

Building a reputation as a socially responsible firm*

Aleix Calveras[†]

Juan-José Ganuza[‡]

This Draft: May 15th, 2014.

ABSTRACT

Many of the attributes that make a good “socially responsible” are credence attributes that cannot be learned by consumers either through search or experience. Consumers aggregate information about them from several channels (media, advertisement, NGOs, etc.). Since these sources may send contradictory messages, the information available to consumers is noisy. In this paper we model such informational environment and show the positive relationship between the accuracy of the information transmitted to consumers and CSR. We also show that firms may be tempted to adding noise to the information channel (e.g., through lobbying of the media), which might reduce the supply of the credence attributes and even harm firms themselves. As a consequence, firms might find profitable, for instance by means of forming a partnership with an NGO, to commit to not manipulate the information. Finally, we also show that such self-commitment by firms is a strategic substitute of transparency regulation by the public sector.

JEL classification numbers: D72, H42, L51, M14, Q52.

Key words: credence good, information asymmetry, corporate social responsibility, regulation

*We benefited from comments by audiences at FEDEA and Universidad de Salamanca. Aleix Calveras gratefully acknowledges financial support of the Spanish Ministry of Science and Technology under grant ECO2010-21393-C04-02. Juan-José Ganuza acknowledges the hospitality of FEDEA and the financial support of the the Barcelona GSE Research and the Spanish Ministry of Education and Science through project SEJ2006-09993/ECON.

[†]Department of Business Economics, Universitat de les Illes Balears. Cra Valldemossa Km 7 07122 Palma de Mallorca, Spain. E-mail: aleix.calveras@uib.es.

[‡]Department of Economics and Business, Universitat Pompeu Fabra. C/ Ramon Trias Fargas 25-27 08005 Barcelona, Spain. E-mail: juanjo.ganuza@upf.edu.

1 INTRODUCTION

Many of the attributes that make a good or a service “green” or, more generally, socially responsible, are credence attributes, not directly observable by consumers (Nelson, 1970; Tirole, 1988, pp. 106-129; Baron, 2009). In addition to their physical and performance characteristics, products have (unobservable) characteristics that consumers cannot learn either through search or experience: these are the so-called credence attributes of the good. Examples of such credence attributes are numerous: the conditions under which the product is produced, including any externalities associated with production (e.g. pollution), how workers are treated and how well they are paid, hidden hazards associated with consumption of the product, etc.

This asymmetry of information between the firm and consumers (and other stakeholders) regarding business’ practices poses a threat to the viability of corporate social responsibility (CSR).¹ CSR is mainly driven by demand by “conscious” consumers (namely, consumers who value one of these ‘socially responsible’ credence attributes, and are willing to pay a higher price for a good that includes such attribute) and thus the level of information available to consumers is key.² Absent credible information, the market might fail to provide the credence attributes valued by consumers: if consumers are uncertain about the attributes of the good, then they might not be willing to pay a premium for it and, thus, firms will not supply such attributes in the first place (Akerloff, 1974).

Thus, building a reputation for social responsibility is key in the development of CSR.³ The firm must convince, communicate consumers in a credible manner the presence of such (socially responsible) credence attributes in the goods and services that it sells. The purpose of this paper is precisely to study the role of informational issues in the promotion of CSR and the firm’s

¹By corporate social responsibility we denote those “voluntary actions that firms take over and above compliance with minimum legal requirements, to address both its own competitive interests and the interests of the wider society” (as defined by the UKs Department of Trade and Industry).

²Evidence regarding the valuation by some conscious consumers on some credence attributes (and who are willing to pay a higher price for them) can be found in relation to the labor conditions of a firm (Hiscox and Smyth, 2009), to charity linked products (Elfenbein and McManus, 2007), or the environmental goodness of a product (Casadesus-Masanell et al., 2009). See also, for instance, Mohr et al., 2001, and Murray and Volgel, 1997).

³We follow Cabral (2005) in that the seller’s reputation would be the buyer’s beliefs that the seller is good; namely, in our case, that it sells socially responsible goods.

reputation as socially responsible. More specifically, we want to analyze:

(i) In which way the level of information accuracy (or lack of it) in the market impact in the ability of the firm to build a reputation for social responsibility. There exist in the market several institutions designed to cope with the asymmetry of information, with the lack of information by consumers on the credence attributes supplied by firms (i.e., the social responsibility of firms' practices). These institutions include certifications, whether provided by a single firm (Bottega and de Freitas, 2008) or a group (club) of them (Baron, 2009); the information provided by activists such as NGOs (Feddersen and Gilligan, 2001) in the context of private politics (Baron, 2003); direct (advertisement) or indirect communication by firms; and information provided to consumers/citizens by the media (Dyck and Zingales, 2002). Previous literature mostly focuses on one of these institutions and assume that such institutions may solve completely the information asymmetry. We start by considering that consumers receive information from several channels and that they aggregate it in a noisy signal. The signal is noisy since it is likely that consumers receive contradictory messages from different sources (media, the firm, NGOs...). The accuracy of this signal is likely to vary across markets and, thus, it is important to assess in which way the quality of the signal impacts on the firm's ability to build a reputation for social responsibility.

(ii) How information about the behaviour of the firm is produced and the incentives of agents to provide it. The number of agents that (may) play a role in the transmission of information to consumers on businesses' practices are wide. In particular we focus on the incentives of the firm to manipulate the information provided to consumers (through advertisement, media, etc.) and the consequences it has on CSR.

(iii) The role of regulation and other institutional arrangements (such as commitment through external agents) to promote transparency and indirectly to foster CSR. An example of a transparency regulation is the European Union Directive 1999/94/EC which requires car makers to inform consumers on fuel economy and CO₂ emissions of each car. Such a Directive has the explicit aim both of increasing the consciousness of consumers and of allowing already conscious consumers to take informed decisions in accordance to their preferences, thus giving incentives to

firms to sell less polluting cars.⁴ Alternatively, we also study the possibility of individual commitment of firms to higher transparency; namely, to not manipulate the information (e.g., by means of forming some form of partnership between the firm and NGO). Finally, we also study the interaction between transparency regulation and such self-commitment.

Our analysis starts in a simple framework in which a firm chooses with which technology it produces a good: either a clean (and more expensive to the firm) technology that produces a unit of a public good for each unit of the good sold, or a dirty and cheaper technology which does not produce any unit of the public good. Even though each consumer is small (negligible) in relation to the market and might thus free-ride in its purchasing decision, we assume that consumers are, to a varying degree, "conscious" (or, alternatively derive a warm-glow from purchasing the good produced with the clean technology). As a consequence, a firm might have an incentive to invest in the clean technology if, afterwards, is able to charge a higher price to consumers. There is, however, asymmetric information between the firm and consumers with regards to the technology effectively used (the technology used by the firm is its own private information). Consumers, however, receive a noisy signal on the choice on which technology the firm is actually using.

Our framework thus allows us to discuss in which way the accuracy of the information that consumers receive influences the incentives of firms to be socially responsible. We show that, intuitively, the higher the accuracy of the information provided to consumers on the true technology used by the firm, the more a firm has incentives to build a reputation for social responsibility by investing in the clean technology (which, of course, makes it more likely that the equilibrium will be one with socially responsible businesses practices). This result also provides a rationale for many forms of transparency regulations present in the markets, such as the European Directive 1999/94/EC explained above. Additionally, we also show that the more consciousness consumers in the market are, the more incentives a firm has to invest in reputation as a socially responsible firm and, thus, it is more likely that there will be an equilibrium in which the firm chooses the

⁴More specifically, the labelling Directive requires the display of a label on fuel consumption and CO2 emissions on all new cars, the publication of national guides on the fuel efficiency of new cars, the display of posters at the dealerships and the inclusion of fuel efficiency information in printed promotional literature.

clean technology.

Next, we take a step backwards in the analysis of the informational channel between firms and consumers (and possibly other agents). More specifically, we analyze the incentives of firms to provide information and show that, if possible, firms would manipulate (e.g., through advertisement) the information provided to consumers, hence increasing the probability of a dirty technology to go undetected. Such manipulation, however, by decreasing the accuracy of the information provided to consumers (increase the noise of the signal they receive), might destroy the incentives of the firm to invest in reputation for being socially responsible. Thus, the possibility to manipulate might eliminate the possibility of an equilibrium with the clean technology.

Moreover, such manipulation might even harm the firm herself and, as a consequence, and in some cases, firms would like to be able, whenever possible, and as a way to build a reputation for being socially responsible, to commit ex ante not to manipulate the information provided to consumers (or, in other words, to increase the accuracy of the information provided to the public). For the firm, a way to commit might imply the involvement of the firm with a third party (such as a NGO) whose independent reputation might assure in a credible way that the firm does not manipulate the information provided to consumers and, thus, increase the accuracy of the signal received by consumers. This result might show one of the rationales behind many partnerships we see in real life between firms and NGOs; e.g., Starbucks with the environmental NGO Conservation International, or the partnership between the multinational fruit company Chiquita with the NGO Rainforest Alliance.

Finally, we also discuss the interaction between such a self-commitment by firms and transparency regulation by the public sector. We show in the paper that both institutional arrangements (that increase the incentives of the firm to be socially responsible) are strategic substitutes.

1.1 Related Literature

Like Besley and Ghatak (2007) and Bagnoli and Watts (2003), and without loss of generality, we model CSR as the provision of a public good. This is analogous to the supply of a credence

attribute in a joint manner with the good the firm is selling (as in Baron, 2009). Our paper focuses on the informational issues between firms and consumers, and also other actors in the economy (e.g., NGOs). Several papers with different approaches have analyzed this. One of such perspectives is the literature on certification and eco-labels. Bottega and de Freitas (2009), for instance, analyze certification by a third party (either a private firm or an NGOs) and the effect it may have on the scope for public regulation. An important difference with our analysis is that, in their framework, the certifier credibly informs (with certitude) about the credence attribute that the firm is supplying. Thus, in their framework, informational accuracy and manipulation is not an issue. The same, for instance, can be said about Baron (2009) which analyzes the supply of credence attributes by firm members of a club (a voluntary organization) that is in charge of verifying that the established standard is met.

The media and other actors (interest groups, firms themselves) do also play a role (in addition to third party certifiers) in the transmission of information to consumers about businesses practices. Generally, governments, interest groups (such as NGOs), and firms themselves generate and aggregate information that the media then process and selectively communicate to consumers/citizens. Dyck and Zingales (2002) provides both anecdotal and systematic evidence that media affect companies' policy towards the environment. They look at the effects of the press on the private sector's responsiveness to environmental concerns. They show (in accordance to our result in proposition 1) that the press (using as a proxy the circulation of daily newspapers normalized by population) has a positive effect on firm's responsiveness to environmental concerns. Baron (2003, 2005) discuss the private politics involved in such a process of aggregating, selecting and transmitting information in which a myriad of agents interact (firms, NGOs, etc.). Feddersen and Gilligan (2001) analyze in an incomplete information theory the role of an activist in the provision of information to consumers regarding the credence attributes of products (businesses' practices).

Although our framework is very simple, our paper is, to our knowledge, the first one to discuss the incentives of a firm to invest in reputation for selling socially responsible goods in a

context in which the firm can manipulate the information regarding business practices provided to consumers. There are many papers dealing with the manipulation of information by agents⁵, but our model incorporates an initial stage in which firms decide whether or not to make a “green” investment. In terms of modeling our framework is similar to Diamond (1989) in that the reputation mechanism includes both an adverse selection (the type of firm) as well as a moral hazard (the choice of the technology) problem. Our framework is, however, static, contrary to Diamond’s dynamic setup. In our case, then, the incentives of the firm to invest in reputation rests in the presence of signals that transmits inaccurate (noisy) information to consumers with regards to the choice of the firm. (As opposed to in Diamond (1989) where reputation acquisition relied on the information available to consumers with regards to the past behaviour of firms).

In section 2 we present the benchmark model which we solve in section 3 obtaining and discussing proposition 1 and 2, namely, the role of the accuracy of information and the consciousness of consumers in the promotion of CSR or, in other words, in the incentives of a firm to invest in reputation for being socially responsible. Next, in section 4, we analyze the provision of information by the firm and its incentives to manipulate such information, as well as its consequences. Then, in section 5 we analyze the incentives of a firm to commit not to manipulate, and also discuss its substitutability with transparency regulation by the public sector. We then conclude in section 6.

2 THE BENCHMARK MODEL

The model consists of a perfect competitive market in which firms sell an homogeneous good. One firm among all may differentiate from the others by attaching a credence attribute to this good. We consider that this firm may choose to produce with a cleaner (and more expensive) technology than the rest.⁶ In the economy there is a continuum of consumers who by consuming

⁵See for example, the cheap talk literature, where informed agents may lie (Crawford and Sobel, 1982; Farrell and Rabin (1996), etc.), and in persuasive games, where informed agents only may hide information (Dye, 1985; Glazer and Rubinstein, 2004; Grossman, 1981; Shin, 1994; etc).

⁶Other analogous examples might be better working conditions in the factories such as excluding child labor, or paying a wage above the market wage.

the good produced with the clean technology exert a positive externality of value G on each consumer in the market. The technology used is private information of the firm, while all consumers receive a noisy signal regarding the type of technology used by the firm.

2.1 Firms and Technologies

Competitive firms do not make profits, and the market price and the marginal cost of the homogenous good are normalized to 0. One firm (henceforth, the firm) may differentiate in the market by attaching a credence attribute to the good it sells. We model this by allowing the firm to choose a clean technology rather than the dirty technology with which the rest of the firms produce. The firm, thus, may choose with which technology t to produce, whether to produce the good with a clean technology (C) or with the dirty and standard technology (D), i.e. $t \in \{C, D\}$. The clean technology entails a fixed cost $F \geq 0$, whereas the dirty technology entails no fixed cost. In either case, the marginal cost of production is 0. If the firm uses the clean technology, it generates for each unit produced a positive externality of value G to all potential consumers of the market (producing with the dirty technology generates no positive externality).⁷

The firm that can differentiate may be one of three different types, depending on the size of the fixed cost in which it incurs in case it uses the clean technology. With ex ante probability $\frac{1-\theta}{2}$ the firm has a fixed cost of $F = \infty$ of choosing the clean technology. As a consequence, independently of market conditions, this type of firm will never choose the clean technology. With probability $\frac{1-\theta}{2}$, for the firm choosing the clean technology has a fixed cost of $F = 0$. This type of firm will always choose the clean technology. Finally, with a probability θ the firm can use the clean technology with a fixed cost of \hat{F} , with $0 < \hat{F} < \infty$. This firm, as we analyze below, will be 'strategic' regarding its decision on which technology to use. Notice that our analysis will focus on the relevant case in which the positive externality G is larger than the fixed cost \hat{F} , i.e. the clean technology is efficient. The firm learns its type before the choice of the technology, and

⁷Thus, consumption of a unit produced with the clean technology can be seen as either equivalent to the private contribution to a public good (as in Besley and Ghatak, 2007), or to the consumption of a good with a credence attribute attached to it (as in Baron, 2009).

consumers do not observe the type of the firm.

2.2 Consumers

There is a mass of unit 1 of consumers with utility

$$u = v + \alpha g - p,$$

where v is the valuation of the standard good, and g is the magnitude of the positive externality, with $g \in \{0, G\}$. More specifically, $g = G$ in case the the firm uses the clean technology, and 0 if the good consumed is produced by the firm using the dirty technology. Let α represent the type of the consumer with regard to his/her degree of consciousness (or altruism). Thus, for instance, α denotes the consumer's valuation of a clean environment or of consuming a good not produced with child labour. More specifically, we assume that α is distributed over the interval $[0, 1]$ with a distribution function $H(\cdot)$. We also assume that the reliability function of the distribution, *i.e.* $\overline{H}(\cdot) = 1 - H(\cdot)$, is logconcave. Finally, p is the price paid by the consumer.

2.3 Signals and Information

The technology effectively used by the firm is not observable by the consumers. Consumers, however, receive one of two signals concerning the technology used by the firm. The signal is s , where $s \in \{s_C, s_D\}$. The probability that consumers receive one or other signal does of course depend on the technology that the firm is using. Thus, the probability that the signal is s given that the technology chosen is t is $\Pr(s | t)$. More specifically,

$$\Pr(s_C | C) = 1,$$

$$\Pr(s_D | C) = 0,$$

$$\Pr(s_C | D) = 1 - \gamma$$

$$\Pr(s_D | D) = \gamma,$$

where $\gamma \in [0, 1]$. That is, if the firm uses the clean technology, the signal will be s_C with certainty. However, if the firm uses the dirty technology there is some noise and thus consumers may receive either signal. Notice that γ represents the accuracy of the signal, with a higher γ implying a more informative signal. More specifically, notice that with $\gamma = 0$ the signal is non-informative whatsoever since consumers never receive signal s_D .

2.4 The Market Game

The timing of the game is as follows.

1. Nature chooses the type of firm, namely, the level of the fixed cost F of the dirty technology.
2. The firm chooses the technology with which it is going to produce, $t \in \{C, D\}$. The rest of the firms produce and sell the standard good (at zero price and cost).
3. Nature chooses the signal $s \in \{s_C, s_D\}$ on the technology used by the firm according to the previous probabilities. All consumers receive the same signal.
4. The firm sets its price p .
5. Each consumer decides whether to buy or not from the firm. The alternative is to buy the standard good from the competitive fringe (at zero price).
6. Profits are realized.

3 THE MARKET EQUILIBRIUM

3.1 Demand

As usual, we solve the game backwards and, thus, start determining demand which depends on the marginal consumer α^* that is indifferent between buying the good to the firm or buying the standard good from the competitive fringe (at zero price). Namely:

$$v + \alpha^* \cdot \Pr(C | s)G - p = v. \tag{1}$$

Then,

$$\alpha^* = \frac{p}{\Pr(C | s)G}. \quad (2)$$

Thus, those consumers with $\alpha \geq \alpha^*$ will buy the "differentiated" good from the firm, while those with $\alpha < \alpha^*$ will buy the standard good. As a consequence, and given the distribution function of α over $[0, 1]$, the demand faced by the firm is $1 - H\left(\frac{p}{\Pr(C|s)G}\right)$.

We see that demand depends on the reputation of the firm for being socially responsible; namely, on the signal received by the consumer and the posterior probability that the technology chosen by the firm is the clean one. We analyze this further below.

3.2 Firm's Profits

The profit of the firm (gross of fixed cost, if any) is demand times price, namely,

$$\pi(p, s) = \left[1 - H\left(\frac{p}{\Pr(C | s)G}\right)\right] \cdot p. \quad (3)$$

Given this profit function, and given the signal received by all consumers, which is the price set by the firm, and its posterior profits? We answer this question in Lemma 1 next, where we see that the price (and profits) of the firm depends on the signal received by consumers:

LEMMA 1 *The price set by the firm is $p^*(s) = \Pr(C | s)Gr^*$, whereas firm's profits (gross of fixed costs, if any) are then $\pi^*(s) = [1 - H(r^*)]\Pr(C | s)Gr^*$, with $r^* = \frac{[1-H(r^*)]}{h(r^*)}$.*

On the one hand, if consumers receive signal s_D , the posterior probability that the technology chosen by the firm is the clean one is zero. As a consequence, and since all consumers are then homogeneous, the price then set by the firm is 0, i.e., the willingness that the consumer has to pay for the product when it believes that it is produced using the dirty technology (in which case consumption entails no public good, no positive externality). In such a case, the firm's profits are 0.

On the other hand, if consumers receive signal s_C , then they have a positive willingness to pay for the product and the price will no be longer 0. The optimal price characterized in Lemma

1, is linear in the conditional probability that the clean technology has been used. As we will show in the next section, this conditional probability will be increasing in the accuracy of the signal. Given this optimal price, the more activist consumers will buy from the firm, and the other consumers will buy from the (rest of the) market. It turns out that the marginal consumer (and consequently the demand) is independent of the information structure, and only depends on the distribution of the consumer "consciousness" parameter, $\alpha^* = r^*$. Firm's profit (gross of the fixed cost) are linear both in the equilibrium price and in the conditional probability that the clean technology has been used, $\pi^*(s) = \Pi \Pr(C | s)$, where $\Pi = [1 - H(r^*)] Gr^*$. We can interpret Π as the profits of the firm producing with a clean technology under perfect information; namely, the market share of the green good, $[1 - H(r^*)]$, multiplied by the willingness to pay of the marginal consumer, Gr^* .

3.3 The Perfect Bayesian Equilibrium

Given this above, which will be the technology chosen by the firm in equilibrium? Or, in other words, is the firm going to invest in reputation for being socially responsible by choosing the clean technology? Since such a choice is private information of the firm, consumers are going to have beliefs on such choice based on the signal received, hence solving the game requires solving for the perfect Bayesian equilibrium. As defined in Fudenberg and Tirole (1991), a perfect Bayesian equilibrium (PBE) is a set of strategies and beliefs such that, at any stage of the game, strategies are optimal given the beliefs, and the beliefs are obtained from equilibrium strategies and observed actions using Bayes rule. Note the link between strategies and beliefs: the beliefs are consistent with the strategies, which are optimal given the beliefs.

Clearly, the choice will depend on the actual type of the firm. In any case, the type with $F = \infty$ will always choose the dirty technology while the type with $F = 0$ will always choose the clean technology. We then have left to discuss what the strategic type with $0 < \hat{F} < \infty$ is going to do. We focus on pure strategies equilibrium, and we (need to) study when the strategic type is going to choose the clean technology (the socially responsible equilibrium) and when it is going

to choose the dirty technology (the not socially responsible equilibrium).

3.3.1 The Socially Responsible Equilibrium We denote as socially responsible the equilibrium in which the strategic type firm chooses the clean technology whereby building a reputation as a socially responsible firm. In equilibrium it must be that priors and beliefs are consistent with strategies. Then, in such a case, (and since the $F = 0$ type by assumption also chooses the clean technology), priors (beliefs of consumers) are that the clean technology is chosen with probability $\Pr(C) = \frac{1-\theta}{2} + \theta = \frac{1+\theta}{2}$, and the dirty technology is chosen with probability $\Pr(D) = \frac{1-\theta}{2}$.

For the firm to be optimal to choose the clean technology over the dirty one, it must be that the expected profits (before the realization of the public signal) when choosing the clean technology are larger than the profits using the dirty technology, i.e.

$$\pi(t = C) \geq \pi(t = D). \quad (4)$$

Let $\pi_C^*(s) \equiv \Pi \Pr_C(C | s)$ be the strategic firm's profits in the case in which it chooses the clean technology (whereas as π_D^* will stand for the case in which the strategic type will choose the dirty technology) and the realization of the public signal on the technology chosen by the firm is s . (Recall from above that $\Pi = [1 - H(r^*)] Gr^*$).

Using bayes rule, we obtain $\pi_C^*(s_D) = 0$ and $\pi_C^*(s_C) = \frac{\Pi(1+\theta)}{2-\gamma(1-\theta)}$. This profit function reflects the consumers' willingness to pay for the good. This willingness to pay is 0 if the realization of the signal is s_D since in this case the consumers learn that the technology was the dirty one with certainty. When the realization of the signal is s_C then, the willingness to pay is increasing on the accuracy of the signal γ .

The expected profits are

$$\begin{aligned} \pi(t = C) &= \Pr(s_C | C) \pi_C^*(s_C) + \Pr(s_D | C) \pi_C^*(s_D) - \widehat{F} \\ &= \pi_C^*(s_C) - \widehat{F}, \end{aligned}$$

since $\Pr(s_C | C) = 1$ and $\Pr(s_D | C) = 0$. On the other hand,

$$\pi(t = D) = \Pr(s_C | D) \pi_C^*(s_C) + \Pr(s_D | D) \pi_C^*(s_D)$$

$$= (1 - \gamma)\pi_C^*(s_C)$$

Then, the condition over expected profits becomes

$$\begin{aligned}\pi_C^*(s_C) - \widehat{F} &\geq (1 - \gamma)\pi_C^*(s_C), \\ \gamma\pi_C^*(s_C) &\geq \widehat{F}.\end{aligned}$$

Plugging in the expression of the profits, we obtain the necessary condition for the socially responsible equilibrium

$$\gamma\pi_C^*(s_C) = \frac{\gamma(1 + \theta)\Pi}{2 - \gamma(1 - \theta)} \geq \widehat{F},$$

where recall that $\gamma \in [0, 1]$. Notice that a necessary condition for the socially responsible equilibrium is that the fixed cost F is smaller than Π .

3.3.2 The Not Socially Responsible Equilibrium The analysis is analogous to the previous one. Suppose now that the strategic type firm chooses the dirty technology. Then, in such a case, priors (beliefs by consumers) that a strategic firm chooses the clean technology are $\Pr(C) = \frac{1-\theta}{2}$, whereas that a firm chooses the dirty technology are $\Pr(D) = \frac{1-\theta}{2} + \theta = \frac{1+\theta}{2}$. As we specified above, for notational purposes, let $\pi_D^*(s_C) = \Pi \Pr(C | s)$ be the profit of the firm (gross of fixed costs) when the signal is good (s_C) and the strategic type firm has chosen the dirty technology. Using the bayes rule, we obtain $\pi_D^*(s_C) = \frac{(1-\theta)\Pi}{2-\gamma(1+\theta)}$. These profits are linked to the positive willingness to pay of consumers that is increasing in the accuracy of the signal. As in the previous case, when the realization of the public signal is bad the profits are zero, i.e $\pi_D^*(s_C) = 0$. Finally, the necessary condition for the strategic type to choose the dirty technology is that

$$\pi(t = D) \geq \pi(t = C). \tag{5}$$

For the same arguments than above, this is equivalent to

$$\begin{aligned}(1 - \gamma)\pi_D^*(s_C) &\geq \pi_D^*(s_C) - \widehat{F}, \\ \gamma\pi_D^*(s_C) &\leq \widehat{F}.\end{aligned}$$

Plugging in the expression of the profits, we obtain the necessary condition for the not socially responsible equilibrium

$$\gamma\pi_D^*(s_C) = \frac{\gamma(1-\theta)\Pi}{2-\gamma(1+\theta)} \leq \widehat{F}.$$

3.4 Results

Once we have characterized the equilibrium of the game, we can analyze the impact of the accuracy of the consumers' information in the payoff of the strategic firms and in the likelihood that the necessary conditions specified above are satisfied.

LEMMA 2 *The expected profits of the strategic firm in the socially responsible equilibrium (non socially responsible equilibrium) are increasing (decreasing) in the accuracy of the signal γ .*

Lemma 2 shows that the accuracy has a positive (negative) effect when the strategic firm chooses the clean (dirty) technology. Then, the necessary condition for the socially responsible equilibrium (non socially responsible equilibrium) is more (less) likely to be met when the accuracy of the signal is large. The proposition states this result in the case in which $\Pi \geq \widehat{F}$, since if the fixed cost is too large, the only equilibrium is the not socially responsible one, for all levels of accuracy.

PROPOSITION 1 *When the level of accuracy γ is small, the only equilibrium is that the firm is not socially responsible (chooses the dirty technology), whereas when the level of accuracy is large enough, the only equilibrium is that the firm is socially responsible (invests in reputation by choosing the clean technology). For intermediate levels of accuracy, both strategies are part of an equilibrium.*

The intuition is straightforward. Only when there is enough information regarding the action followed by the firm, consumers will be willing to pay a higher price for a good labeled as socially responsible. Given that consumers' behavior depends on the information they hold, the firms' incentives to undertaking socially responsible actions also does. Mathematically, we can see this

analysing the effect of γ on the probability that the technology used is the clean one when the signal is s_C (see Figure 1). We see that this probability increases with γ , and accordingly does the willingness to pay of the consumer. As a consequence, the incentives of the firm to be socially responsible (to choose the clean technology) are higher. As a consequence, we can state that the better information there is regarding the firm's practices, the more incentives the firm has to build a reputation as a socially responsible firm.

[INSERT FIGURE 1]

We can also analyse the impact of the level or degree of 'consciousness' in society on the strategic choice of the firm regarding technology (the incentives to invest in reputation). This we do in the following proposition, measuring the degree of consciousness in society according to the first-order stochastic dominance. Let $H'(\cdot)$ be an alternative distribution of α such that H' that first-order stochastically dominates H , $H'(\cdot) \geq H(\cdot)$ for all $\alpha \in [0, 1]$.

PROPOSITION 2 *In a more 'conscious' society, the firm is more likely to invest in the clean technology.*

Propositions 1 and 2 have two important implications over the total welfare in the economy.

COROLLARY 3 *Total surplus (welfare) is increasing in the likelihood of a socially responsible equilibrium. Namely, (i) for a given level of information accuracy γ , the total surplus is increasing in the level of 'consciousness' of society; and (ii) for a given distribution of consciousness H , the total surplus is increasing in γ .*

All consumers derive utility from the firm adopting the clean technology, even though only a part of them are willing to pay for it (knowing that their individual decision would not have an effect over the firm's decision). We, however, do not consider that activist consumers may obtain additional utility for their prosocial behavior.

4 ENDOGENOUS INFORMATION STRUCTURES

In this section we consider that the accuracy of the public signal will depend on the behavior of the firm. In particular, we consider in a reduced form model that the ratio between good and bad news about the firm determines the probability of signal realization: the probability of receiving a good realization s_C (a bad realization s_D) is $P(s_C) = \frac{N_G}{N_B+N_G}$, ($P(s_D) = \frac{N_B}{N_B+N_G}$) where N_G (N_B) are good (bad) news.

News are produced by the firm and other private actors (such as NGOs, the media, etc.). It is natural to assume that the firm will only provide good news. The other actors, on the other hand, may either be biased in favor of the firm (for example, due to advertisement or lobbying), or will convey truthful information (such as NGOs). We can explain our initial information structure using this simple reduced form. On the one hand, when the firm chooses the clean technology, then $\Pr(s_C | C) = 1$, since all agents will provide good news since there is no conflict of interest between the biased (the firm) and the unbiased (NGOs) actors. On the other hand, when the firm chooses the dirty technology, then $\Pr(s_C | D) = 1 - \gamma$, because there is a conflict between neutral and biased actors, and hence there will be good and bad news.

In this section thus we study the incentives of the firm to manipulate the information, and the impact that such manipulation has on the equilibrium and on welfare, including on the own profits of the firm.

4.1 Manipulation

Consider the case in which the firm can increase the amount of good news when the technology used is the dirty one. We start by assuming that this manipulation process is costless. This means in our model that the firm can increase $\Pr(s_C | D)$, namely, it decreases the accuracy of the information available to consumers, decreasing γ to γ' . Then, the timing of the game goes as follows:

1. Nature chooses the type of firm, F .
2. The firm chooses the technology with which it is going to produce, $t \in \{C, D\}$. The rest of the firms produce and sell the standard good (at zero price and cost).
3. The firm either manipulates or not, $m \in \{M, NM\}$, which is private information of the firm. In case of manipulation, the firm increases $\Pr(s_C | D)$, from $1 - \gamma$ to $1 - \gamma'$, where $\gamma' < \gamma$.
4. Nature chooses the signal $s \in \{s_C, s_D\}$ on the technology used by the firm according to the information structure determined by $\Pr(s_C | D)$. All consumers receive the same signal.
5. The firm sets its price p .
6. Each consumer decides whether to buy or not.
7. Profits are realized.

4.1.1 Equilibrium with manipulation Lemma 2 above shows that the expected profits of the strategic firm decrease with accuracy when it is using a dirty technology, and increase when it is using the clean technology. For the same token, manipulation is going to be used only by firms using a dirty technology. If the firm uses the clean technology, it is not interested in manipulation because, by assumption, the signal will be the clean one. If the firm uses the dirty technology manipulating is good for the firm's profits because taking (in equilibrium) the value of the signals as given, manipulation increases the probability that the signal is good, s_C . In short, then, manipulation makes relatively more attractive the dirty technology, and this implies that it is more difficult to find an equilibrium in which the strategic firm chooses the clean technology. The next Lemma states this result.

PROPOSITION 4 *The possibility of manipulation reduces the likelihood of a socially responsible equilibrium. The new necessary condition is $\frac{\gamma'(1+s)\Pi}{2-\gamma'(1-s)} \geq F$.*

In particular, if the fixed cost of the strategic types lies over the interval $[\frac{\gamma'(1+s)\Pi}{2-\gamma'(1-s)}, \frac{\gamma(1+s)\Pi}{2-\gamma(1-s)}]$ the possibility of manipulation makes the socially responsible equilibrium not feasible. The intuition

goes as follows. When manipulation is costless, the non strategic type with $F = +\infty$ will always manipulate. This worsens the value of the good signal and consequently makes less attractive the clean technology. Moreover, as we said above, the possibility of manipulation makes it more profitable to deviate from the socially responsible equilibrium since it increases the probability that the signal is s_C without incurring in the fixed cost. Both effects lead to make it more difficult that the necessary conditions for the socially responsible equilibrium are met, and thus, it is less interesting for the firm to invest in reputation for social responsibility.

It is interesting to know who benefits from the possibility of manipulation. Consumers are clearly worse off since it is less likely than the good equilibrium arises. The non strategic type with $F = 0$ and the strategic type with \hat{F} such that it chooses a clean technology independently of manipulation, decrease their profits due to manipulation since it reduces the value of the good signal. The non strategic type with $F = +\infty$ and the strategic type with \hat{F} such that it chooses a dirty technology independently of manipulation, increase their profits due to manipulation since increasing the probability a good signal overcomes that the good signal is less valuable.

It is specially interesting the case in which the possibility of manipulation makes unfeasible the socially responsible equilibrium when it was possible without such manipulation. In such a situation, there are strategic types that choose a clean technology when manipulation is not possible, but when this possibility exists, they choose a dirty technology and reduce the accuracy to γ' afterwards. This, however, does not mean that these types are better off with manipulation; in fact, next Lemma states that some of these strategic types are worst off when manipulation is a possibility.

LEMMA 3 *The possibility of manipulation may lead a strategic firm (that, absent such possibility, would choose the clean technology) to choose the dirty technology and then manipulate. Furthermore, some of these firms (may) show a reduction in their profits due to such manipulation.*

4.2 Transparency regulation

We have taken as given the outcome of the manipulation process that leads accuracy from γ to γ' and we have also assumed that this process was costless. In practice, however, manipulation may involve a cost, and γ' should be the solution to an optimization problem. More specifically, as the cost of manipulation is likely to increase with the stringency of transparency regulatory measures, it is also likely then that more stringent transparency regulatory policies reduce the net benefits of manipulation, hence reducing the distance between γ and γ' . As a result we are able to state the following corollary:

COROLLARY 5 A more stringent transparency regulation that reduces the distance between γ and γ' increases the likelihood of a socially responsible equilibrium.

That is, making transparency regulation more strict, increases the costs of manipulating the information and, as a consequence, makes such manipulation less likely and a socially responsible equilibrium more likely to occur.

5 COMMITMENT THROUGH EXTERNAL AGENTS

The previous results imply that some firms may favor an increase in the transparency regarding their business practices (the choice of technology in our model). Such higher transparency would have the effect of, on the one hand, increase the incentives of firms to adopt responsible business practices and, on the other hand, might also increase the own profits of the firm.

A higher transparency might be achieved through different institutional settings (e.g., through regulatory policies). In this section, however, we focus first on the incentives that some firms might have to commit unilaterally to a higher transparency; namely, to commit not to manipulate the information that will be available to consumers regarding the technology choice of the firm. As we discuss in the following subsection, an example of such a business strategy can be found in the many partnerships between firms and NGOs whose logic lies in the role of the NGO as a an

independent and reputable certifier of some of the firm's business practices.

To undertake such an analysis, consider the previous model with an additional period in which firms may choose whether or not to commit to not manipulate. Moreover, it is natural to assume that this commitment is visible by consumers and that they update their valuation of the product using the commitment decision as well as the realization of the signal. Then, the timing of this new signaling game goes as follows:

1. Nature chooses the type of firm, F .
2. The firm decides whether or not to commit to not manipulate, $d \in \{Co, NCo\}$. This commitment decision is visible.
3. The firm chooses the technology with which it is going to produce, $t \in \{C, D\}$. The rest of the firms produce and sell the standard good (at zero price and cost).
4. If the firm has not committed in period 2, it decides whether to manipulate or not, $m \in \{M, NM\}$, which is private information of the firm. In case of manipulation, the firm increases $\Pr(s_C | D)$, from $1 - \gamma$ to $1 - \gamma'$, where $\gamma' < \gamma$.
5. Nature chooses the signal $s \in \{s_C, s_D\}$ on the technology used by the firm according to the information structure determined by $\Pr(s_C | D)$. All consumers receive the same signal and they update their beliefs using the signal realization and the (visible to all) commitment decision.
6. The firm sets its price p .
7. Each consumer decides whether to buy or not.
8. Profits are realized.

5.1 Equilibrium with commitment and manipulation

We have to characterize the perfect bayesian equilibrium of the game, the equilibrium strategies for every type $\{(d_0^*, t_0^*, m_0^*), (d_{\widehat{F}}^*, t_{\widehat{F}}^*, m_{\widehat{F}}^*), (d_{\infty}^*, t_{\infty}^*, m_{\infty}^*)\}$ and the belief system $\lambda(d, s) = \{\lambda(0 | d, s), \lambda(\widehat{F} | d, s), \lambda(\infty | d, s)\}$, where (d_F^*, t_F^*, m_F^*) is the equilibrium strategies of type F , and $\lambda(F | d, s)$ is the posterior belief of consumers of the firm being type F when a decision d and a signal realization s have been observed by consumers. Notice that in the previous sections we have focused on the posterior probability, $P(C|s)$, now using the equilibrium strategies and the beliefs, we can construct $P(C|s, \lambda)$ which is key in the analysis. Since we can have a conflict between the realization of the signal and extreme values of the beliefs, we will assume that independently of the beliefs, a bad signal realization always reveals the type of technology, $P(C|s_D, \lambda) = 0$.

LEMMA 4 *The perfect bayesian equilibria in pure strategies must be a pooling equilibria in which all types of firms follow the same strategy regarding the commitment decision.*

The intuition why a separating equilibrium does not exist is as follows. There does not exist a separating equilibrium in which the type $F = \infty$ chooses a different strategy than the other two types since, independently of the realization of the signal, it would not get any profits and mimicking the type $F = 0$ generates positive profits. In summary, $F = \infty$ has incentives to be with the other two types. For the opposite argument, there does not exist a separating equilibrium in which the type $F = 0$ chooses a different strategy than the other two types, since the other two types have incentives to mimic $F = 0$. Finally, there does not exist a separating equilibrium in which the strategic type chooses a different strategy than the other two types: the strategic type, depending on the parameters, has the same preferences than $F = 0$ or $F = \infty$, which implies that the previous arguments apply.⁸

⁸The previous arguments and Proposition 4 depend on the fact that we assume that it is costless to commit to not manipulate. It is very intuitive to see that if we introduced a cost of committing, separation equilibria might arise. For example, consider a constellation of parameters such that firms with $F = 0$ and $F = \widehat{F}$ choose the clean technology and committing has a cost c . Then, if $\pi_C^*(s_C) - \widehat{F} \geq c \geq (1 - \gamma)\pi_C^*(s_C)$ we could build a separating equilibrium in which firms with $F = 0$ and $F = \widehat{F}$ would commit to not manipulate, whereas a firm with $F = \infty$ would not commit. Notice that such an equilibrium would have $\lambda(\infty | NCo, s) = 1$.

Focusing thus in the pooling equilibria, we can construct two pooling equilibria in which the three types choose either to commit or not to commit, beliefs on the equilibrium path are the priors and consumers have a belief outside of the equilibrium path $\lambda(\infty | d, s) = 1$ (namely, a deviating firm is $F = \infty$). However, it seems less natural an equilibrium in which “good” firms lose the opportunity to differentiate through commitment. In fact, as we state in the next proposition, this equilibrium may not pass the intuitive criterium of Cho-Kreps.

PROPOSITION 6 *The pooling equilibrium in which all types of firms follow a commitment strategy is the only perfect bayesian equilibrium that satisfies the intuitive criterium of Cho-Kreps for all parameter values.*

The pooling equilibria requires beliefs out of the equilibrium path that give more weight (probabilities) to the types that will produce with a dirty technology. However, the intuitive criterium of Cho & Kreps allows us to show that good types ($F = 0$ and strategic types producing with the clean technology) may find it optimal to deviate from the pooling equilibrium in which all the types choose *NCo*. The intuitive criterium establishes two conditions to eliminate a perfect bayesian equilibrium. First, that the type that deviates obtains larger profits out of the equilibrium path if it is identified for his true type. In our case, if for example, type $F = 0$ would obtain larger profits by choosing *C* and being identified as type $F = 0$, that it is obtaining with *NCo* and pooled with $F = \infty$.

Second, and applying the requirement to our model, the bad types ($F = \infty$ and strategic types producing with the dirty technology) may obtain larger profits in the equilibrium path (*NCo*) that out of the equilibrium path for all possible beliefs. In our case, this implies that they must obtain larger profits by being pooled with the good types and having the opportunity to manipulate than by committing to not manipulate (obtaining with lower probability a low signal) with the more favorable beliefs $\lambda(0 | C, s_C) = 1$ (that is, being taken as an $F = 0$), this second condition is satisfied only if manipulating is a sufficiently attractive, that is, if $\gamma - \gamma'$ is large enough.

5.2 Commitment through external agents vs transparency regulation

As we explained above, there exist alternative institutional arrangements to cope with the problems derived from the possibility of manipulating the information by the firm. Namely, we have mentioned transparency regulation by the public sector and self-commitment by the firm through external agents. The analysis in the previous subsection allows us to discuss the interaction between such institutional arrangements in the following corollary:

COROLLARY 7 Transparency regulation that reduces the distance between γ and γ' is a strategic substitute of self-commitment to transparency by firms.

We do this statement in the sense that only when $\gamma - \gamma'$ is large enough, we can guarantee that the only equilibrium in pure strategies (that satisfies the intuitive criterium) is the pooling equilibrium in which all types of firms follow a commitment strategy.

5.3 Partnerships between firms and NGOs

The previous analysis shows that, under some circumstances, and so as to be able to build a reputation for being socially responsible, firms may have incentives to commit not to be able to manipulate the information available to consumers with regards to the technology used in production. In other words, to increase the transparency with regards to its choice of technology. It is our opinion that such result lies behind the rationale for many of the partnerships between a firm (or a group of firms) and an NGO (or a group of them) that we observe in many industries.⁹ According to Yaziji (2004), one of the strengths that NGOs have (as apposed to corporations) is 'legitimacy'¹⁰. As Yaziji (2004) explains, and according to a poll conducted by the Edelman public relations firm, both Americans and Europeans said they found NGO spokespeople more

⁹The rationales behind such partnerships may vary and are not confined to the rationale presented in this paper. For instance, Brugman and Prahalad (2007) discuss such alliances between a firm and an NGO for the purpose of developing some entrepreneurship and business model in the developing world. Such a partnership allows firms and NGOs to share some knowledge and capabilities that are specific to each one.

¹⁰The other three are awareness of social forces, distinct networks and specialized technical expertise.

credible than either a company's CEO or Public Relations representative. Some fraction of the public, specially in Europe, sees NGOs as dedicated first and foremost to serving an aspect of the general social welfare. This is what gives credibility to their positions regarding social issues as, e.g., are the environmental ones. Such 'legitimacy' is precisely the reason why in our framework NGOs can be used by firms as a way to commit to the public (consumers) to a certain course of action and can increase the transparency of their actions. Furthermore, Yaziji (2008) also stresses that partnering with NGOs, and advertising it, can draw stricter scrutiny from the public, the press, the regulators, and so on than your company formerly received. Notice that such effect of partnering with an NGO is analogous to increasing transparency in our framework, making it more difficult to manipulate the information.

Examples of such partnerships between a firm and an NGO abound. In the garment industry, for instance, the firm GAP, in its aim to try to ensure a proper treatment of workers in the factories that are part of its supply chain, provides two independent evaluations of GAP's factory inspection program by the NGOs Social Accountability International and Verite. In another example, the multinational firm Starbucks has developed a partnership with an NGO, environmental group Conservation International, with the aim of increasing transparency in their operations and assuring that the operations were done under sound conditions.

6 CONCLUDING REMARKS

The rise in the importance of the phenomena of corporate social responsibility that has taken place in the last 15 or 20 years is inextricably linked to both an increase in the "consciousness" of markets (consumers, investors, workers) with regards to social and environmental issues, as well as an increase in the transparency of the market and non-market behaviour of firms. Our focus in this paper has been the study of the role that informational issues play in the promotion of CSR and, more specifically, in the incentives that a firm has to build a reputation for social responsibility.

The framework for the analysis has been one with a demand by consumers of a good with some

socially responsible credence attributes (e.g., that the good be produced with a green technology). Since such an attribute is not directly observable by consumers, demand and willingness to pay must depend on firm's reputation which, in turn, depends on some indirect information: in our set-up a signal that all consumers receive on regards to the technology used by the firm. This signal is a reduced form modelling of the information that consumers receive about firms (e.g., through media). The first result in our paper is quite intuitive: the better the information available to consumers, the more consumers are willing to pay for a good labeled as 'green', and, accordingly, the more incentives firms have to adopt a 'green' mode of production. In addition to numerous anecdotal evidence, there is some empirical evidence in support of this intuitive result. In Dyck and Zingales (2002) it is shown that a higher diffusion of the press in a country (a construct of better available information to consumers on firms' practices) implies a higher responsiveness of firms towards environmental issues, that is, firms are more likely to be 'green'.¹¹

Next, we acknowledge that the availability of information to consumers on the firm's practices is endogenous, namely, it is dependent on many actors' behaviour. Such information is dependent not only on media behaviour (the press, TV), but also other stakeholder's behaviour, such as NGOs, activist shareholders and institutional shareholders, financial analysts, and the information provided by the firm itself. Hence, the second part of our analysis has endogenized the information available to consumers by allowing the firm to manipulate such information in a way that decreases the accuracy of the signal received. We show that, as a result of such manipulation capability, an equilibrium with socially responsible business practices becomes less likely. Since consumers know that the information they have is likely to have been manipulated, they are less willing to pay a premium for the supposedly 'green' product. As a consequence, a firm then has less incentives to build a reputation for social responsibility by adopting the (more costly) 'green' mode of production.

More surprising, though, is the result that some of these firms end-up worse-off because of their possibility to manipulate the information provided to consumers. This is so because such

¹¹See also Xia et al. (2008) for empirical evidence on the impact of media freedom in the adoption by firms (and the corresponding global diffusion) of the environmental ISO 14001 certification.

manipulation possibility destroys a socially responsible equilibrium in which the firm provided the good with the credence attribute (green mode of production), and the consumer paid a premium price. As a consequence of the decrease in profits due to the manipulation possibilities, these firms (the ones that, absent the possibility to manipulate, would invest in reputation for social responsibility) would favor any measure that ties their hands and impedes them to manipulate the information; in other words, a measure that increases the transparency in the market regarding the firm's mode of production. Alternative institutional arrangements may play such role, for instance, transparency regulation by the public sector (e.g., the European Union Directive 1999/94/EC which requires car makers to inform consumers on fuel economy and CO₂ emissions of each car).

We mostly focus our analysis, however, on a decentralized solution to increase market transparency of business practices, namely, the observed partnerships between firms and NGOs (e.g., GAP and the NGO Verité; the multinational fruit company Chiquita and Rainforest Alliance, Starbucks and Conservation International, etc). While acknowledging that such partnerships may serve several purposes, we provide a rationale behind such alliances in the way that an independent and reputable NGOs may be capable of credibly communicating consumers that the information, eventhough maybe still noisy, has not been manipulated by the firm. We show that when such partnerships become available to firms, the only intuitive (à la Cho-Kreps) perfect bayesian equilibrium in pure strategies is that in which such partnerships are formed with the purpose of firms to credibly and visibly commit in the eyes of consumers to not manipulate the information they receive. In such a way, transparency in the market is increased, and a socially responsible equilibrium becomes more likely. Finally, we also discuss the interaction between such decentralised institutional arrangement (partnerships) and transparency regulation and show that they are strategic substitutes. Namely, this means that a more strict transparency regulation by the public sector will make self-commitment by firms less likely.

A APPENDIX

PROOF OF LEMMA 1: As $\overline{H(\cdot)} = 1 - H(\cdot)$ is logconcave, $\pi(p, s)$ is quasiconcave on p (see Bagnoli and Bergstrom (2005)) and the optimal price is given by the first order condition

$$-h\left(\frac{p^*}{\Pr(C|s)G}\right) \cdot \frac{p^*}{\Pr(C|s)G} + 1 - H\left(\frac{p^*}{\Pr(C|s)G}\right) = 0$$

Taking $r = \frac{p}{\Pr(C|s)G}$, the solution is characterized by $r^* = \frac{[1-H(r^*)]}{h(r^*)}$. As r^* is a feature of the distribution H and it is independent of $\Pr(C|s)$ and p , we can characterize the optimal price and firm's posterior profits as functions of r^* .

$$p^*(s) = \Pr(C|s)Gr^* \text{ and } \pi^*(s) = [1 - H(r^*)] \Pr(C|s)Gr^*$$

■

PROOF OF LEMMA 2:

i) The expected profits of the strategic firm in the socially responsible equilibrium are:

$$\pi(t=C) = \pi_C^*(s_C) - \widehat{F} = \frac{(1+\theta)\Pi}{2-\gamma(1-\theta)} - \widehat{F}$$

if we take the derivate over the accuracy of the signal γ , we obtain:

$$\frac{d\pi(t=C)}{d\gamma} = \frac{-(1-\theta)(1+\theta)\Pi}{(2-\gamma(1+\theta))^2} > 0$$

ii) The expected profits of the strategic firm in the not socially responsible equilibrium are:

$$\pi(t=D) = (1-\gamma)\pi_D^*(s_C) = \frac{(1-\gamma)(1-\theta)\Pi}{2-\gamma(1+\theta)}$$

if we take the derivate over the accuracy of the signal γ , we obtain:

$$\frac{d\pi(t=D)}{d\gamma} = \frac{-(1-\theta)(1-\theta)\Pi}{(2-\gamma(1+\theta))^2} < 0$$

■

PROOF OF PROPOSITION 1: The characterization of the equilibrium depends on the both condition stated in the main text.

$$\begin{aligned} \gamma\pi_C^*(s_C) &= \frac{\gamma(1+\theta)\Pi}{2-\gamma(1-\theta)} \geq \widehat{F}, \\ \gamma\pi_D^*(s_C) &= \frac{\gamma(1-\theta)\Pi}{2-\gamma(1+\theta)} < \widehat{F}. \end{aligned}$$

Notice that as $\gamma \in [0, 1]$, $\pi_C^*(s_C) - \pi_D^*(s_C) = \frac{4\theta(1-\gamma)\Pi}{(2-\gamma(1-\theta))(2-\gamma(1+\theta))} \geq 0$. In particular, $\pi_C^*(s_C) > \pi_D^*(s_C)$ if $\gamma < 1$, and the profits' difference converges to 0 when $\gamma = 1$. Moreover, $\pi_C^*(s_C)$ and $\pi_D^*(s_C)$ are increasing on γ . These results implies that if $\Pi > \widehat{F}$, there exist two values of γ , $\gamma^* < \gamma^{**}$ such that, for $\gamma < \gamma^*$ only the condition $\gamma\pi_D^*(s_C) < \widehat{F}$ holds, whereas for $\gamma > \gamma^{**}$ only the condition $\gamma\pi_C^*(s_C) < \widehat{F}$ is satisfied. For intermediated values of γ , both conditions hold. ■

PROOF OF PROPOSITION 2: $\Pi' = [1 - H'(r^{**})]Gr^{**}$ where $r^{**} \in \arg \max\{[1 - H'(r)]Gr\}$. Notice that

$$\Pi' = [1 - H'(r^{**})]Gr^{**} \geq [1 - H'(r^*)]Gr^* \geq [1 - H(r^*)]Gr^* = \Pi.$$

This implies that, $\pi_C^*(s_C) = \frac{(1+\theta)\Pi'}{2-\gamma(1-\theta)} \geq \pi_C^*(s_C) = \frac{(1+\theta)\Pi}{2-\gamma(1-\theta)}$. For the same token, $\pi_D^*(s_C) \geq \pi_D^*(s_C)$. The proposition follows for the previous results and the equilibrium conditions. ■

PROOF OF COROLLARY 3: Given that the size of the market of the clean firm is fixed $r^* = \frac{[1-H(r^*)]}{h(r^*)}$. Surplus only depends of the likelihood of implementing the clean technology. Hence the results are just an application of Propositions 1 and 2. ■

PROOF OF PROPOSITION 4: As the non strategic type with $F = \infty$ will manipulate, the value of a good signal, $\pi_C^*(s_C)$, decreases from $\frac{(1+\theta)\Pi}{2-\gamma(1-\theta)}$ to $\frac{(1+\theta)\Pi}{2-\gamma'(1-\theta)}$. Moreover, manipulation make more attractive to deviate since the good signal is more likely, then the condition of the equilibrium becomes:

$$\pi_C^*(s_C) - F \geq \Pr(s_C | D) \pi_C^*(s_C) = (1 - \gamma')\pi_C^*(s_C)$$

then

$$\gamma'\pi_C^*(s_C) = \frac{\gamma'(1+\theta)\Pi}{2-\gamma'(1-\theta)} \geq \widehat{F}$$

This condition is more difficult to satisfied that $\frac{\gamma(1+\theta)\Pi}{2-\gamma(1-\theta)} \geq \widehat{F}$, because $\frac{\gamma(1+\theta)\Pi}{2-\gamma(1-\theta)} \geq \frac{\gamma'(1+\theta)\Pi}{2-\gamma'(1-\theta)}$

PROOF OF LEMMA 3: Take for example the limit firm with fixed cost equal to $\frac{\gamma'(1+\theta)\Pi}{2-\gamma'(1-\theta)} + \varepsilon$. the payoff of this firm in the clean equilibria is $\frac{(1+\theta)\Pi}{2-\gamma(1-\theta)} - \frac{\gamma'(1+\theta)\Pi}{2-\gamma'(1-\theta)}$ which is higher than its payoff in the dirty equilibrium $\frac{(1-\gamma')(1-\theta)\Pi}{2-\gamma'(1+\theta)}$

$$\frac{(1+\theta)\Pi}{2-\gamma(1-\theta)} - \frac{\gamma'(1+\theta)\Pi}{2-\gamma'(1-\theta)} > \frac{(1-\gamma')(1+\theta)\Pi}{2-\gamma(1-\theta)} > \frac{(1-\gamma')(1-\theta)\Pi}{2-\gamma'(1+\theta)}$$

The last inequality, follows from

$$\frac{(1 - \gamma')(1 + \theta)\Pi(2 - \gamma'(1 + \theta))}{(2 - \gamma(1 - \theta))(2 - \gamma'(1 + \theta))} > \frac{(1 - \gamma')(1 - \theta)\Pi(2 - \gamma(1 - \theta))}{(2 - \gamma'(1 + \theta))(2 - \gamma(1 - \theta))}$$

Then

$$(1 + \theta)(2 - \gamma'(1 + \theta)) > (1 - \theta)(2 - \gamma(1 - \theta))$$

which simplifies to

$$2(1 + \theta) - \gamma'(1 + 2\theta + \theta^2) > 2(1 - \theta) - \gamma(1 - 2\theta + \theta^2)$$

and finally

$$(\gamma - \gamma')(1 + \theta^2) + 2\theta(2 - \gamma' - \gamma) > 0$$

This is true given that $1 \geq \gamma > \gamma'$. ■

PROOF OF COROLLARY 5: The proof follows straightforward from proposition 4. ■

PROOF OF LEMMA 4: In order to rule out the partially separating equilibrium (given that we have two possible strategies and three types) we have to consider several cases.

- i) Consider that $F = 0$ chooses $d_0 = Co$ and the other two types $d_{\hat{F}} = d_{\infty} = NCo$. Thus $P(C|s_C, \lambda(s_C, Co)) = 1$. There are two cases. If the strategic type in equilibrium chooses a dirty technology, then $P(C|s_C, \lambda(s_C, NCo)) = 0$ and both types $F = \hat{F}$ and $F = \infty$ are willing to mimic $F = 0$. If the strategic type chooses a clean technology, then it would increase its profits by mimicking $F = 0$ and choosing $d_{\hat{F}} = Co$.
- ii) Consider that $F = 0$ chooses $d_0 = NCo$ and the other two types $d = Co$, then $P(C|s_C, \lambda(s_C, NCo)) = 1$. Thus the other two types are better by deviating and mimicking $F = 0$.
- iii) Consider that $F = \infty$ chooses $d_{\infty} = NCo$ and the other two types $d = Co$, then $P(C|s_C, \lambda(s_C, NCo)) = 0$ and the profits of $F = \infty$ will be 0. Therefore, $F = \infty$ would be better by deviating and mimicking the other two types.
- iv) The case in which $F = \infty$ chooses $d_{\infty} = Co$ and the other two types $d = NCo$ is completely analogous to case iii).

v) Consider now that the strategic type chooses a different strategy than the other two, there are two possibilities. First, the strategic type chooses a clean technology, then $P(C|s_C\lambda(s_C, d_{\hat{F}})) = 1$ and $F = 0$ would prefer to mimic the strategic type, as in the last case of i). Second, the strategic type chooses a dirty technology but then, the posterior belief will be $P(C|s_C\lambda(s_C, d_{\hat{F}})) = 1$ and its profits will be 0. In this case, the strategic type would be better by deviating and by mimicking the other two types, as in case iii).

Finally, the pooling equilibrium on any strategy C and NC is always an equilibrium because we can have arbitrary beliefs out of the equilibrium path, in particular $\lambda_\infty(s_C, \cdot) = 1$. ■

PROOF OF PROPOSITION 6: The pooling equilibrium in which all the types choose Co , posterior beliefs on equilibrium path are equal to priors beliefs and the belief out of the equilibrium path is $\lambda_\infty(NCo, \cdot) = 1$, is a perfect Bayesian equilibrium and satisfies the intuitive criterion of Cho-Kreps. The first part, is direct because independently of the realization of the signal, the profits of the firm when choosing NCo are 0. Then, all the types prefer Co to NCo . Moreover, this equilibrium satisfies the Cho-Kreps criterion since the maximum payoff of type $F = \infty$ are achieved with NCo and $\lambda_0(NCo, \cdot) = 1$, then we cannot rule out the belief $\lambda_\infty(NCo, \cdot) = 1$.

The pooling equilibrium in which all the types choose NCo , posterior beliefs on equilibrium path are equal to priors beliefs and the belief out of the equilibrium path is $\lambda_\infty(Co, \cdot) = 1$, is a perfect Bayesian equilibrium but for some parameters does not satisfy the intuitive criterion of Cho-Kreps. In particular, consider that parameters are such that the strategic type chooses a clean technology and does not have incentives to manipulate. Then, the equilibrium payoff of the $F = \infty$ type on the equilibrium path is $\frac{(1-\gamma')(1+\theta)\Pi}{2-\gamma'(1-\theta)}$. Then, if $\frac{(1-\gamma')(1+\theta)\Pi}{2-\gamma'(1-\theta)} > (1-\gamma)\Pi$, type $F = \infty$ will be better in the equilibrium path than choosing Co with the most favorable belief $\lambda_0(Co, \cdot) = 1$. This rule out $\lambda_\infty(Co, \cdot) = 1$ and consequently the previous pooling equilibrium. ■

PROOF OF COROLLARY 7: This is due to the fact that the condition for ruling out the pooling equilibrium on NCo , $\frac{(1-\gamma')(1+\theta)\Pi}{2-\gamma'(1-\theta)} > (1-\gamma)\Pi$, is likely to be satisfied when $\gamma - \gamma'$ is large, since both sides of the inequality are decreasing. ■

REFERENCES

- [1] Akerlof, G. (1970), "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism", *Quarterly Journal of Economics*, 84, 3, 488-500.
- [2] Bagnoli, M. and T. Bergstrom (2005), "Log-concave probability and its applications", *Economic Theory*, 26, 2, 445-469.
- [3] Bagnoli, M and S. Watts (2003), "Selling to Socially Responsible Consumers: Competition and the Private Provision of Public Goods", *Journal of Economics & Management Strategy*, 12, 3, 419-445.
- [4] Baron, D. (2003), "Private Politics", *Journal of Economics and Management Strategy*, 21, 31-66.
- [5] Baron, D., (2005), "Competing for the public through the news media", *Journal of Economics and Management Strategy*,.
- [6] Baron, D., (2009), "Credence attributes, voluntary organizations, and social pressure", working paper.
- [7] Bendell, J., "Chiquita", Case Studies, Business and Sustainable Development, BSD-global.com, as visited March 2010.
- [8] Besley, T. and M. Ghatak (2007), "Retailing public goods: The economics of corporate social responsibility", *Journal of Public Economics*, vol. 91(9), pages 1645-1663, September.
- [9] Bottega, L and J. de Freitas (2009), "Public, Private and Nonprofit Regulation for Environmental Quality", *Journal of Economics and Management Strategy*, Volume 18 Issue 1.
- [10] Bruggmann, J. and CK Prahalad, (2007), "Cocreating Business's New Social Compact", *Harvard Business Review*, February.
- [11] Cabral, L., (2005), "The economics of trust and reputation: a primer", mimeo.

- [12] Cashore, B., G. Auld, and D. Newson, (2005), *Governing through Markets: Forest Certification and the Emergence of Non-State Authority*, New Haven, CT: Yale University Press.
- [13] Casadesus-Masanell, R., M. Crooke, F. Reinhardt, and V. Vasishth, (2009), "Households' Willingness to Pay for 'Green' Goods: Evidence from Patagonia's Introduction of Organic Cotton Sportswear", *Journal of Economics and Management Strategy*. 18: 203-233.
- [14] Crawford, V.P. and J. Sobel (1982), "Strategic Information Transmission", *Econometrica*, 50, 1431-1451.
- [15] Diamond, D., (1989), "Reputation acquisition in debt markets", *Journal of Political Economy*, vol. 97 n 41.
- [16] Dyck, A. and L. Zingales (2002), "The Corporate Governance Role of the Media", in R. Islam: *The Right to Tell: The Role of Mass Media in Economic Development*. Washington, D.C.: World Bank, 2002, pp. 107-40.
- [17] Dye, R.A (1985), "Disclosure of Nonproprietary Information", *Journal of Accounting Research*, 23, 123-145.
- [18] Elfenbein, D. W. and B. McManus (2007), "A Greater Price for a Greater Good? Evidence that Consumers Pay More for Charity-Linked Products", working paper, University of North Carolina.
- [19] Farrell, J and M. Rabin (1996), "Cheap Talk", *Journal of Economic Perspectives*, 10, 3, 103-118.
- [20] Feddersen, T.J. and T.W. Gilligan (2001), "Saints and Markets: Activists and the Supply of Credence Goods," *Journal of Economics and Management Strategy*, 10(1), 149–171.
- [21] Fudenberg, D. and J. Tirole (1991), *Game Theory*, MIT Press.
- [22] Glazer, J. and A. Rubinstein (2004), "On Optimal Rules of Persuasion", *Econometrica*, 72, 1715-1736.

- [23] Grossman, S.J. (1981), “The Informational Role of Warranties and Private Disclosure about Product Quality”, *Journal of Law and Economics*, 24, 561-483.
- [24] Heal, G. (2008), *When principles pay*, Columbia University Press.
- [25] Hiscox, Michael J. and Nicholas F.B. Smyth (2006), “Is There Consumer Demand for Improved Labor Standards? Evidence from Field Experiments in Social Product Labeling.” Working paper, Harvard University.
- [26] Mohr, L.A., D.J. Webb, and K. Harris (2001), “Do Consumers Expect Companies to Be Socially Responsible? The Impact of Corporate Social Responsibility on Buying Behavior,” *Journal of Consumer Affairs*, 35(1), 45–72.
- [27] Murray, K. and C. Vogel, (1997), “Using a Hierarchy of Effects Approach to Gauge the Effectiveness of Corporate Social Responsibility to Generate Goodwill Toward the Firm: Financial versus Nonfinancial Impacts,” *Journal of Business Research*, 38.
- [28] Nelson, P. (1970), “Information and consumer behaviour”; *Journal of Political Economy* 78: 311-329.
- [29] Shin, H.S (1994), “The Burden of Proof in a Game of Persuasion”, *Journal of Economic Theory*, 64, 253-264.
- [30] Tirole, J. (1988), *The theory of industrial organization*, MIT Press.
- [31] Xia, J., J. Wang, Y. Wang, and Xing, R. (2008), ”Stakeholder pressures and the global diffusion of the ISO 14001 initiative: a resource dependence perspective”, *International Journal of Sustainable Society*, vol. 1, n^o 1.
- [32] Yaziji, Michael, 2004: “Turning gadflies into allies”, *Harvard Business Review* February 2004.

[natheight=7.4996in, natwidth=9.9998in, height=3.6348in, width=4.8369in]
C:/Users/uib/AppData/Local/Temp/graphics/Figure1-R₁.jpg