FAERE
French Association
of Environmental and Resource Economists

Working papers

Optimal Environmental Radical Activism

Mireille Chiroleu-Assouline, Ariane Lambert-Mogiliansky

WP 2020.07

Suggested citation:

M. Chiroleu-Assouline, A. Lambert-Mogiliansky (2020). Optimal Environmental Radical Activism. *FAERE Working Paper, 2020.07.*

ISSN number: **2274-5556**

www.faere.fr

Optimal Environmental Radical Activism

Mireille CHIROLEU-ASSOULINE * Ariane LAMBERT-MOGILIANSKY †

February 26, 2020

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. Disruptive actions and the threat thereof are used to force firms to concede i.e., to engage in self-regulation. We address the optimal use of scarce activist resources in face of incomplete information by looking at a general mechanism, directly adapted from Myerson's (1981) optimal auction theory. The characterization 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 for civil disobedience. We find that optimal radical activism is relatively resilient to repression. In an extension that accounts for asymmetry between firms'abatement cost, we find that the mechanism optimizes the allocation of abatment efforts and creates incentives for innovation. We discuss some other welfare properties of optimal activism.

KEYWORDS: Activism, self-regulation, mechanism design, repression.

JEL: D44, D73, D82, F21, G22, H23

1 Introduction

The failure of governments across the world to respond to the global environmental crisis in spite of the scientific consensus of its urgency (IPCC 2015, IPBES 2019) is pushing ordinary citizens to engage in various forms of activism to pressure firms and governmental institutions to take action.

*Paris School of Economics – Paris 1, Mireille.Chiroleu-Assouline@univ-paris1.fr

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

In this paper we are interested in the phenomenon of radical environmental activism of the kind of Extinction Rebellion (XR). XR is a non-violent civil disobedience movement that engages in radical actions in order to pressure firms and institutional players to take action in response to the climate emergency (Gunningham, 2019).

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 in 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, states 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 the powerful to take action, they argue that the situation is a one that legitimates civil disobedience.

In this paper we address radical environmental activism with tools developed in economics to investigate the question as to what is the best use of activist resources in view of her objectives. More precisely, we are interested in how an activist can use its disruptive power optimally to achieve the largest change in firms' (or government agencies') practice in terms of the preservation of the environment. The starting point is that firms are unwilling to change because those changes involve costs. Next, according to activists governments have for decades demonstrated a lack of effective commitment to tackle climate emergency. Therefore, activists have to resort to coercion to force them. The questions we are interested in are the following, given the activists' finite disruptive resources how should those resources be used to maximize the 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 optimal activism?

To answer these questions we develop a model where the activist threatens with disruptive actions privately informed firms with the aim of extracting concessions in terms of self-regulation measures that reduce their environmental damage. Typically, the number of firm in the population is larger than the number of firms the activist actually can harm. We are interested in how she can exploit her limited disruptive power by relying on competition between potential targets for avoiding being harmed. A first motivation for our approach is of empirical nature. 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. A second rationale is that when putting firms in competition with each other (to avoid being harmed), the activist achieves significantly larger gains from a given disruptive power than when focusing exclusively on the firms that she actually can harm.

We closely follow Myerson (1981) to characterize the optimal activist mechanism. The mechanism is optimal in the sense that it maximizes firms' collected concessions given the activist limited disruptive power. As in Koessler and Lambert-Mogiliansky (2014) who study optimal extortion, we find that the situation shows strong similarities with an auction. The activist sells promises "not to target a firm with an action" in exchange for a concession. At first sight, the setting differs from Myerson (1981) in several respects. First, the activist can sell as many promises as she wants. Second, she may be forced to sell some of them by force of a resource constraint. Third, the activist's valuation of the promises depends on the types of the firms that do not receive it, i.e., that are subjected to disruptive action. Last, utility may not be perfectly transferable because of the technology firms use to deliver concession i.e., reduce their environmental harm. Nevertheless, as in Koessler and Lambert-Mogiliansky (2014) we show that Myerson's general model is almost directly applicable. This model covers quite general situations in which the political gain (or cost) from taking action for the activist varies across firms, and firms may be heterogeneous with respect to their expected vulnerability to action and their efficiency in damage reduction.

An optimal activist mechanism is a campaign that simultaneously addresses a population of firms. The optimal mechanism is characterized by the formula for the thresholds for non-action and the equilibrium concessions. When the resource constraint does not bind the thresholds also determine the magnitude of the concessions the firms have to make to avoid action. 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 which we measure by three indicators: the probability of action, the magnitude of concession to avoid action and the global gain from activism which includes the sum of the concessions and the political gains from taking action. We show that the higher the political gain to the activist from taking action, the higher the reserve concession below which the activist chooses to engage in action against the firms and therefore the higher the risk of action. The total gain increases with the action political value both directly and indirectly through higher concessions. The second determinant of the activist's strength is the resource constraint. By definition that constraint is a limit on the number of firms that can be acted against and therefore on the risk of action. We also show that it reduces the magnitude of the concessions.

Since we are dealing with civil disobedience, the impact and the role of repression is central. Accepting to take physical and legal risks has always been a constituent part of civil disobedience. Ghandi and Martin Luther King viewed it as its fundamental strength: the readiness to sacrifice oneself to expose the illegitimacy of a policy and the immorality of repression. Siegel (2011) investigates how the extent to which repression reduces participation, and how the extent to which an angry backlash against repression increases participation, depend critically on the structure of the social network in place. He also stresses the role of emotional responses to the violent repression, such as anger and fear, that act as, respectively, incentives or disincentives to participate in response to the removal of another within one's social network. Even without considering network effects, we find that in our model where we do not assume that the political gains outweigh the cost of repression, the class of situations in which repression is counter-productive is non empty: optimal activism is quite resilient. Increased repression may lead to more actions, larger concession and larger global gain. The impact of repression exhibits some non-monotonicity: the expected impact of moderate repression is to embolden the activist but harsher repression reduces her strength. So it leaves the government facing rather extreme choices: either it effectively concedes to disobedience (implementing a strategy of 'negotiated accommodation' as advocated by Smith (2012)) or applies harsh repression.

We next consider heterogeneity in firms' abatement efficiency. The optimal mechanism is shown to deliver an efficient allocation of the global abatement cost between firms. In addition the mechanism exhibits interesting cross effects that stimulate innovation in damage reduction. Finally, we consider some extensions including the optimal selection the population of firms, the case when the potential targets are government agencies and alternative formulation of the political gain function. We develop some remarks on welfare and end with a discussion linking up with existing literature.

Related Literature. There exists as for today a very small 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. The present work is a follow-up on some of these conjectures. In particular, they note that activism shows similarity with extortion and that the activist can benefit from competition between firms. Baron (2016) is most closely related to the present work. He investigates the choices of activists who use the threat of campaign to induce firms to self-regulate. Our approaches are however quite different. Baron's model is a one of market for activists where activists monitor firms in bilateral relationships and donors fund activists. Our focus is on radical activism defined by the exploitation of the threat of using disruptive action against firms, acting as a "bad cop" coercing companies to reduce their environmental impact (Lyon, 2012).[1] We are interested in the best use of limited capacity to harm firms in order to maximize the magnitude of self-regulating measures (concessions) from firms. As in Baron, the activists do not ex-ante know the vulnerability of firms. However in Baron the activist learns the vulnerability of the firm that has been randomly assigned to her if she chooses to monitor that firm. In contrast our optimal mechanism induces the self-selection of firms which materializes in their choice of the extent of self-regulation aimed at preventing attacks. The maximization of global concession is obtained by creating competition between firms to avoid being targeted under the constraint put by limited capacity. Baron does not exploit interaction between firms.[2] In the Discussion section we address some distinctions in the predictions of our respective approaches.

A main contribution of this paper is to use the technics of mechanism design to characterize optimal environmental activism aimed at maximizing damage reduction relying on limited disruptive resource. The exercise shows how the activist can exploit the threat of disruptive action by creating competition between firms so as to induce maximal self-regulation. We also address the impact of repression and provide some support for the value of the sacrificial spirit of non-violent civil disobedience.

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

[1] Another strand of this literature considers the informational role of activists, who may complement public regulation by deciding to oppose certain hazardous projects led by the industry (Daubanes and Rochet, 2019).

[2] Heyes, Lyon and Martin (2018) develop a different framework where an industry and NGO play a salience game to influence limited public attention to social or environmental impacts. Firms compete to avoid being targeted by hiding the damage they cause.

2 Optimal Activist Mechanism

2.1 Basic Model

Firms

We have a population of firms $N = \{1, \ldots, n\}$ of (risk neutral) firms exercising some activity which has detrimental impact on the environment. This impact can be reduced at cost for the firms. In absence of any environmental effort, all firms earn the same profit π. We let each individual firm i be characterized by its "vulnerability" i.e., the value of the harm when targeted by an action (see below), the value of the environmental damage $D_i \in R^+$resulting from its activity and the abatement efficiency α_i. Vulnerability is denoted $h_i \in H_i \equiv [a_i, b_i]$, where $0 \leq a_i < b_i < \pi$. The vulnerability of a firm reflects a variety of features which combine in a idiosyncratic way. For instance, a visible brand tends to increase vulnerability to actions.[3] The existence of close substitutes makes a firm vulnerable as consumers can more easily switch to a substitute to support the activist's action. 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 harm to the firm of an action is assumed to be unknown to the activist (and the public) as it depends on the combination of a number factors including the firm's technical and economic ability to mitigate some of the harm.[4] 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.

We thus assume that only firm i knows the true value of its vulnerability or "type" h_i. Public information about vulnerability is given by $f_i : H_i \to \mathbb{R}_+$ 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

$$\pi - h_i.$$

[3]This is because consumer can easily identify its product and reduce their consumption when learning thanks to the action that the firm is responsible for environmental damages.

[4]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 unit cost for abatement $\alpha_i = \alpha$ for all i, and it is normalized $\alpha = 1$. We relax this assumption in section 5.

Activist

There is one (risk neutral) activist organization, referred to as "she" the activist. Her objective is to protect the environment by forcing firms to take measures (make concessions) to reduce the detrimental impact of their activity on the environment e.g., to abate their CO_2 emissions. The activist engages in a campaign against a population of firms. She has the capacity to carry out disruptive actions $a \in \{0,1\}$ against firms. The threat of such actions is aimed at inducing concessions. 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 with an action, the activist receives $w(h_i, D_i; Z) \in R$. 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 function $w(.)$ is parametrized by Z , the (exogenous) level of repression. We assume that $\frac{\partial w}{\partial Z} > 0$ 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 activist. For most part of the paper we take Z as given so we write $w(h_i, D_i; Z) = w(h_i, D_i)$. In section 4, we investigate the impact of repression on the equilibrium solution.

Action

The term action is used to refer to a collection of measures that reduces the firms's ability to earn profit in a broad sense. Actions cover symbolic disruptions like die-ins, tagging but also boycott, blockage and sabotage. They are aimed at affecting the firm's reputation, the demand for its products or directly disrupting its production activity. While the action is a binary variable $a \in \{0,1\}$, its impact on the firm and on the activist's objective depends on the type of the firm. The cost to the firm is the harm h_i and its political value for the activist is captured by $w(h_i, D_i) \in \mathbb{R}$. For most part of the paper, we assume that $\frac{\partial w(h_i,D_i)}{\partial h_i} \leq 0$ and $\frac{\partial w(h_i,D_i)}{\partial D_i} \geq 0$ so the political gain is continuously decreasing in the vulnerability of the firm and increasing in its environmental damage. 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 conciliatory firms.[5] On the other side, the political gain does not give any weight to possible change in practice following the action. In an extension where we consider the case when $\frac{\partial w(h_i,D_i)}{\partial h_i} \geq 0$ which we interpret as the political gain associated with the action yielding an impressive impact (high harm).

Mechanism

The objective of the mechanism is for the activist to use her disruptive power to achieve maximal improvement in firms' production practice or equivalently to minimize the total harm to the environment from a population of firms. A mechanism determines which firms are targeted (if any) and how much firms "pay" to avoid being targeted. The payment is in terms of concessions to environmental demands. We shall be interested in the risk for firms of being targeted by a disruptive action, the magnitude of firms' abatements (concessions) and the global gain (sum of the value of conceded abatements plus political gain from action). We adopt a mechanism design approach, assuming that firms' bargaining power is minimized.[6] We also assume that the activist does not know the values that firms attach to avoiding actions (their vulnerability) or the maximal concession that each firm is willing to make to avoid being targeted. If the activist perfectly knew each firm's vulnerability, he would be able to obtain the maximal concession from each of them by threatening any non-obedient firm with action. 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" 'promises not to target' to firms in exchange for concessions. The activist aims at maximizing her revenue, here the sum 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.[7]

The optimal mechanism that we characterize is obtained by slightly adapting the design of an optimal auction in Myerson (1981). In our setting the differences are very close to those in Koessler et Lambert-Mogiliansky (2014, hereafter KLM). We formulate them as follows: (i) The seller (activist) sells several (homogeneous) goods (a 'good" in our model being a guarantee of *not* targeting the firm) and each buyer (firm) needs one good at most (each firm demands

[5] In the extension section we show that our results do not depend on this assumption.

[6] The growing concern of the business world for eco activism suggests that activists already have some significant bargaining power which is most likely to grow with the climate crisis.

[7] Given the constraint on resource facing activist groups, there is often close cooperation between them. They are viewed as a single player with respect to our concern here.

at most one guarantee not to be targeted); (ii) The seller has some constraint on the minimal number of goods he should sell (i.e., the activist is capacity constrained with respect to carrying out actions); (iii) The seller's valuation for a good depends on the types of the buyers who do not receive the good (i.e., the value of the concessions may depend on the characteristics of the firms that are targeted); (iv) 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 is similar to the one of an optimal auction mechanism.

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 $h = (h_1, \dots h_n)$, $p_i(h)$ 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 ($\alpha_i = \alpha = 1, \forall i \in N$). 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

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

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. This is the coercive feature of the mechanism which we have in common with the extortion set-up in KLM. Condition (4) means that firms have no incentive 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. The resource constraint allows introducing competition between firms in a tractable way. We below characterize the activist's global gain when he optimally exploits competition.

2.2 Feasible and Optimal Mechanisms

The objective of the activist is to choose 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

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

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

As is standard in the literature, we make an assumption of regularity that secures that state-by-state optimization of the program above implies that $p_i(h_i, h_{-i})$ is increasing in h_i :

Assumption 1

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

$$c_i(h_i) = h_i - w(h_i, D_i) - \frac{1 - F_i(h_i)}{f_i(h_i)} \tag{7}$$

is strictly increasing in h_i.

The function $w(h_i, D_i)$ 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., $\partial w(h_i, D_i)/\partial h_i \leq 0$. This secures that $c_i(h_i) = h_i - w(h_i, D_i) - \frac{1-F_i(h_i)}{f_i(h_i)}$ is strictly increasing in h_i which 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.[8] We immediately 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*

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

where the virtual type $c_i(h)$ *of firm* i *is given by (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. The concession of firm* i *to the activist demand is given by* $x_i(h)$ *defined in (6).*

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 $\min_i^k x_i = x_k$. Let

$$y_i(h_{-i}) = \min\{s_i \in H_i : c_i(s_i) \geq 0 \text{ or } c_i(s_i) \geq \min_{j \neq i}{}^k c_j(h_j)\}, \tag{8}$$

be the smallest type of firm i such that firm i is not targeted when other firms' types are given by h_{-i}. The optimal mechanism can therefore be reformulated as follows:

$$p_i(h) = \begin{cases} 1 & \text{if } h_i > y_i(h_{-i}), \\ 0 & \text{if } h_i < y_i(h_{-i}), \end{cases} \quad \text{and} \quad x_i(h) = \begin{cases} y_i(h_{-i}) & \text{if } h_i > y_i(h_{-i}), \\ 0 & \text{if } h_i < y_i(h_{-i}). \end{cases} \tag{9}$$

For each firm i the optimal mechanism involves a (possibly firm specific) threshold value $c_i^{-1}(0)$ for no-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

[8] The public may learn about a targeted firm's vulnerability from the market's reaction to the action.

to grant a guarantee of no-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 $y_i(h_{-i})$ given by Equation (8), 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.

Remark

We note that truth revelation is the only equilibrium strategy for firms. Firms can calculate the threshold in equilibrium so they also know whether their vulnerability is below that threshold. However, the presence of the resource constraint implies that whether or not a firm is targeted always depends on the profile of all firms so truth revelation is the only equilibrium. This in turn is important to our mechanism since the equilibrium concession when the RC is binding is $\min^k_{j\neq i} c_j(h_j)$ *which could not be computed in a pooling equilibrium.*

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-expropriation. 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 below h_0 are targeted and concede nothing, and the others are not targeted and make the same concession of value: $\min^k_{j\in N} h_j$, 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 ($\min^k_{j\in N} h_j$) may be strictly lower than the activist's "reserve price" (h_0).[9]

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 $x = \min\{h_0, \min^k_{j\in N} o_j(h_j)\}$. 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

[9] 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.

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 $w(h_i, D_i) = (\gamma D - h_i)$, 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-expropriation 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 $\min\{h_0, \min^k_{j \in N} h_j\}$.

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 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 $w(y, D_i) = w(y, D_j)$, 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.

3 The Strength of Radical Environmental Activism

In this subsection we are interested in the strength of radical activism in terms of inducing changes in firms' practice and raising awareness for the environmental cause (political gain). We measure that strength with three indicators. The first is the *probability of action* implied by the threshold values for no-action $c_i^{-1}(0)$. Recall from Proposition 1 that firm i is never targeted if $c_i(h_i) > 0$, i.e., $h_i > c_i^{-1}(0)$; otherwise, if $h_i < c_i^{-1}(0)$, firm i is targeted whenever the activist resource constraint is not binding. Hence, the threshold values determine the ex-ante probability for action. The higher the thresholds the more likely the firms' types are lower than the thresholds and therefore the more likely they could be targeted. The second indicator of the strength of activism is the *magnitude of the concessions* that firms make to avoid being targeted. The larger the concessions, the stronger the activist. Finally, a third indicator is the *global gain*, which comes both from concessions and actions. The larger the global gain from acting or threatening to act, the larger the strength of activism.

3.1 The Political Gain from Actions

A determinant of the virtual type of firm i, and hence of the threshold value $c_i^{-1}(0)$, is the function $w(\cdot) : H_i \to R$ that determines the value to the activist for targeting firm i with an action as a function of firm i's type $h_i \in H_i$ and the damage D_i given the level of repression. In Example 1 we assumed $w(h_i, D_i) = \gamma D_i - h_i$ so the threshold value $h_{i0} = \frac{\gamma D_i + b}{3}$ is increasing in D_i. This means that the ex-ante probability that firm i is targeted is increasing with the perceived environmental damage caused by that firm, in the example captured by γD_i. So is the concession (equal to h_{i0}) made by the types of firm i that are not targeted. When the RC is binding concessions are constant in D_i, and equal to $h_{k0} = \frac{\gamma D_k + b}{3}$ the threshold for the k-th highest firm among the set of targeted firms. The next proposition shows 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.

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 $w(\cdot)$.*

Proof. Consider a political gain function $\widetilde{w}(\cdot)$ of some firm $i \in N$ such that $\widetilde{w}(\cdot) > w(h_i, D_i)$ for every $h_i \in H_i$. 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 $\widetilde{w}(\cdot)$ than with $w(\cdot)$. To show that the global gain of the activist is also higher with $\widetilde{w}(\cdot)$ than with $w(\cdot)$ it suffices to notice that the optimal mechanism with $w(\cdot)$ is also feasible with $\widetilde{w}(\cdot)$ 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 $\widetilde{w}(\cdot)$ 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 larger the total expected gain to the activist from the campaign. Since the political gain is an increasing function of the damages D_i, 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 also note that a decrease in the political value of the action against some firm j - for instance because j adapts to become less vulnerable (or it reduces its damage) - may be detrimental for another firm i when the RC is binding and i's type is below its threshold for non-targeting ($h_i < c_i^{-1}(0)$). Indeed smaller values of $w_j(\cdot)$ imply higher values of j's virtual types $c_j(\cdot)$, and hence the virtual types of any other firm i becomes smaller relative to j's virtual type. The resulting concession and risk of being targeted of firm i could therefore increase as $y_i(h_{-i})$ of Equation (8) increases with $c_j(\cdot)$.

3.2 The Resource Constraint

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. In the Example the k lowest types below $h_0 = \frac{D+b}{3}$ are expropriated, and the others pay $\min\{h_0, \min_{j \in N}^k h_j\}$. Hence, the weaker the RC (the larger k) the larger the probability of targeting and the concessions made by each firm. Since k only appears as a constraint in the activist's optimization program (through Equation (5)), his revenue is also increasing in k. These results extend to the general case and we have:

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 an action.*

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

We view the resource constraint of radical activism as primarily capturing the extent of radical engagement in the population. It is an expression of the time, the energy activists are willing to surrender to the implementation of actions as well as the amount of physical and judicial risk they are willing to take. The more radical the activists and the more numerous they are, the less binding RC i.e., the larger k. It may also capture donors' financial support to the activist organization.[10]

We conclude that the strength of activism, in terms of the three indicators we use, is unambiguously increasing with the political value of action and with the number of firms that she has the resource to target with an action.

4 Optimal Activism and Repression

Since we are dealing with civil disobedience, actions violate laws and thus give rise to police and judicial repression. We assume that police and judicial repression aims at securing the rule of law and maintaining public order (Smith, 2012). One of the objectives is to intimidate citizens by the prospect of facing costs in terms of physical harm, fines and jail terms. However, the very point of civil disobedience is to confront the state. When the state represses activists, they are given a chance to advocate for their cause by denouncing the use of force and laws to counter illegal but altruistic action while serious deeds from profit seeking firms and/or captured government agencies are left unpunished. One example is a recent action in France amounting to taking down the portrait of President Macron in town halls around the country. This action formally qualifies as "stealing in organized group" which is punishable up to 5 years in prison. The objective of the action is to raise awareness and denounce President Macron's inaction with respect to environmental issues including his non-compliance with the Paris agreement. The activists hope to create public demand for the necessary change in policy.

As argued in the Introduction, accepting to take physical and legal risks has always been a constituent part of civil disobedience. Recently, in April and October 2019, Extinction Rebellion

[10] It is quite remarkable that some environmental organizations like Extinction Rebellion France function nearly fully without external funding.

in London openly aimed at having as many people arrested as possible. Repression and judicial prosecution serves the activist's objective because it brings publicity to the subject matter of environmental crisis. It also serves the cause by highlighting the senselessness of treating as ordinary offenders people who take high personal risks to promote better environmental protection. The gamble of civil disobedience is that taking the cost of exposing this inadequacy in trial will increase further the support from the population and counter the goal of intimidation. In addition, exposing the activist body and freedom to state violence is a signal of the depth of the activist conviction.

Our model allows studying some impact of increased repression. We remind that $w(h_i, D_i) = w(h_i, D_i; Z)$ with $\frac{\partial w}{\partial Z} > 0$. This assumption captures the fact that 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. Presumably the use of violence on non-violent altruistic people tends to outrage more people the more excessive the force. This in turn broadens the support for activists. But repression also affects the resource constraint. 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. Relying on our results in Propositions 2 and 3 above, Proposition 4 characterizes the impact of repression on optimal environmental activism.

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

(i) 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.

(ii) 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.

The proof follows immediately from Propositions 2 and 3 above.

The result in Proposition 4 establishes that optimal activism is by construction quite resilient to repression. In spite of the assumption that repression tightens the resource constraint. This is because optimal activism does not always exhaust its 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 sole effect of harsher repression is to increase the political gain from action against any firm i, by Proposition 2 above we know that this strengthens activism.

In a situation when the resource constraint is binding - because a large number of firms turn out to be little vulnerable - increased repression contains activism by reducing the probability of action and the magnitude of the concessions. However, the political gains e.g., in terms of a public backlash may still overweigh the loss. Therefore the impact on the global gain is ambiguous. Clearly when repression brings down k to 0 (e.g., by putting all activists in jail) it effectively prevents activism.

So we find that the model is consistent with the sacrificial gamble of civil disobedience i.e., the cause may benefit from (costly to the activist) repression. The effect is obtained as the equilibrium response of the optimal mechanism where the only source of uncertainty is firms' vulnerability i.e., its cost of facing a disruptive action. Importantly, it does not require that the direct political gain from repression looms particularly large.

We thus find that when dealing with optimal activism, the expected impact of repression is non-monotonous. The optimal choice for a government that aims at containing its strength 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 weakens the strength of the activist (by reducing the political value) but may also encourage a radicalization of actions.[11]

Example 2 contd. Let $w(h_i, D_i; Z) = Z + \gamma D - h_i,\ D \geq 0$ 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-expropriation 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.

[11] The strategy to the French government in October 2019 was to concede to Extinction Rebellion - letting activists occupy central Paris for 5 days. This significantly reduced the visibility of the action.

5 Optimal Activism 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 it is more realistic to assume that firms differ in their abatement cost, i.e. of 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 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.[12] The higher the marginal abatement cost of firm i, and 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 marginal abatement cost α_i: in the absence of environmental regulation, firms select the level of damage that maximizes profit. Condider two firms i and j, if we have $D_j > D_i$, we expect $\alpha_j < \alpha_i$.

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

$$U_i(p, x; h_i) = \int_{H_{-i}} (h_i p_i(h) - x_i(h)) f_{-i}(h_{-i}) dh_{-i}, \tag{10}$$

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

$$U_0(p, x) = \int_H \left(\sum_{i \in N} (1 - p_i(h)) w(h_i, D_i) + \alpha_i x_i(h) \right) f(h) dh. \tag{11}$$

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

[12] Is is common to consider that abatement costs are privately known by the firms. As initial damages are common knowledge, the activist should however be able to make valuable inference about their abatement cost efficiency.

mechanism is, as before, characterized by

$$x_i(h) = p_i(h)h_i - \int_{a_i}^{h_i} p_i(s_i, h_{-i})ds_i, \tag{12}$$

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

$$\int_H \sum_{i \in N} \underbrace{\left(\alpha_i h_i - w(h_i, D_i) - \alpha_i \frac{1 - F_i(h_i)}{f_i(h_i)}\right)}_{c_i(h_i)} p_i(h)f(h)dh, \tag{13}$$

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

$$c_i(h_i) = \alpha_i h_i - w(h_i, D_i) - \alpha_i \frac{1 - F_i(h_i)}{f_i(h_i)} \tag{14}$$

is strictly increasing in h_i.

Equation (14) 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 $w(h_i, D_i)$, 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) a lower (the) concession to be paid to avoid being targeted.*

Proof: *The threshold value for no targeting is $c_i^{-1}(0)$ is defined by $\alpha_i \left(h_{i0} - \frac{1-F_i(h_{i0})}{f_i(h_{i0})} \right) - w(h_{i0}, D_{i0}) = 0$, and the rhs of the equation rises in α_i since $h_{i0} - \frac{1-F_i(h_{i0})}{f_i(h_{i0})} > 0$. By regularity again this implies that h_{i0} decreases with α_i. Hence 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., CO_2 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 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 3.1 we noted that the result in Proposition 2 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 $h_i^B \in [a, \overline{b}]$ with $\overline{b} > \underline{b}$. Computing the threshold values in the basic version of the example we get $c^{-1}(0) = h_0^A = \frac{\gamma D + \underline{b}}{3} < h_0^B = \frac{\gamma D + \overline{b}}{3}$so the magnitude of the reserve concession is higher for the more vulnerable firms (B population). 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. So it must be that the activist is better off with the B population from which he extracts higher concessions.

This suggests that to the extent that the activist chooses the population for her campaign, she should (among equally polluting populations) go for a population with the higher average expected vulnerability. Indeed since the objective of the mechanism is to maximize the sum of expected concessions, the higher the average expected vulnerability, the higher the expected sum.

While this finding is not surprising, the question has, to the best of our knowledge, never been raised. It deserves further investigation to precise the statistical characteristics of largest interest for selection. In particular, this is because optimal coercive mechanisms have broader relevance (see e.g., Koessler and Lambert-Mogiliansky 2014). We also note that this prelimanry insight in favor of vulnerable firms is consistent with the empirical literature that found that activists tend to campaign against more vulnerable firms see Discussion below (section 8).

6.2 Government agencies as targets

We have developed the analysis for a campaign against private firms. However, it can easily be adapted to address the situation in which the activist focuses on public entities like municipalities, regional and central government agencies.

The model would then be formulated as follows. We have a population of $N = \{1, \ldots, n\}$ (risk neutral) government agencies or elected public entities (in the following we refer to them as government agencies of 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 carry out policies, secure the provision of services and infrastructures, maximize re-election probability, etc.... We let each individual GA i be characterized by its "vulnerability" i.e., the value 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 (political) costs to the incumbent administration when unable to use the whole of its budget to honour its commitments. One source could be if the GA has close link with industrial interest groups.[13] Another source of vulnerability is when a GA has been elected on a environmental program but is failing to deliver it. Yet another is a short

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

majority for the party or coalition in power. An action against a GA can take the form of an occupation of territories to prevent a new project with detrimental impact on the environment, symbolic actions in front of public buildings to request the declaration of a state 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. 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 payoff is $B - h_i$. Agencies can make concessions $x_i(h) \in \mathbb{R}$ to avoid being targeted so the budget is $B - x_i(h)$. The concessions corresponds to public measures and commitments to improve the ecological records of the GA.

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 effect 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).

6.3 When the political gain grows 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.[14] This makes the regularity condition more demanding. We need that for every $i \in N$, the virtual type

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

is strictly increasing in h_i. When the political gain(loss) to the activist, $w(h_i, D_i)$, 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 efficiency in abatement costs, regularity is secured for large enough α_i. The α_i coefficients could also capture the activist's (common) subjective valuation of damage reduction.

[14] In contrast a model where targeted firms make concessions valued by the activist would be have to be very different.

All the formal results in the paper hold with this alternative assumption provided the regularity condition is met. To see some implications, we return to our lead example.

Example 4 Assume that $w(h_i, D_i) = \gamma D + \beta h_i$, 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-expropriation 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 $\min\{h_0, \min_{j \in N}^k h_j\}$.

So compared with the original formulation of the example, the threshold is higher when the political gain grows in the vulnerability of firms. This is not surprising because taking action is more valuable for the activist. This also means that 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. When concessions are fully transferable, the profit of companies is reduced, for non-targeted firms by the amount that is gained by the activist $\sum_{i \in \{\text{non targeted}\}} x_i$. In the basic model concessions are pure redistribution from the firms to the activist. 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 for the firms. For those targeted by the activist's action, they lose an amount corresponding to their total vulnerability $\sum_{i \in \{\text{targeted}\}} h_i$ which is not directly comparable with the political value. Transfers to compensate harmed firms are precluded because the threat of being harmed is at the heart of the mechanism. Clearly, it would be preferable to achieve the same damage reduction with regulation without the cost brought forth by disruptive actions. But the very reason for radical activism is that governments have effectively shown incapable (unwilling) to enact and enforce needed regulations in time. 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 action are not only a necessary cost, but also a technology to "activate" awareness and make possible welfare improving environmental actions.

It is worth noting here that, in our model, attacking a corporation does not provide any environmental gain but only a political gain to the activist.[15] Accounting for possible concessions from targeted firms following an action would only increase the welfare value of the mechanism.

Next, similarly to an optimal auction our optimal activist 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 type lies below the threshold for non-targeting and they with suffer the harm from the disruptive action. The mechanism is also endowed with some remarkable efficiency properties:

- The least vulnerable firm are targeted which minimizes the cost of actions to the firms;
- When firms differ in abatement cost 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.

Finally, when it comes to consumers, no clear cut evaluation can be produced but we develop a few arguments of some relevance. When the harm inflicted to the firm takes the form of a loss of reputation, the action informs consumers about the damage associated with the production of the good. That allows for more informed decisions reflected in a alternative consumption basket including a reduction of the demand for the targeted firm's products. When an action disrupts supply, this is likely to induce losses for consumers.

For conceding firms that invest in damage reduction, we may have different cases. Either the firm passes over the new costs through higher prices in which case consumer welfare decreases. Or the firm cannot do so because of say competition and they suffer reduced profit see above.

8 Discussion

In this paper, we characterized the optimal way for an activist engaged in non-violent civil disobedience to exploit her disruptive power in order to achieve the maximal gains in terms of environmental damage reduction. We find that given a population of firms, she should let firms

[15] In our static context such concession are not justified.

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 their concessions. 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. The mechanism is endowed with some other nice properties, in particular it minimizes the harm needed to achieve global damage reduction. 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 to get non-targeted companies to make concessions. In contrast and as in most models involving activists, Baron (2016) assumes that the activist maximizes concessions from the companies under attack and there is no interaction between firms. As a result, in his model, the most vulnerable firms are targeted to induce them to concede. In addition in our setting, the level of concessions is determined by competition between firms to avoid being targeted. In Baron that level is determined by the firms expectation about being monitored (a random draw of Nature) and if monitored their expectation to be subject 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.[16] Finally, they are not concerned with the impact on threatened firms but only on firms subject to attack. Briscoe and Gupta (2016) emphasize the issue posed by lacking data on firms that are just threatened which makes the investigation of the spillover of activists actions difficult. Recently however, using a database of more than 9000 French firms, Beaumais and Chiroleu-Assouline (2020) find that the intensity

[16] 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.

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 dataset of 129 actions (including only boycotts and proxy votes) between 1988 and 1995, Gupta and Innes (2014) find no significant impacts of firm emissions : environmental performance appear not to be a central driver of NGOs' decisions on targeting of environmental boycotts.

As 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 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 results that the targeted firms are not the one that are vulnerable and the success of activists should not be measured by their concessions but by the concessions made by threatened firms.

Consistently with our modelling of the role of the political gain King (2011) studies the mechanisms of disruption that grant activists power over corporations ; he highlight 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 target organizations. 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.

We did not expect that activists actually use an optimal activist mechanism, there are however empirical indications that actions are used to intimidate other firms into concessions rather than (primarily) to obtain concession from targeted firms. This is in line with the thesis that threat and spill-over effects plays a central role in activist campaigning. As suggested above more empirical research is needed to precise the whole picture. Finally, from a more theoretical point of view, we 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

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

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

Beaumais O. and M. Chiroleu-Assouline (2020) "Informal versus Formal Corporate Social Responsibility: a Tale of Hidden Green Attitude", mimeo.

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

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

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

Eesley, C., Decelles, K. A., and 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.

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

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

Hiatt, S. R., Grandy, J. B., and 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.

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

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.

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

Koku, P. S., Akhigbe, A., and T. M. Springer (1997) “The financial impact of boycotts and threats of boycott”, *Journal of Business Research*, 40(1), 15-20.

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

Lenox, M. J. and 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.

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

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

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

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

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.

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):

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

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:

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

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

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:

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

The term $\sum_{i\in N} w(h_i,D_i)$ 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 (12), 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:

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

which yields the modified virtual type (Eq. 14)

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