition between firms to avoid being targeted stimulates innovation in damage reduction technology.
- Resilience to repression increases the environmental gains from activism.

[30] It should be noted that private politics models generally do not address welfare issues, with the notable exception of Maxwell et al. (2000).

[31] Moreover, the collected harm is not directly comparable with the political gain to the activist.

Finally, when it comes to consumers, the impact depends very much on the form of the action. When the action against a firm informs about the damage associated with the production of its goods, this allows for better informed consumer decisions which improves expected consumers' welfare. When an action disrupts supply, this is likely to induce losses for consumers.

For conceding firms that invest in damage reduction, we may also have different cases. Either the firm passes over the new costs through higher prices in which case consumer welfare decreases. Or the firm cannot increase its prices in which case consumer welfare weakly increases.[32]

8 Discussion

In this paper, we have characterized the optimal way for an activist engaged in non-violent civil disobedience to exploit her disruptive power in order to achieve maximal gains in terms of environmental damage reduction. We find that given a population of firms, she should let firms compete to avoid attacks and target the less vulnerable big polluters with disruptive actions. The more vulnerable firms are granted a guarantee not to be harmed in exchange for a concession. When firms are asymmetric in abatement cost efficiency, the ones that are the most efficient at damage reduction make concessions and avoid action. We find that the optimal mechanism is quite resilient to police and judicial repression in line with the credo of civil disobedience. In particular, competition to avoid being targeted creates a free riding problem in prosecution for firms. The mechanism is endowed with some other nice properties. In particular it minimizes the cost of damage reduction under the constraint on available instruments. The analysis suggests interesting dynamic properties of optimal activism that deserve further investigation.

The adopted mechanism design approach is concise and normative by construction. In the most closely related paper by Baron (2016), the activist also maximizes concessions but in a more structured context. Some of our results coincide but some do not. Primarily this is due to our different approaches. Our focus is on the activist's strategy when addressing a population of firms. The strategy exploits the threat of disruptive action in a competitive setting in order to get non-targeted companies to make concessions. In contrast and as in most models involving activists, Baron (2016) models a campaign as a bilateral relationship between the activist and a firm. There is no interaction between firms. As a consequence, in his model, the activist maximizes concessions by engaging with the most vulnerable firms to induce them to concede. In addition in our setting, the level of concessions is determined by competition between firms

[32] Not all consumers may value the reduction in environmental damage.

to avoid being targeted. In Baron (2016), that level is determined by the firms expectation about being monitored (a random draw of Nature) and if monitored by their expectation to be subjected to a successful campaign (not all monitored firms are attacked and not all attacks yield concessions).

When it comes to the scarce empirical literature on the selection of target and the impact of action, it is in place to talk about mixed and contradictory evidence. The existing empirical findings are moreover quite difficult to relate to our theoretical results. There are several reasons for that. While many papers focus on the significance of vulnerability for target selection, they use diverse ways of defining and measuring it. In addition the papers do not distinguish between the choice of population and the choice within a population.[33] Finally, they are not concerned with the impact on threatened firms but only on firms subjected to attack. Briscoe and Gupta (2016) emphasize the problem posed by the lack of data on firms that are threatened into self-regulation. This makes the investigation of the "spill-over of activists' actions" difficult. Recently however, using a database of more than 8,800 French firms, Beaumais and Chiroleu-Assouline (2020) find that the intensity of NGOs' attacks on large firms in a sector spurs CSR behavior in smaller firms in the same sector.

One result common to our setting and Baron's finds some support in the empirical literature: the target of radical activist campaign tend to be well-known big polluters. Using a database of 552 private environmental activists campaigns directed against firms during the period 1988-2003, Lenox and Eesley (2009) find that activists tend to target larger firms with larger toxic emissions, but also firms belonging to certain industries, particularly polluting industries. But, on a narrower data set of 129 actions (including only boycotts and proxy votes) between 1988 and 1995, Gupta and Innes (2014) find no significant impacts of firms' emissions: environmental performance appear not to be a central driver of NGOs' decisions on targeting of environmental boycotts.

With regard to the correlation between vulnerability and the probability to be targeted, King (2008) shows that companies with a strong brand name or a high level of reputation are more likely to be targeted by activists' campaigns; for King and Soulé (2007), companies that are large, visible, and financially successful seem to be preferred targets. Gupta and Innes (2014) find that boycotts target larger firms with larger market shares, firms that are more intensively inspected for compliance with Clean Air laws and firms with strong reputations for

[33] We remind that in our setting it is optimal to select a population of vulnerable firms and within that population to target the least vulnerable firms.

CSR. But proxy actions, like shareholders initiatives, are favored against 'resistant targets' with particularly sketchy reputations for social progressivism (based on indices constructed from KLD data on 'non-environmental categories'). Most of the literature thus seems to indicate that the companies most attacked by activists are the less vulnerable companies with the exception of Lenox and Eesley (2009), for whom firms with smaller levels of cash that could be used to fight a private political campaign, are more likely to be targeted for a campaign. But they also find that activists are more likely to attack companies that have already been the target of a boycott or that are in industries that have been frequently targeted in the past (King, 2008) which could be explained by the fact that they have been more resistant than others. Interestingly, King (2008) finds that activists tend to target large and highly visible firms with positive reputation but these characteristics do not predict the likelihood of a boycott's success which is generally low. King finds that activists may "often set themselves up for failure". We note that this is consistent with our approach that targeting a firm is not primarily to obtain concession from that firm. The success of activists should not be measured by the concessions obtained from attacked firms but rather by the concessions made by threatened firms.

King (2011) studies the mechanisms of disruption that grant activists power over corporations. He highlights the impact of large media coverage on a boycott's damage to reputation. The disruptiveness of boycotts depends on the ability of boycotters to draw media attention and on the selection of ideal targets, specifically large companies as they are more sensitive to declines in their stock price and more likely to attract media coverage. Eesley et al. (2016) confirm that protests and boycotts are associated with greater media attention, whereas lawsuits and proxy votes are associated with investor perceptions of risk.

Concluding, we find that the current limited empirical literature provides some support for our thesis that disruptive actions are used to intimidate other firms into concessions rather than (primarily) to obtain concession from targeted firms. More empirical research is needed to precise the whole picture.

From a theoretical point of view, we conjecture that the study of optimal coercive mechanisms can open the way for a new approach in law and economics e.g., for the study of anti-corruption campaign. We also believe that an interesting avenue of research is to integrate networks considerations (see Billard, 2020) to the analysis of spill-over effects from activism.

References

Ayanian, A. H., & N. Tausch (2016). "How risk perception shapes collective action intentions in repressive contexts: A study of Egyptian activists during the 2013 post-coup uprising", *British Journal of Social Psychology*, 55(4), 700-721.

Aytaç, S. E., Schiumerini, L., & S. Stokes (2018) "Why do people join backlash protests? Lessons from Turkey", *Journal of Conflict Resolution*, 62(6), 1205-1228.

Baron, D. P. (2012) "The industrial organization of private politics", *Quarterly Journal of Political Science*, 7(2), 135-174.

Baron, D. P. (2016) "Self-Regulation and the Market for Activism", *Journal of Economics & Management Strategy*, 25(3), 584-607.

Baron, D.P., & D. Diermeier (2007) "Strategic Activism and Non-Market Strategy", *Journal of Economics and Management Strategy*, 6/3, 599-634.

Beaumais, O., & M. Chiroleu-Assouline (2020) "Informal versus Formal Corporate Social Responsibility: a Tale of Hidden Green Attitude", Paris School of Economics working papers series WP 2020-81.

Billard C. (2020) "Technology Contagion in Networks", WP 2020-01, FAERE Working Papers.

Briscoe, F., & A. Gupta (2016) "Social activism in and around organizations", *The Academy of Management Annals*, 10(1), 671-727.

Brown, J. M. (2022) Civil Disobedience, Sabotage, and Violence in US Environmental Activism. In *The Oxford Handbook of Comparative Environmental Politics.*

Calveras, A., Ganuza J. J., & G. Llobet (2007). "Regulation, corporate social responsibility and activism", *Journal of Economics & Management Strategy*, 16(3), 719-740.

Chenoweth, E., & M.J. Stephan (2011). *Why civil resistance works: The strategic logic of non-violent conflict.* Columbia University Press.

Culling, J. (2016),"Reviving Up your Reputation," *World Trademark Review*, December 2016-January 2017:56-59.

Daubanes, J., & J.C. Rochet (2019) "The Rise of NGO Activism", *American Economic Journal: Economic Policy*, 11(4), 183-212.

Davenport, C. (2007) "State repression and political order", *Annual Review of Political Science*, 10, 1-23.

DeNardo, J. (1985). *Power in numbers: The political strategy of protest and rebellion.* Princeton,NJ: Princeton University Press.

Edwards, P. (2021). "The politics of nonviolent mobilization: Campaigns, competition, and social movement resources", *Journal of Peace Research*, Vol. 58(5) 945–961.

Eesley, C., Decelles K. A., & M. Lenox (2016) "Through the mud or in the boardroom: Examining activist types and their strategies in targeting firms for social change", *Strategic Management Journal*, 37(12), 2425-2440.

Egorov, G., & B. Harstad (2017),"Private Politics and Public Regulation", *Review of Economic Studies*, Oxford University Press, vol. 84(4), pages 1652-1682.

Feinberg, M., Willer, R., & C. Kovacheff (2020). "The activist's dilemma: Extreme protest actions reduce popular support for social movements", *Journal of personality and social psychology*, 119(5), 1086.

Francisco, Ronald A. (1996). "Coercion and Protest: An Empirical Test in Two Democratic States", *American Journal of Political Science,* 40 (4): 1179–1204.

Galbraith, J.K. (1952), *American Capitalism: The Concept of Countervailing Power*, Houghton Mifflin.

Gans, J., & S. Kaplan (2017). *Survive and thrive: Winning against strategic threats to your business.* Dog Ear Publishing.

Gunningham, N. (2019) "Averting Climate Catastrophe: Environmental Activism, Extinction Rebellion and Coalitions of Influence", *King's Law Journal*, 1-9.

Gupta, D. K., Harinder S., & T. Sprague. (1993). "Government Coercion of Dissidents: Deterrence or Provocation?". *Journal of Conflict Resolution*, 37: 301–39.

Gupta, S., & R. Innes (2014) "Private politics and environmental management", *Journal of Environmental Economics and Management*, 68(2), 319-339.

Heyes, A., Lyon, T. P., & S. Martin (2018) "Salience Games: Private Politics when Public Attention is Limited", *Journal of Environmental Economics and Management*, 88, 396-410.

Hiatt, S. R., Grandy, J. B., & B. H. Lee (2015) "Organizational responses to public and private politics: An analysis of climate change activists and US oil and gas firms", *Organization Science*, 26(6), 1769-1786.

IPBES (2019) *Global assessment report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*, Brondízio, E. S., Settele, J., Díaz, S., Ngo, H. T. (eds). IPBES secretariat, Bonn, Germany.

IPCC (2014) *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* IPCC, Geneva, Switzerland.

IBS Center of Management Research 2011, "Greenpeace, Nestlé and the Palm Oil Controversy: Social Media Driving Change?" ICMR 911-010-1.

den Hond, F., & F. G. de Bakker (2007). "Ideologically motivated activism: How activist groups influence corporate social change activities". *Academy of Management Review*, 32(3), 901-924.

King, B. G. (2008). "A political mediation model of corporate response to social movement activism", *Administrative Science Quarterly*, 53(3), 395-421.

King, B. G. (2011). "The tactical disruptiveness of social movements: Sources of market and mediated disruption in corporate boycotts", *Social Problems*, 58(4), 491-517.

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

Koessler, F., & A. Lambert-Mogiliansky (2014) "Extortion and Political-risk Insurance", *Journal of Public Economics*, 120, 144-156.

Konrad, K. A., & S. Skaperdas (1997) "Credible Threats in Extortion", *Journal of Economic Behavior and Organization* 33, 23-39.

Konrad, K. A., & S. Skaperdas (1998) "Extortion", *Economica* 65, 461-477.

Lefkowitz, D. (2007) "On a Moral Right to Civil Disobedience", *Ethics*, 117(2), 202-233.

Lenox, M. J., & C. E. Eesley (2009) "Private Environmental Activism and the Selection and Response of Firm Targets", *Journal of Economics & Management Strategy*, 18(1), 45-73.

Lyon, T. P. (2012). *Good cop/bad cop: Environmental NGOs and their strategies toward business.* Routledge.

Lyon, T. P., & J. W. Maxwell (2004). *Corporate environmentalism and public policy.* Cambridge University Press.

McDonnell, M. H., & B. King (2013). "Keeping up appearances: Reputational threat and impression management after social movement boycotts", *Administrative Science Quarterly*, 58(3), 387-419.

Maxwell, J. W., Lyon, T. P., & S. C. Hackett (2000). "Self-regulation and social welfare: The political economy of corporate environmentalism", *The Journal of Law and Economics*, 43(2), 583-618.

Myerson, R. B. (1981) "Optimal Auction Design", *Mathematics of Operation Research,* 6(1), 58-73.

Opp, K. D., & W. Roehl (1990). "Repression, micromobilization, and political protest", *Social Forces*, 69(2), 521-547.

Rawls, J. (2009) *A Theory of Justice.* Harvard University Press.

Rawls, J. (2013). "The justification of civil disobedience", *Arguing about law*, 244-253.

Shadmehr, M., & R. Boleslavsky (2022). "International Pressure, State Repression, and the Spread of Protest", *The Journal of Politics*, 84(1), 148–165.

Siegel, D. A. (2011). "When does repression work? Collective action in social networks", *The Journal of Politics*, 73(4), 993-1010.

Smith, W. (2012) "Policing Civil Disobedience", *Political Studies*, 60(4), 826-842.

Thoreau, Esq., H.D. (1849). "Resistance to Civil Government". *Æsthetic Papers,* edited by Elizabeth P.Peabody. Boston and New York: The Editor and G.P. Putnam. pp. 189–211.

Young, L. E. (2019). "The psychology of state repression: Fear and dissent decisions in Zimbabwe", *American Political Science Review*, 113(1), 140-155.

9 Appendix

Following Myerson (1981), we below derive the formulation for the virtual type when considering asymmetric concessions closely following the proof of his Lemma 3 (p.64-66).

Let us first rewrite the activist's objective function (2):

$$\begin{aligned}
U_0(p,x) &= \int_H \left(\sum_{i\in N}(1-p_i(h))w_i(h_i,D_i)+\alpha_i x_i(h)\right) f(h)dh \\
&= \int_H \left(\sum_{i\in N}\alpha_i\left[\frac{w_i(h_i,D_i)}{\alpha_i}(1-p_i(h))+x_i(h)\right]\right) f(h)dh \\
&= \sum_{i\in N}\alpha_i\left[\int_H \left(\frac{w_i(h_i,D_i)}{\alpha_i}(1-p_i(h))+x_i(h)\right) f(h)dh\right] \\
&= \sum_{i\in N}\alpha_i\left[\int_H \frac{w_i(h_i,D_i)}{\alpha_i}f(h)dh+\int_H p_i(h)\left(h_i-\frac{w_i(h_i,D_i)}{\alpha_i}\right)f(h)dh+\int_H (x_i(h)-p_i(h)h_i)\,f(h)dh\right]
\end{aligned}$$

Equation (4.10) in Myerson is immediately transferable to our case (same use of his Lemma 2, unchanged by α_i), yielding:

$$\int_H (x_i(h)-p_i(h)h_i)\,f(h)dh = -U_i(p,x,a_i)-\int_H (1-F_i(h_i))\,p_i(h)f_{-i}(h_{-i})dh$$

Substituting this expression into $U_0(p,x)$ above gives us:

$$\begin{aligned}
U_0(p,x) &= \sum_{i\in N}\alpha_i\int_H \frac{w_i(h_i,D_i)}{\alpha_i}f(h)dh \\
&\quad+\sum_{i\in N}\alpha_i\int_H p_i(h)\left(h_i-\frac{w_i(h_i,D_i)}{\alpha_i}\right)f(h)dh \\
&\quad+\sum_{i\in N}\alpha_i\left[-U_i(p,x,a_i)-\int_H (1-F_i(h_i))\,p_i(h)f_{-i}(h_{-i})dh\right]
\end{aligned}$$

$$\begin{aligned}
U_0(p,x) &= \sum_{i\in N}\alpha_i\int_H \frac{w_i(h_i,D_i)}{\alpha_i}f(h)dh \\
&\quad+\sum_{i\in N}\alpha_i\int_H p_i(h)\left(h_i-\frac{w_i(h_i,D_i)}{\alpha_i}\right)f(h)dh \\
&\quad-\sum_{i\in N}\alpha_i U_i(p,x,a_i)-\sum_{i\in N}\alpha_i\int_H (1-F_i(h_i))\,p_i(h)f_{-i}(h_{-i})dh
\end{aligned}$$

Using the joint density function of individual vulnerabilities under the assumption that these are stochastically random variables, the last term is

$$\int_H (1-F_i(h_i))\,p_i(h)f_{-i}(h_{-i})dh = \int_H \frac{(1-F_i(h_i))}{f_i(h_i)}p_i(h)f(h)dh$$

which leads to:

$$\begin{aligned} U_0(p,x) &= \int_H \left(\sum_{i \in N} \alpha_i \left(h_i - \frac{w_i(h_i, D_i)}{\alpha_i} - \frac{1 - F_i(h_i)}{f_i(h_i)} \right) p_i(h) \right) f(h) dh \\ &+ \int_H \left(\sum_{i \in N} w_i(h_i, D_i) \right) f(h) dh - \sum_{i \in N} \alpha_i U_i(p, x, a_i). \end{aligned}$$

The term $\sum_{i \in N} w_i(h_i, D_i)$ is a constant for the activist and, from the individual rationality constraint (3), the incentive-compatibility constraint (4) and the rule of choice of concessions of the optimal mechanism (10), it follows that $\sum_{i \in N} \alpha_i U_i(p, x, a_i) = 0$, which is the best possible value for this term.

The activist's objective function can thus be simplified as maximization of the first term of the previous formula:

$$\begin{aligned} & \int_H \left(\sum_{i \in N} \alpha_i \left(h_i - \frac{w_i(h_i, D_i)}{\alpha_i} - \frac{1 - F_i(h_i)}{f_i(h_i)} \right) p_i(h) \right) f(h) dh \\ = & \int_H \left(\sum_{i \in N} \left(\alpha_i h_i - w_i(h_i, D_i) - \alpha_i \frac{1 - F_i(h_i)}{f_i(h_i)} \right) p_i(h) \right) f(h) dh \end{aligned}$$

which yields the modified virtual type (Eq. 12)

$$c_i(h_i) = \alpha_i h_i - w_i(h_i, D_i) - \alpha_i \frac{1 - F_i(h_i)}{f_i(h_i)}. \blacksquare$$