The Informational Role of Non-Governmental Organizations to Induce Self-Regulation: Cheering the Leaders or Booing the Laggards?

Matthieu Glachant*, Gabrielle Moineville†

20th July 2012

Abstract

Non-governmental organizations (NGOs) play a key role in creating incentives for firms to develop a Corporate Social Responsibility (CSR) policy by disclosing publicly self-regulatory corporate efforts. Their informational behavior is heterogeneous: Some NGOs mostly disclose information on firms that do not behave responsibly (e.g., Greenpeace). Others are specialized in revealing firms that are socially or environmentally responsible (e.g., the Marine Stewardship Council). We develop a model describing the interactions between a NGO, a continuum of firms and a representative stakeholder to explain what drives the NGO communication choice and its impact on the level of self-regulation.

We show that the NGO specializes in equilibrium: depending on the size of its budget, it either chooses to cheer the leaders or to boo the laggards. We extend the model to the case with multiple NGOs. We also introduce the possibility of NGO-corporate partnerships and derive policy implications.

Keywords: Non governmental organisations, Corporate Social Responsibility, incentives, self regulation.

JEL classification: D82.

*MINES ParisTech, CERNA Corresponding author: 60 bd St Michel, 75006 Paris. Email: glachant@ensmp.fr. Tel: +33(0)140519229. Fax: +33(0)140519145.

†MINES ParisTech, CERNA.

1 Introduction

Firms frequently take socially and environmentally friendly actions not required by law, thereby privately providing public goods, or voluntarily internalizing externalities. For example, the oil company Total finances the rehabilitation of the Anggana mangrove forest in the Mahakam Delta in Indonesia. The chemical and pharmaceutical company Bayer develops programs to promote employees gender diversity. Komatsu[1] seeks to address employee mental health concerns. The Coca-Cola Company has implemented a comprehensive corporate policy, including quantified objectives regarding packaging recycling, water stewardship, energy conservation, etc. Albeit ill-defined, the concept of Corporate Social Responsibility (CSR) is a convenient umbrella term to designate such activities.

As self-regulatory efforts are arguably costly, the reasons why corporations are willing to self-regulate have been extensively explored in the economic literature. Many works stress the fact that some stakeholders are willing to reward CSR leading firms or, alternatively, to punish laggards. Some consumers may accept to buy their products at a higher price or boycott 'dirty' corporations (Bagnoli and Watts, 2003; Arora and Gangopadhyay, 1995). Employees may work in a more productive way, or they may accept lower wages in environmentally- or socially-responsible firms (Brekke and Nyborg, 2008). Socially responsible investors can help leading firms to reduce the cost of capital (Heinkel and al., 2001).

The problem is that self-regulation is a credence good whose benefit is impossible for an individual consumer, employee or investor to ascertain. It provides the rationale for the existence of specialized actors that have sufficient resources to observe self-regulatory activities of individual firms and to convey this information to the stakeholders. Without such a monitoring, stakeholders are not able to reward individual efforts or punish laziness, thereby reducing self-regulation in the first place. This paper focuses on NGOs. Beside labeling organizations or rating agencies, they now play a key role in monitoring and communicating over CSR corporate efforts (Feddersen and Gilligan, 2001).

The starting point of the paper is the observation that, in the real world, NGOs have heterogeneous informational behaviors. Some NGOs are specialized in the transmission of good news: they certify that a firm – or a product – is socially or environmentally responsible. An illustration is the Marine Stewardship Council certification which rewards

[1] A multinational corporation that manufactures construction, mining, and military equipment.

sustainable fishing by certifying about 12 percent of the world catch intended to human consumption. Other NGOs mostly send bad news: Greenpeace is the emblematic example of an organization that fiercely conducts campaigns against firms that they do not deem to behave responsibly. Lie in between neutral NGOs which send both and bad news, as for example WWF which alternates between cooperative and offensive behavior toward firms depending on their compliance with WWF demands. In the following, we refer to the NGOs which cheer CSR leaders as good cops whereas bad cops are NGOs which boo the laggards. Note that this heterogeneity may reflect selective monitoring - the NGO focuses on the observation of positive activities - or selective disclosure - the NGO chooses to communicate only some of its monitoring results.

In this paper we seek to understand when and why NGOs choose between sending good news or bad news. Another objective of the paper is to study how differences in NGO communication shape the firms' incentives to self regulate.

To answer these questions, we build a simple model with asymmetric information that describes the interaction between a continuum of firms that may individually decide to self-regulate ($a = 1$) or not ($a = 0$), a stakeholder who is willing to pay for self-regulation and a NGO that seeks to maximize the average level of self-regulation under a budget constraint and that provides the stakeholder with imperfect information on a. We also extend the analysis to a setting with multiple NGOs.

Information disclosure by the NGO is asymmetric. More specifically, the NGO observes action and conveys this information to the stakeholder with a probability g when $a = 1$ (good news) and with a probability b when $a = 0$ (bad news). The parameters g and b are endogenously selected by the NGO which can decide to be neutral ($g = b$), a good cop ($g > b$) or a bad cop ($g < b$). We use the model to investigate how the NGO disclosure probabilities (g, b) shape firms' behavior and then to characterize the probabilities chosen in equilibrium by the NGO.

We show that the NGO decides to specialize: If its budget is limited, it chooses to be a good cop ($g > 0$ and $b = 0$). When more resources are available, it opts for being a bad cop ($b > 0$ and $g = 0$). The intuition is that it is more effective to go against the stakeholder expectations: A low budget limits disclosure probabilities, implying low firms' incentives to self-regulate. Being aware of this, the stakeholder holds pessimistic beliefs on the expected level of self-regulation. By filtering good firms out, a good cop improves the belief about firms whom she receives no news about.

We then extend the analysis by assuming multiple NGOs which non cooperatively select their technology. It gives birth to a coordination problem whereby NGO choices are biased towards friendliness. We also study the potential role of a welfare-maximizing regulator and we introduce the opportunity for firms to form partnerships with NGOs.

The economic literature on self-regulation and Corporate Social Responsibility has already explored the role of NGOs and activists in triggering corporate socially- and environmentally-responsible investments. For instance, "Good Cop/Bad Cop" is the title of a recent book edited by Tom Lyon (2010) which contrasts the heterogeneity of NGO strategies towards business. But none of the contributions in this book develops a theoretical analysis of what drives such strategies. There exist theoretical works which specifically deal with bad cops. For instance, Baron (2001) and Lyon and Maxwell (2011) investigate the impact of NGOs which are able to penalize firms which do not make sufficient environmental or social efforts. Others deal with good cops, and in particular with the role of NGOs in product labeling (ecolabels, fair trade, etc.). For example, Bottega and De Freitas (2009) compare two certifiers: a NGO and a for-profit organization. None of these papers deal with the different types of NGO in a unified framework as we do in this paper.

The paper is structured as follows. We present the base model with a single NGO in the next section. We extend to the case with multiple NGOs and we address regulation issues in the third section. In the fourth section, we analyze the role of corporate-NGO partnership. We summarize the results and conclude in the last section.

2 The base model

2.1 Assumptions

We consider a continuum of firms that may self-regulate or not. Each firm makes a binary decision: $a = 1$ if it self-regulates, $a = 0$ if it carries on with business-as-usual. Self-regulation costs c, which is heterogeneous across firms and uniformly distributed over $[0; 1]$. Assuming that c is positive is in line with the very nature of self-regulation: it consists in improving social and environmental performance beyond the business-as-usual. That is, once all actions at negative cost have been implemented. The assumption that $c \leq 1$ is irrelevant qualitatively. It simplifies the notations. Uniformity of distribution ease interpretation as it allows obtaining closed-form expressions. It also leads to rule out multiple equilibria which can arise in hostile informational environments (see Fleckinger et al., 2012).

We assume the existence of a stakeholder who has a (marginal) willingness to pay $w_1 \geq 1$ for self-regulation and $w_0 \leq 0$ for business-as-usual[2]. But the stakeholder is not able to observe a on her own and bases her desicion to reward or not the firms on the information provided by the NGO. More specifically:

- In the case where a firm chooses $a = 1$, the NGO discloses the value of a with a probability g. With a probability $1 - g$, no information is generated about the firm's action. The stakeholder believes the news (hard information). We rule out any potential concerns about the NGO credibility.

- In the case where $a = 0$, the disclosure probability is b. With probability $1 - b$, she receives no news.

If the stakeholder learns that $a = 1$, she transfers w_1 to that firm. Hence, we assume that the firm is able to fully extract the stakeholder's surplus.[3] Conversely, she punishes the firm with a negative transfer w_0 when she learns that $a = 0$. When she receives no news about the action of a given firm, she forms a belief μ that the firm self-regulates and transfers $\mu w_1 + (1 - \mu) w_0$ to the firm. The stakeholder is sophisticated in that she relies on Bayes' rule to form her belief.

The disclosure technology is thus fully described by the probabilities (g, b). These are endogenously chosen by the NGO, which can decide to be neutral $(g = b)$, a good cop $(g > b)$, or a bad cop $(g < b)$. However, the NGO has limited resources which prevent to obtain and/or convey perfect information about a. More specifically, we introduce the assumption that $g + b \leq \alpha$.[4]

The timing of the game is as follows:

1. The NGO selects the disclosure probabilities (g, b).

2. Each firm decides to self-regulate ($a = 1$) or not ($a = 0$).

3. The NGO discloses the value of a with probabilities g and b if $a = 1$ and $b = 0$, respectively.

[2]At some point, we normalize $w_1 - w_0$ to 1, which does not alter our results qualitatively

[3]We could assume that the firms could only extract a share, say α, of the surplus. This would not affect qualitatively the results.

[4]As we will see below the linearity of this resource constraint simplifies the analysis by generating simple corner solutions. But relaxing this assumption would not qualitatively alter the results.

4. The stakeholder transfers w_0, w_1 or $\mu w_1 + (1-\mu)w_0$ depending on the messages received.

Note that, if information on a were perfect, all firms would self-regulate, which is socially optimal as $c \leq w_1$ for all c. Our analysis explore how information asymmetry prevents from reaching this first best outcome. In this respect, the fact that the NGO has limited resources is crucial. Without this hypothesis, the NGO would provide perfect information ($g = b = 1$), inducing self-regulation by all firms.

We now proceed in two steps. First, we characterize firms' response to a given monitoring technology (g, b). Second we identify the technology selected by the NGO which seeks to maximize CSR under the constraint $g + b = \alpha$.

2.2 Firm's choice

Consider a firm of type c. Its expected payoff is $\Pi(a=1) = gw_1 + (1-g)\left(\mu w_1 + (1-\mu)w_0\right) - c$ under self-regulation while its payoff is $\Pi(a=0) = bw_0 + (1-b)\left(\mu w_1 + (1-\mu)w_0\right)$ otherwise. Hence, the firm self-regulates if:

$$c \leq (w_1 - w_0)((g - (g-b)\mu)$$

An almost immediate consequence of this incentive constraint is that the CSR equilibrium is characterized by a cost threshold below which firms self-regulate and above which they do not. Let c^* denote that threshold and normalize $w_1 - w_0 = 1$ to simplify notations. We have

$$c^* = g - (g-b)\mu \tag{1}$$

The belief that is consistent in the Bayesian sense with this cutoff is then

$$\mu^* = \frac{(1-g)c^*}{(1-g)c^* + (1-b)(1-c^*)} \tag{2}$$

The two equations (1) and (2) defines the Bayesian equilibrium.

It will prove more convenient to adopt a standard fixed point representation of the equilibrium to investigate equilibrium existence properties. Combining (1) and (2), the cutoff equilibrium satisfies

$$c^* = \Phi(c^*) \quad \text{where} \quad \Phi(c) \equiv \frac{g(1-b) - (g-b)c}{(1-b) - (g-b)c} \tag{3}$$

We establish properties of Φ that will be useful in the following:

Lemma 1 *1) $\Phi(0) = g$, $\Phi(1) = b$. 2) Φ is increasing if $g > b$ and decreasing if $g \leq b$. 3) Φ is concave.*

Proof. 1) Obvious. 2) Differentiation of Φ yields $\frac{d\Phi}{dc}(c) = -\frac{(g-b)(1-g)(1-b)}{[1-b-(g-b)c]^2}$, which is positive if $g \leq b$ and negative otherwise. 3) We have $\frac{d^2\Phi}{dc^2}(c) = \frac{(g-b)^2(1-g)(1-b)}{[(g-b)c-(1-b)]^3}$, which have the same sign as that of $(g-b)c-(1-b)$. This expression is obviously negative when $g - b \leq 0$. When $g - b > 0$, it is maximized in $c = 1$ and $(g-b)c - (1-b) = -(1-g) < 0$ in this case. ■

Then we use these properties to draw the function Φ in Figure 1. In this graph, any equilibrium corresponds to the intersection of Φ with the 45 degree line. When $g > b$ (Fig 1a), there exists a unique equilibrium as Φ is decreasing, starts above the 45 degree line ($\Phi(0) = g > 0$) and ends up below ($\Phi(1) = b < 1$).

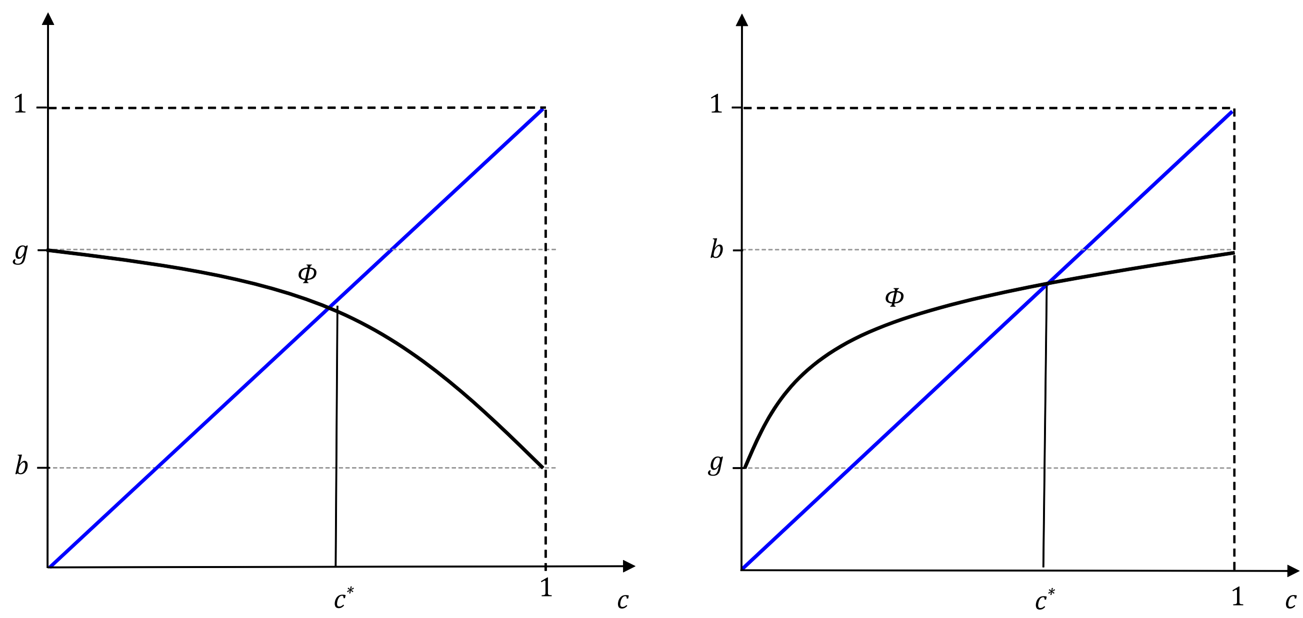


Figure 1: Good cop ($g > b$) versus bad cop ($g < b$) equilibria

In the case where $g < b$, Φ is increasing. If $g > 0$, we have $\Phi(0) = g > 0$, meaning that Φ starts above the 45 degree line in $c = 0$. It is the case depicted in Fig. 1b. Property

1 and the intermediate value theorem ensures existence. The equilibrium is also unique: as Φ crosses the 45 degree line from above in c^*, concavity of Φ ensures it will never cross it for higher values of c.

Things are slightly more complicated when $g = 0$. To begin with, $c = 0$ is always an equilibrium ($\Phi(0) = 0$). If $d\Phi/dc(0) \leq 1$, which is equivalent to $b \leq \frac{1}{2}$, this equilibrium is unique for Φ is concave as argued before. When $d\Phi/dc(0) > 1$, there will be a second equilibrium as Φ is higher than c when deviating upward from zero.

We now summarize our findings.

Proposition 1 (Existence) *There always exists a (subgame) equilibrium which is defined by the condition (3). But the equilibrium is not necessarily unique. More precisely:*

- *When $g > 0$, there exists a unique equilibrium c^* which is strictly positive. c^* is defined by (3).*
- *When $g = 0$ and $b \leq \frac{1}{2}$, the unique equilibrium is $c^* = 0$.*
- *When $g = 0$ and $b > \frac{1}{2}$, there exists two equilibria. The first is $c^* = 0$ and the second is strictly positive.*

In the following, we rely on simulations to establish some of our results in order to avoid cumbersome calculations. We thus need an explicit expression of c^* in the case where it differs from zero.

Lemma 2 *Equilibrium with a strictly positive amount of CSR is defined as follows:*

- *When $g = b$, we have $c^* = b = g$.*
- *When $g \neq b$:*

$$c^* = \frac{1}{2(g-b)}\left(g - 2b + 1 - \sqrt{(g-2b+1)^2 - 4g(1-b)(g-b)}\right)$$

Proof. See in Appendix. ■

2.3 NGO choice

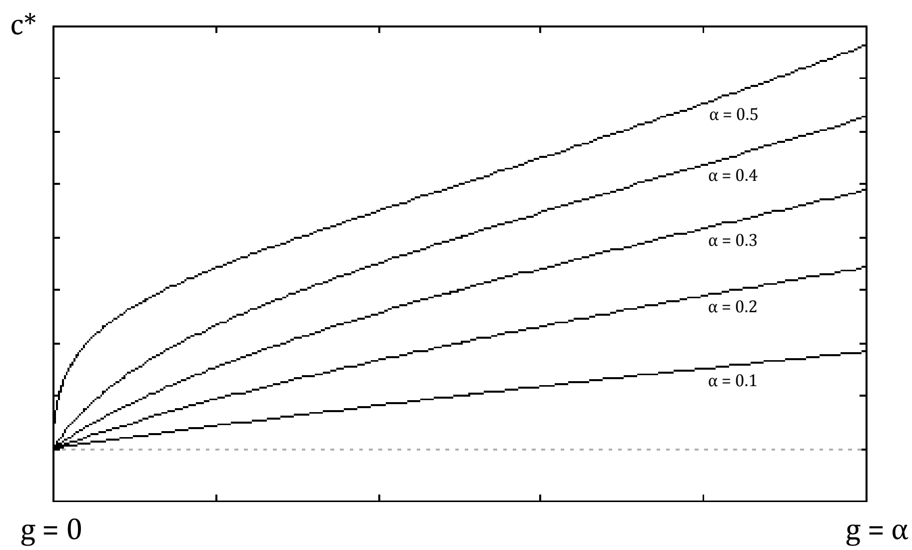
Our objective is now to characterize the disclosure probabilities (g^*, b^*) that will be selected by the NGO, given the firms' response. Formally, the NGO solves

$$\max_{g,b} c^*(g,b) \text{ subject to } g + b = \alpha \tag{4}$$

where c^* is the equilibrium defined by Proposition 1.

The possibility that the NGO chooses g^* and b^* such that $c^* = 0$ is excluded since it precisely seeks to maximize c^*[5].We can thus concentrate on the equilibrium characterized in Lemma 2.

We first run simulations to show that there is no interior maximum. The results are displayed in Figure 2 which plots c^* as a function of g for different values of α. We can see that either c^* increases with g when α is low or the curve is U-shaped with an interior minimum.



[5]This also applies to the zero equilibrium when it coexists with the positive equilibrium in the case where $g = 0$ and $b > 1/2$: as soon as g deviates upward from zero with an infinitesimal amount, this equilibrium disappears.

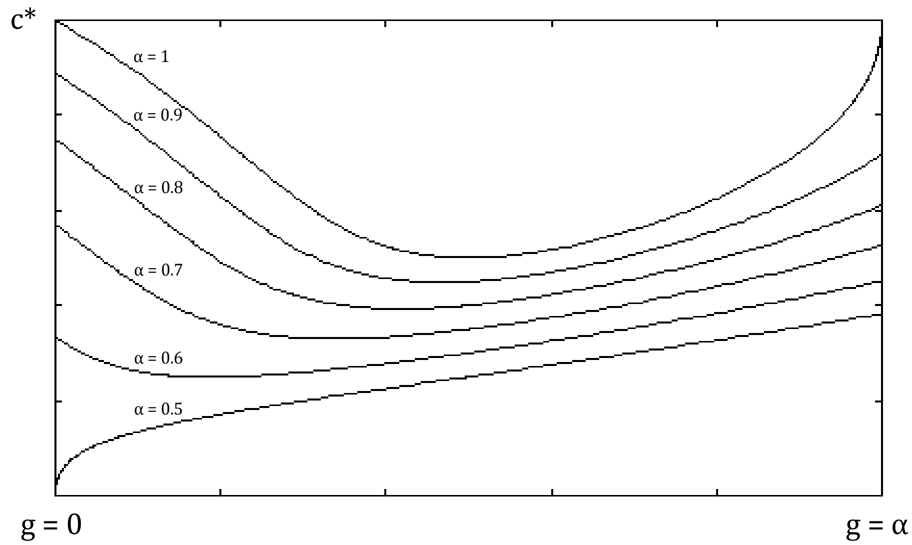


Figure 2: c^* **as a function of** g **for different value of** α

Having ruled the existence of interior solutions out, we compare the two corner solutions ($g = \alpha$ and $g = 0$) in Appendix. We obtain a new proposition.

Proposition 2 *The NGO chooses to be a bad cop with* $g = 0$ *and* $b = \alpha$ *if* $\alpha > \frac{2}{3}$. *Otherwise it chooses to be a good cop (*$g = \alpha$ *and* $b = 0$*).*

Proof. See in Appendix. ■

Let us now discuss the conditions $\alpha \gtrless \frac{2}{3}$. A high α means that the NGO produces a lot of information about firms' behavior. Incentives for firms to invest in self-regulation are thus high, implying that the stakeholder holds an optimistic belief that one firm about whom she does not receive any news is actually self-regulating. In this context, it is more effective for the NGO to be hostile in order to deter firms from shirking on this optimistic state of mind. It mitigates their incentives to free ride on the lack of information. And conversely when the stakeholder is pessimistic about the level of self-regulatory efforts (because α is low), it is more effective to motivate firms by increasing the probability g to get the full reward when self-regulating rather than the probability b to get close nothing because of the stakeholder's low belief when exerting no effort. The general point here is that the best strategy consists in going against the stakeholder's a priori built upon the environment characteristics captured by the parameter α[6].

[6]One should not give too much importance to the fact that we only obtain corner solutions for it is due

As the NGO maximizes the average level of self-regulation and as self-regulation improves social welfare, note that:

Corollary 1 *The NGO choice is socially optimal.*

3 Multiple NGOs

In reality, firms are under the scrutiny of several NGOs. That is why we extend the base model to account for the existence of n NGOs. Each has the same budget $\alpha(\alpha < 1)$. Let g_i and b_i denote the values of g and b selected by the NGO i ($i = 1, ..n$). Each NGO has the budget constraint $g_i + b_i = \alpha$. The overall probability G that a firm with $a = 1$ is observed by at least one NGO is thus

$$G(g_1, ..g_i, ..g_n) = g_i + (1 - g_i)\left(1 - \prod_{j \neq i}(1 - g_j)\right) \tag{5}$$

That is, either the firm is observed by the NGO i or it does not, but at least another NGO $j \neq i$ does so[7]. Similarly, the probability that $a = 0$ is disclosed is

$$B(b_1, ..b_i, ..b_n) = b_i + (1 - b_i)\left(1 - \prod_{j \neq i}(1 - b_j)\right). \tag{6}$$

The level of self-regulation is still defined by Proposition 1 except that we replace g and b by the aggregate probabilities G and B.

3.1 Best reply

The best reply of the NGO i to the $n - 1$ other NGOs strategies is the couple (g_i, b_i) which solves

$$\max_{g_i, b_i} c^* \left(G(g_1, ..g_i, ..g_n), B(b_1, ..b_i, ..b_n)\right) \text{ subject to } g_i + b_i = \alpha$$

to the linear budget constraint $g + b = \alpha$. Other functional forms would probably lead to interior solutions, but this would not modify qualitatively the result that the NGO does not select $g = b$ and the nature of the influence of the environment characteristics.

[7] $\prod_{j \neq i}(1-g_j)$ is the probability that none of the $n - 1$ other NGOS discloses $a = 1$. Hence, the probability that at least one discloses it is $1 - \prod_{j \neq i}(1-g_j)$.

Note that:

$$\begin{aligned}\frac{\partial G}{\partial g_i} &= \prod_{j\neq i}(1-g_j) > 0,\\ \frac{\partial B}{\partial b_i} &= \prod_{j\neq i}(1-b_j) > 0.\end{aligned}$$

These properties simplify the problem as raising g_i or b_i basically means raising G or B. The aggregate probability G will be equal to $G(g_1,..,g_{j-1},0,g_{j+1}..g_n)$ if the NGO selects $g_i = 0$ and to $G(g_1,..,g_{j-1},\alpha,g_{j+1}..g_n)$ if $g_i = \alpha$. Hence, selecting g_i amounts to select a value of G in the interval $A_i \equiv \left[G(g_1,..,g_{j-1},0,g_{j+1}..g_n), G(g_1,..,g_{j-1},\alpha,g_{j+1}..g_n)\right]$. Consider then Figure 2. A_i is a segment of the curves that are depicted in this graph. In the case where c^* is increasing over the whole interval A_i, the NGO i is better off by increasing g_i as much as possible. Hence, it chooses $g_i = \alpha$. In the case where it is decreasing, $g_i = 0$. In the case where the curve is U-shaped over A_i, the NGO's best reply is either $g_i = 0$ or $g_i = \alpha$, depending on the precise location of A_i on the curve.

The common point to every case is that the best reply is a corner choice whatever the others' strategies. The last step of this analysis consists in identifying the precise condition that defines the best of the two candidate corner solutions. We obtain the following:

Lemma 3 *All NGOs have the same best reply function. If*

$$\frac{\alpha + \left(2 - \prod_{j\neq i}(1-g_j) - \prod_{j\neq i}(1-b_j)\right)(1-\alpha)}{2-\alpha} > \frac{\prod_{j\neq i}(1-b_j)}{\prod_{j\neq i}(1-g_j) + \prod_{j\neq i}(1-b_j)}, \tag{7}$$

the best reply is to be a bad cop ($g_i = 0$ and $b_i = \alpha$). Otherwise, they prefer to be good cops ($g_i = \alpha$ and $b_i = 0$).

Proof. See in Appendix. ■

3.2 Nash equilibrium

Having identified the best reply function of each NGO in Lemma 3, we now identify the resulting Nash equilibrium. The intermediate cases where some NGOs would opt for being a good cop and others for being a bad cop can be easily excluded. It would result in an equilibrium value of G strictly higher than $G(0,..,0)$ and strictly lower than $G(\alpha,..,\alpha)$.

But, given the function c^* depicted in Figure 2, some NGO would then have an incentive to deviate towards one of the corner values of G. We can thus concentrate the analysis on the corner cases.

Consider first the equilibrium in which all NGOs choose to be bad cops: $g_i = 0$, $b_i = \alpha$. Plugging these values in (5) and (6), the resulting overall probabilities are $G = 0$ and $B = 1 - (1-\alpha)^n$. Hence, the constraint (7) becomes

$$1 > 2(1-\alpha)^n + (1-\alpha)^{2n-1}$$

Taking Note that the left-hand side decreases with α. Furthermore, the inequality is satisfied when $\alpha = 1$ and not when $\alpha = 0$. Hence, there exists a unique value of $\alpha \in (0,1)$, denoted α_{lim}, such that:

$$1 \equiv 2(1-\alpha_{\text{lim}})^n + (1-\alpha_{\text{lim}})^{2n-1} \tag{8}$$

For all the values of α such that $\alpha > \alpha_{\text{lim}}$, there exists a Nash equilibrium where all the NGOs choose to be a bad cop.

Turning next the equilibrium where all NGOs choose to be good cops ($g_i = \alpha, b_i = 0$), the best reply condition (7) is

$$1 - (1-\alpha)^n < \frac{(1-\alpha)+1}{(1-\alpha)^{n-1}+1}$$

It is easy to see that this condition is always satisfied as the left-hand side is higher than 1 whereas the right-hand side is less than 1. We can thus conclude.

Proposition 3 *The equilibrium depends on whether the NGOs' budget is higher or lower than a threshold α_{lim} defined by (8).*

- *If $\alpha < \alpha_{\text{lim}}$, there exists only one Nash equilibrium where all NGOs choose to be good cops.*
- *If $\alpha \geq \alpha_{\text{lim}}$, there exists two equilibria. In the first one, all NGOs are good cops. In the second, all are bad cops.*

This proposition highlights a coordination problem that arises when there are a lot of NGOs (α_{lim} decreases with the number of NGOs) with a confortable budget. The equilibrium is thus not always socially efficient contrary to the single-NGO case (see Corollary 1). This creates the opportunity for a welfare-improving public intervention. The next proposition presents a simple decision rule to help NGOs to select the Pareto dominant equilibrium.

Proposition 4 *1. If $\alpha < \alpha_{\lim}$, public intervention is not necessary*

2. If $\alpha_{\lim} \leq \alpha \leq 1 - 3^{-\frac{1}{n}}$, the regulator should help the NGOs to select the good cop equilibrium.

3. If $\alpha > 1 - 3^{-\frac{1}{n}}$, it should help to select the bad cop equilibrium.

Proof. See in Appendix. ■

Subsidizing NGOs is another way of improving social welfare. It amounts to increasing the budget α available to each NGO, which in turn leads to higher disclosure probabilities G or B. Ignoring the opportunity cost of such subsidies, welfare obviously improves as higher probabilities raise the level of self-regulation:

Proposition 5 *Subsidizing NGOs improves social welfare as it raises disclosure probabilities and thus the level of self-regulation ($\frac{dc^*}{dG} > 0$, $\frac{dc^*}{dB} > 0$).*

Proof. See in Appendix. ■

4 Corporate-NGO partnership

In reality, firms have the opportunity to develop partnerships with NGOs. They transfer resources to some organizations, which are used to finance monitoring and reporting activities, resulting in better information on firms' behavior. In this section, we introduce in our setting the possibility for firms to enter partnerships and we examine how it influences the results obtained so far.

A NGO may commit to publicize specifically the partnering firm's social or environmental performance. For instance, the NGO grants a product label. In other cases, the firm contributes with a donation, enabling the NGO to increase its overall activities, but without targeting specifically its donors. In the following, we refer to the second scenario as a collective partnership as the increase in disclosure probabilities concerns all firms whereas the first is referred to as an individual partnership. We successively analyze these two forms.

4.1 Collective partnership

Under a collective partnership, a firm's contribution leads to an increase in the overall probabilities G or B. We pose the problem as follows. In the status quo, n NGOs with budget α have selected their communication strategy. We assume the equilibrium is socially efficient (because they were able to coordinate, eventually with a regulator's support). Hence if $\alpha < 1 - 3^{-\frac{1}{n}}$, they are all good cops. Otherwise, they are all bad cops. We then study whether an individual firm's equilibrium profit increases with disclosure probabilities G or B. To begin with, we establish a Lemma describing the marginal effect of G and B on equilibrium profits.

Lemma 4 *We have:*

1. $\frac{d\pi(a=1)}{dB} > 0$ *and* $\frac{d\pi(a=1)}{dG} > 0$
2. $\frac{d\pi(a=0)}{dB} > 0$ *and* $\frac{d\pi(a=0)}{dG} > 0$ *if* $\alpha < 1 - 3^{-\frac{1}{n}}$ *and* ≤ 0 *otherwise.*

Proof. See in Appendix. ■

We can immediately derive that:

Proposition 6 *If the status quo is the good cop equilibrium ($\alpha < 1 - 3^{-\frac{1}{n}}$), all firms are willing to enter collective partnerships. If it is the bad cop equilibrium ($\alpha \geq 1 - 3^{-\frac{1}{n}}$), the self-regulating firms are the only potential contributors.*

This result is very intuitive: firms which do not self-regulate do not contract with bad cops. Note however that the bad cop equilibrium emerges only when the average level of CSR is high (because $\alpha \geq 2/3$), meaning that the share of firms which does not self-regulate is low in relative terms. Note also that this analysis only looks at firms' preferences, ruling out free riding issues stemming from the fact that the benefit of collective partnership is a public good. Hence, the proposition only gives a necessary condition for the existence of collective partnerships.

4.2 Individual partnership

We model individual partnership as follows: Each firm has the option to contract with a NGO which commits to disclose the firm's action with a probability equal to one[8] in

[8] Note that the distinction between a good cop and a bad cop is not relevant anymore when the NGO is engaged in a partnership. If the NGO is initially a good cop which increase its probability to $g = 1$. This implies that receiving no news means $a = 0$ for the stakeholder (as if $b = 1$). Conversely, if the NGO is a bad cop commiting to $b = 1$, receiving no news means $a = 1$ (as if $g = 1$).

exchange of a transfer t. Let s denote a binary variable equal to one if the firm enters a partnership and 0 it it does not. We assume the stakeholder observes s. To simplify notations, we also assume $w_0 = 0$ and $w_1 = 1$ (hence, we still have $w_1 - w_0 = 1$).

In this new version of the game, a firm's strategy is the pair (a, s) and its payoff is

$$\Pi(a,s) = \begin{cases} G + (1-G)\mu_1 - c & \text{if } a = 1,\ s = 0 \\ 1 - c - t & \text{if } a = 1,\ s = 1 \\ (1-B)\mu_2 & \text{if } a = 0,\ s = 0 \\ -t & \text{if } a = 0,\ s = 1 \end{cases}$$

Note that beliefs can be different when the strategy pair is $(a = 1, s = 0)$ and $(a = 0, s = 0)$. We can immediately rule out the strategy $(a = 0, s = 1)$ because it is strictly dominated by $(a = 0, s = 0)$: it is irrational to costly reveal $a = 0$.

Consider now a firm that self-regulates $(a = 1)$. The firm has an incentive to contract with a NGO if $G + (1-G)\mu_1 - c < 1 - c - t$. Hence, if $t < (1-G)(1-\mu_1)$. This condition does not depend on c; it implies that either all self-regulating firms decide to signal their action or none of them does so. It simplifies the stakeholder's Bayesian updating: as soon as she observes one partnership, she knows that entering a partnership is profitable for all firms with $a = 1$. Reciprocally, all firms without partnership have thus chosen $a = 0$. In fact, as soon as there is at least one partnership, the belief μ_2 is equal to 0. In this context, each firm decides to self-regulate if its payoff with self-regulation and partnership, $1 - c - t$ is positive (the payoff without self-regulation and without partnership is equal to zero). We get the classical unravelling separating equilibrium in which firms choose $(a = 1, s = 1)$ if $c \leq 1 - t$ and $(a = 0, s = 0)$, otherwise.

Let us now come back to the case where self-regulating firms do not signal their action $(t \geq (1-G)(1-\mu_1))$. As firms with $a = 0$ does not contract with a NGO, the stakeholder does not observe any partnership. It is the standard case without partnership analyzed in the previous sections: firms choose to self-regulate if $c < c^*$ with c^* defined by (1) and $\mu_1 = \mu^*$ defined by (2). We now summarize these results.

Lemma 5 *When firms have the option to contract with a NGO such that its disclosure probability becomes one at cost t, two equilibria are possible:*

1. *If $t \geq (1-G)(1-\mu^*)$, none of the firms contract with a NGO and firms with type $c < c^*$ choose to self-regulate with c^* defined by (1).*

2. *If $t < (1-G)(1-\mu^*)$, the firm of type $c \leq 1-t$ self-regulates and contract with a NGO. The others do nothing.*

In the second case, we get a separating equilibrium which fundamentally differs from the equilibrium obtained in the base model. It is easy to show that:

Proposition 7 *The average level of self-regulation is higher when firms have the possibility to enter individual partnerships.*

Proof. Remember that NGOs choose to be either all good cops ($g = \alpha, b = 0$) or all bad cops ($g = 0, b = \alpha$) in the base model. If they all choose to be good cops, we have (from 1) that $c^* = \alpha(1-\mu^*)$. Besides, we know that $t < (1-\alpha)(1-\mu^*)$, thus $1-t > \alpha - \alpha(1-\mu^*)$.

If they all choose to be good cops we have (from 1) that $c^* = \alpha\mu^*$. Besides, we know that $t < (1-\mu^*)$ which implies $1-t > \mu^* > c^*$. ■

The proposition is very intuitive as partnerships basically increase the amount of information on self regulation. When are then individual partnerships more likely to occur? Answering the question requires to interpret the threshold $(1-G)(1-\mu^*)$. In the bad cops equilibrium ($\alpha < 1-3^{-\frac{1}{n}}$), the threshold simplifies to $1-\mu^*$. When establishing Lemma 5, we have seen that μ^* was increasing with B. Hence, the higher the resources available to the NGOs, the lower this threshold, and thus the less likely individual partnerships. In the good cops equilibrium, we have seen that μ^* was also increasing with G. Hence, the threshold is lower and the partnership less likely. Both results go in the same direction:

Corollary 2 *The higher NGOs budget, the less likely individual partnerships.*

5 Conclusion

The main objective of the paper is to characterize what drives NGOs' communication strategy, contrasting good cops − NGOs which disclose information about firms which self-regulate − and bad cops − NGOs which disclose information on firms which do not. To answer these questions, we consider a model featuring a continuum of firms which can self-regulate or not, a stakeholder who is willing to reward self-regulation but who is not able to observe the firms' behavior and non governmental organizations (NGOs) that induce self-regulation by imperfectly disclosing on firms' level of self-regulation.

We show that the NGO decides to specialize: if its budget is low, it chooses to be a good cop. When the budget is larger, it opts for being a bad cop. The general point here is that the best information disclosure strategy consists in going against the stakeholder's a priori built upon the environment characteristics: a high budget mean that the NGO produces much information about firms' behavior. Incentives for firms to invest in CSR are thus high, implying that the stakeholder holds an optimistic belief that one firm about whom she does not receive any news is actually self-regulating. In this context, it is more effective for the NGO to be hostile as this mitigates their incentives to free ride on the lack of information. And conversely when the stakeholder is pessimistic about the level of self-regulation (because NGOs' budget is limited), it is more effective to motivate firms by increasing the probability g to get the full reward when investing in self-regulation.

We develop several extensions of the model. First, we assume multiple NGOs which non cooperatively select their technology. A coordination problem arises for the two possible equilibria -NGOs being all bad cops and NGOs being all good cops- sometimes coexist. We also introduce the option for firms to cooperate with NGOs. More specifically, we distinguish individual partnership, whereby the NGO specifically publicizes the partnering firm's behavior, and collective partnership whereby industry support leads to an increase of the aggregate disclosure probability.

Although the prime focus of the analysis is positive, it is possible to derive a general policy lesson. When the amount of resources available to NGOs is limited and/or the number of NGOs is low, everything is going smoothly and the role of a welfare-maximizing regulator is limited: There exists a single equilibrium which is socially optimal. As all NGOs are good cops, firms are willing to engage in collective or individual partnerships in order to increase the quantity of information, which in turn increases self-regulation and thus social welfare. Things get more complicated when the quantity of information is such that the social outcome is the bad cop equilibrium. In this case, the coordination problem can lead to the survival of the good cop equilibrium. It is reinforced by the fact that firms which do not self-regulate do not form partnerships anymore and self-regulating firms are less prone to enter in partnership for the benefits get smaller. To sum up, a public intervention promoting bad cops becomes increasingly useful.

References

Arora, S. and Gangopadhyay, S., 1995. Toward a theoretical model of voluntary overcompliance. *Journal of Economic Behavior and Organization* 28, 289-309.

Baron, D., 2001. Private Politics, Corporate Social Responsibility, and Integrated Strategy. *Journal of Economics & Management Strategy* 10(1), 7-45.

Bagnoli, M. and Watts, S., 2003. Selling to Socially Responsible Consumers: Competition and The Private Provision of Public Goods. *Journal of Economics & Management Strategy* 12(3), 19-445.

Bottega, L. and De Freitas, J., 2009. Public, Private and Nonprofit Regulation for Environmental Quality. *Journal of Economics & Management Strategy* 18(1), 105-123.

Brekke, K. and Nyborg, K., 2008. Attracting responsible employees : Green production as labor market screening. *Resource and Energy Economics* 30(4), 509-526.

Feddersen, T. J. and Gilligan, T. W., 2001, Saints and Markets: Activists and the Supply of Credence Goods. *Journal of Economics & Management Strategy*, 10: 149–171.

Fleckinger, P., Glachant, M. and Moineville, G., 2012. Incentives for Quality in Friendly versus Hostile Informational Environments. *Working paper*

Heinkel, R., Kraus, A. and Zechner, J., 2001. The Effect of Green Investment on Corporate Behaviour. *Journal of Financial and Quantitative Analysis*, 36, 431-450.

Heyes, A., 2005. A Signaling-motive for Self-regulation in the Shadow of Coercion. *Journal of Business and Economics* 57, 238-246.

Lyon, T. (Ed.), 2010. *Good Cop/Bad Cop: Environmental NGOs and Their Strategies*. RFF Press, Washington, DC.

Lyon, T.P. and Maxwell, J.W., 2011. Greenwash: Corporate Environmental Disclosure under Threat of Audit. *Journal of Economics and Management Strategy*, forthcoming.

Yaziji, M., and Doh, J., 2009. NGOs and Corporations: Conflict and Collaboration. Cambridge, UK; Cambridge University Press

6 Appendix

6.1 Proof of Lemma 2

To begin, we plug $g = b$ in (3) to obtain c^* in the case where $g = b$. Then, solving (3) for c yields two roots:

$$
\begin{aligned}
c_1 &= \frac{1}{2(g-b)}\left(g - 2b + 1 + \sqrt{(g-2b+1)^2 - 4g(1-b)(g-b)}\right) \\
c_2 &= \frac{1}{2(g-b)}\left(g - 2b + 1 - \sqrt{(g-2b+1)^2 - 4g(1-b)(g-b)}\right)
\end{aligned}
$$

When $g - b > 0, c_1 > 1$ as this inequality is equivalent to $\sqrt{(g-2b+1)^2 - 4g(1-b)(g-b)} > -(1-g)$, which is obviously satisfied.

When $g - b < 0$, $\sqrt{(g-2b+1)^2 - 4g(1-b)(g-b)} \geq |g - 2b + 1|$. Then we have two subcases: If $g - 2b + 1 \geq 0$, $|g - 2b + 1| = g - 2b + 1$, the term in brackets is obviously positive, implying $c_1 < 0$. If $g - 2b + 1 < 0$, $|g - 2b + 1| = -(g - 2b + 1)$. Hence $\sqrt{(g-2b+1)^2 - 4g(1-b)(g-b)} \geq -(g - 2b + 1)$. The term in brackets is thus positive, meaning that $c_1 \leq 0$.

6.2 Proof of Proposition 2

For ease of presentation, we introduce the following notations: $c_g^* \equiv c^*(\alpha, 0)$ and $\Phi_g \equiv \Phi(c)$ when $g = \alpha$ and $b = 0$. Hence $\Phi_g = \frac{\alpha(1-c)}{1-\alpha c}$. Similarly, c_b^* is the stable equilibrium $c^*(0, \alpha)$ and $\Phi_b \equiv \Phi(c) = \frac{\alpha c}{1-\alpha(1-c)}$. Furthermore, let c° denote the value of c where $\Phi_g = \Phi_b$. Straightforward calculations show that $\Phi_g(c^\circ) = \Phi_b(c^\circ) = \frac{\alpha}{2-\alpha}$ and $c^\circ = \frac{1}{2}$. Then we consider two cases: $\Phi_g(c^\circ) = \Phi_b(c^\circ) > c^\circ$ - which is equivalent to $\alpha > 2/3$ - and $\Phi_g(c^\circ) = \Phi_b(c^\circ) \leq c^\circ$.

Case 1: $\alpha > 2/3$ From the intermediate value theorem (IVT) follows directly $c_g^* > c^\circ$: $\Phi_g(c^\circ)$ and $\Phi_g(1) = 0$ are above and below the 45 degree line, respectively. The same theorem also implies $c_b > c^\circ$ as Φ_b is above the 45 degree line in c° and ends up below in $c = 1$ ($\Phi_b(1) = \alpha < 1$).

Since Φ_g is decreasing and Φ_b is increasing (Property 2 in Lemma 1), we have the general property that $\Phi_g(c) < \Phi_b(c')$ for any $c, c' > c^\circ$. It is thus satisfied in the particular cases where $c = c_g^*$ and $c\prime = c_b^*$. It implies $c_g^* < c_b^*$ as $c_g^* = \Phi_g(c_g^*)$ and $c_b^* = \Phi_b(c_b^*)$.

Case 2: $\alpha \leq 2/3$ The IVT implies that $c_g^* \leq c^\circ$. We also have $c_b^* \leq c^\circ$ because (i) Φ_b is now below the 45 degree line in c° , (ii) it ends up below in $c = 0$ ($\Phi_b(0) = \alpha$) and (iii) Φ_b is concave (see footnote X). These three properties prevent Φ_b to intersect with the 45 degree line beyond c°. Then $\Phi_g(c) \leq \Phi_b(c\prime)$ for any $c, c\prime \leq c^\circ$, $c_g^* \leq c^\circ$ and $c_b^* \leq c^\circ$ implies $c_g^* > c_b^*$ in this case.

6.3 Proof of Lemma 3

For ease of presentation, we introduce the following notations :

$$\begin{aligned}
g_{-i} &= \left(1 - \prod_{j \neq i}(1 - g_j)\right) \text{ and } b_{-i} = \left(1 - \prod_{j \neq i}(1 - b_j)\right) c_{b_i}^* \equiv c^*(g_{-i}, b_{-i} + \alpha(1 - b_{-i})) \\
c_{g_i}^* &\equiv c^*(g_{-i} + \alpha(1 - g_{-i}), b_{-i}) \\
\Phi_{g_i}(c) &\equiv \Phi(c) \text{ with } g_i = \alpha \text{ and } b_i = 0 \\
\Phi_{b_i}(c) &\equiv \Phi(c) \text{ with } g_i = 0 \text{ and } b_i = \alpha \\
c^\circ &= \frac{1 - b_{-i}}{2 - (b_{-i} + g_{-i})} \text{ such that } \Phi_{g_i}(c) = \Phi_{b_i}(c) \\
\Phi_{g_i}(c^\circ) &= \frac{\alpha + (b_{-i} + g_{-i})(1 - \alpha)}{2 - \alpha}
\end{aligned}$$

Then we follow closely the proof of Proposition 2. We know that

$$\Phi_{b_i}(0) < \Phi_{g_i}(0) \text{ and } \Phi_{b_i}(1) > \Phi_{g_i}(1), \text{ for any } g_i, g_{-i}, b_i, b_{-i}.$$

It follows that $\Phi_{b_i} < \Phi_{g_i}$ if $c < c^\circ$ and $\Phi_{b_i} \geq \Phi_{g_i}$ if $c \geq c^\circ$. Now we consider two subcases:

- If $\Phi_{g_i}(c^\circ) = \Phi_{b_i}(c^\circ) > c^\circ$, we necessarily have $c_{b_i}^* > c^\circ$ and $c_{g_i}^* > c^\circ$ from the IVT. We also know that $\Phi_{b_i} > \Phi_{g_i}$ if $c > c^\circ$ and $\Phi_{b_i}(1) > \Phi_{g_i}(1)$. Hence, Φ_{g_i} will necessarily the 45 degree line before Φ_{b_i} when departing upward from c°. Hence $c_{b_i}^* > c_{g_i}^*$.
- If $\Phi_{g_i}(c^\circ) = \Phi_{b_i}(c^\circ) < c^\circ$, we symmetrically have $c_{b_i}^* < c^\circ$ and $c_{g_i}^* > c^\circ$ and $\Phi_{b_i} < \Phi_{g_i}$ if $c < c^\circ$ and $\Phi_{b_i}(0) < \Phi_{g_i}(0)$. Hence $c_{g_i}^* > c_{b_i}^*$.

6.4 Proof of Proposition 4

We start characterizing the socially optimal benchmark. In the good cop equilibrium, we have $G = 1 - (1 - \alpha)^n$ and $B = 0$, whereas $G = 0$ and $B = 1 - (1 - \alpha)^n$ if all NGOs are bad cops. For ease of presentation, let P denote $1 - (1 - \alpha)^n$. We thus need to compare

$c^*(0,P)$ with $c^*(P,0)$. It is exactly the comparison we have made in Proposition 2, which establishes that $c^*(P,0) > c^*(0,P)$ if $P < 2/3$. It is equivalent to $\alpha \leq 1 - 3^{-\frac{1}{n}}$ Then we show that $\alpha_{\text{lim}} \leq 1 - 3^{-\frac{1}{n}}$, which is equivalent to $3^{-\frac{1}{n}} < 1 - \alpha_{\text{lim}}$. Hence $2(1-\alpha_{\text{lim}})^n + (1-\alpha_{\text{lim}})^{2n-1} > 2/3 + 3^{\left(\frac{1}{n}-2\right)}$. The lef-hand side is less than 1: it decreases with n and it is equal to 1 when $n = 1$. This completes the proof as $2(1-\alpha_{\text{lim}})^n + (1-\alpha_{\text{lim}})^{2n-1} = 1$ by definition.

6.5 Proof of Proposition 5

Differentiating the equilibrium condition $c^* = \Phi(c^*)$ and rearranging, we obtain:

$$\frac{dc^*}{dG} = \frac{\frac{\partial \Phi}{\partial G}}{1 - \frac{d\Phi}{dc}} \text{ and } \frac{dc^*}{dB} = \frac{\frac{\partial \Phi}{dB}}{1 - \frac{d\Phi}{dc}}$$

The stability condition is $1 - \left|\frac{d\Phi}{dc}(c^*)\right| > 0$ which implies that $1 - \frac{d\Phi}{dc} > 0$. (It is true when $\left|\frac{d\Phi}{dc}(c^*)\right| = -\frac{d\Phi}{dc}(c^*)$ and $\left|\frac{d\Phi}{dc}(c^*)\right| = \frac{d\Phi}{dc}(c^*)$). Hence we just need to look at the signs of $\frac{\partial \Phi}{dG}$ and $\frac{\partial \Phi}{dB}$, which are both positive as

$$\frac{\partial \Phi}{\partial G} = \frac{(1-B)^2 (1-c^*)}{(B-1+c(G-B))^2} \text{ and } \frac{\partial \Phi}{\partial B} = \frac{c^* (1G)^2}{(Bc - B - cG + 1)^2}.$$

6.6 Proof of Lemma 4

Impact of G

We have

$$\frac{d\pi(a=1)}{dG} = 1 - \mu^* + (1-G)\frac{d\mu^*}{dG} \quad (9)$$

$$\frac{d\pi(a=0)}{dG} = (1-B)\frac{d\mu^*}{dG} \quad (10)$$

Rearranging (1) leads to $\mu^* = \frac{G-c^*}{G-B}$. Hence$\frac{d\mu^*}{dG} = \frac{1}{G-B}\left(1 - \mu^* - \frac{dc^*}{dG}\right)$. We have $\frac{dc^*}{dG} = \frac{\partial \Phi}{\partial G}\left(1 - \frac{d\Phi}{dc}\right)^{-1}$ with $\frac{\partial \Phi}{\partial G} = \frac{(1-B)^2(1-c^*)}{1-B-(G-B)c^*}$. Combining $\mu^* = \frac{G-c^*}{G-B}$ with (3) yields $1 - \mu^* =$

$\frac{(1-B)(1-c^*)}{1-B-(G-B)c^*}$. Plugging this expression in $\frac{\partial \Phi}{\partial G}$ leads to $\frac{\partial \Phi}{\partial G} = \frac{(1-B)(1-\mu)}{1-B-c(G-B)}$. Hence

$$\frac{d\mu^*}{dG} = \frac{(1-\mu^*)}{(G-B)\left(1-\frac{d\Phi}{dc}\right)}\left(1-\frac{d\Phi}{dc}-\frac{(1-B)}{(1-B-c\,(G-B))}\right) \tag{11}$$

Then, we have $\Phi(c) = \frac{G(1-B)-(G-B)c}{(1-B)-(G-B)c}$. Inverting this function yields $c = \frac{(1-B)(G-\Phi)}{(G-B)(1-\Phi)}$. As $c = \Phi$ in equilibrium, we have

$$c^* = \frac{(1-B)(G-c^*)}{(G-B)(1-c^*)}. \tag{12}$$

Rearranging this equation, we get $1-B-(G-B)c^* = (1-B)(1-G)/(1-c^*)$. Substituting $1-B-(G-B)c^*$and $\frac{d\Phi}{dc} = -\frac{(G-B)(1-c^*)^2}{(1-B)(1-G)}$in (11) and rearranging, we obtain

$$\frac{d\mu^*}{dG} = \frac{(1-\mu^*)}{(G-B)\left(1-\frac{d\Phi}{dc}\right)}\left(\frac{(G-B)(1-c^*)^2-(G-c^*)\,(1-B)}{(1-B)(1-G)}\right).$$

Rearranging (12) leads to $(1-B)(G-c^*) = c^*(G-B)(1-c^*)$. Hence

$$\frac{d\mu^*}{dG} = \frac{(1-c^*)\,(1-\mu^*)}{\left(1-\frac{d\Phi}{dc}\right)(1-B)(1-G)}\,(1-2c^*)\,. \tag{13}$$

Plugging (13) in (9), we obtain

$$d\pi(a=1)/dG = \frac{(1-\mu^*)}{\left(1-\frac{d\Phi}{dc}\right)(1-G)\,(G-B)}\left((1-G)\,(c^*-B)+(1-c^*)^2\,(G-B)\right)$$

which is always positive: If $G > B$, $c^* - B < 0$, meaning the last term in bracket and the first are both negative. If $G < B$, they are both positive.

Turning next to the firms which do not self-regulate, we have:

$$\frac{d\pi(a=0)}{dG} = (1-B)\frac{d\mu^*}{dG}$$

which is positive if and only if $c^* < 1/2$. If we now substitute

$$c^* = \frac{1}{2\,(G-B)}\left(G-2B+1-\sqrt{(G-2B+1)^2-4G\,(1-B)\,(G-B)}\right)$$

in this inequality, calculations show that $c^* < 1/2$ is equivalent $G < 2/3$ if $B = 0$ and $B < 2/3$ if $G = 0$. As $B = 1 - (1-\alpha)^{n-1}$ when $G = 0$, $B < 2/3$ is equivalent to $\alpha < 1 - \frac{1}{3^{1/n}}$.

Impact of B

We have

$$\begin{aligned} \frac{d\pi(a=1)}{dB} &= (1-G)\frac{d\mu^*}{dB} && (14) \\ \frac{d\pi(a=0)}{dB} &= -\mu^* + (1-B)\frac{d\mu^*}{dB} && (15) \\ \frac{d\mu^*}{dB} &= \frac{1}{G-B}\left(\mu^* - \frac{dc^*}{dB}\right) && (16) \end{aligned}$$

We have $\frac{dc^*}{dB} = \frac{\partial\Phi}{\partial B}\left(1 - \frac{d\Phi}{dc}\right)^{-1}$ with $\frac{\partial\Phi}{\partial B} = \frac{c^*(1-G)^2}{(bc-B-cg+1)^2}$. Combining $\mu^* = \frac{G-c^*}{G-B}$ with (3) yields $\mu^* = \frac{c^*(1-G)}{1-B-(G-B)c^*}$. Hence $\frac{\partial\Phi}{\partial B} = \frac{(1-G)\mu^*}{1-B-(G-B)c^*}$. It follows that

$$\frac{d\mu^*}{dB} = \frac{\mu^*}{(G-B)\left(1-\frac{d\Phi}{dc}\right)}\left(1 - \frac{d\Phi}{dc} - \frac{(1-G)}{1-B-(G-B)c^*}\right).$$

As $1 - B - (G-B)c^* = (1-B)(1-G)/(1-c^*)$ and $\frac{d\Phi}{dc} = -\frac{(G-B)(1-c^*)^2}{(1-B)(1-G)}$, we have

$$\frac{d\mu^*}{dB} = \frac{\mu^*\left((c^*-B)(1-G) + (G-B)(1-c^*)^2\right)}{(G-B)(1-B)(1-G)\left(1-\frac{d\Phi}{dc}\right)}. \tag{17}$$

which is always positive as $G > B$ implies $c^* - B > 0$ and conversely. Therefore $\frac{d\pi(a=1)}{dB} > 0$.

Substituting (17) in (15) and rearranging, we obtain:

$$\frac{d\pi(a=0)}{dB} = \frac{\mu^*(1-c^*)}{\left(1-\frac{d\Phi}{dc}\right)(1-B)}(1-2c^*).$$

which is positive only if $c^* < 1/2$ and thus $\alpha < 1 - \frac{1}{3^{1/n}}$.