

Informative advertising by an environmental group

Pim Heijnen

Received: 10 August 2009 / Accepted: 17 April 2012 / Published online: 10 May 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract Consuming a product does not (necessarily) reveal the environmental damage it may cause. In terms of environmental damage, most goods are credence goods. Therefore, advertising and pricing rarely can transmit such information effectively to consumers. This article considers the scope with which an environmental group (EG) can signal this information to consumers via advertising and campaigning. Although an EG can inform the public, the welfare effects are ambiguous. Therefore, the EG may not have an incentive to acquire information about the firm's production practices. However, when the firm chooses the level of environmental damage and the EG chooses whether to acquire information, there exists a mixed-strategy equilibrium for some parameter values, such that the EG acquires information, and the threat of advertising causes the firm to produce an environmentally friendly good.

Keywords Environmental group · Environmental quality · Monopoly · Informative advertising · Asymmetric information · Information acquisition

JEL Classification L12 · L30 · Q50

1 Introduction

Consuming a product does not necessarily reveal its environmental quality. If consumers care about the environmental quality of the good (and there is evidence that they do; [Kuhn 2005](#)), then how does this information get transmitted to consumers?¹

¹ Environmental quality is a credence good. See [Darby and Karni \(1973\)](#) and [Emons \(1997, 2001\)](#).

Three possible groups could provide this information: the firm that sells the good, the government, and environmental groups (i.e., a non-governmental organization that aims to improve the state of the environment). Starting with the firm, [Mahenc \(2007, 2008\)](#) demonstrates under which conditions a firm can use the price of a good to signal environmental quality. In this case, the marginal cost of producing a good should be increasing with its environmental quality, and the firm's profit should be increasing with consumers' perception of environmental quality. Another instrument at the firm's disposal is ecolabeling, which [Ibanez and Grolleau \(2008\)](#) show firms can use to inform consumers. Government-firm interactions also have been studied extensively. For example, [Maxwell et al. \(2000\)](#) study the effect of (future) mandatory pollution regulation, [Amacher et al. \(2004\)](#) provide a theoretical analysis of ecolabeling, and [Petraakis et al. \(2005\)](#) examine information provision as a public policy instrument.

In contrast, the role of environmental groups (EGs) as information providers has not received much attention. Most literature considers how EGs influence the behavior of firms, such as through the effects of boycotts ([Baron 2003](#); [Innes 2006](#)) or the threat of boycotts ([Diermeier and Baron 2007](#)). This focus has neglected the myriad of activities that EGs engage in. In particular, EGs communicate with consumers in various ways, such as through advertisements or encouraging media coverage of their own activities (see [Baron 2005](#)). EGs generally try not only to persuade the consumer to adopt the organization's viewpoint, but also to provide facts and information that the consumer may not have been aware of.

As an example, consider the 'Viswijzer' (Dutch for fish guide) advertising campaign which features as its centerpiece an informative brochure that tells consumers which fish are on the brink of extinction, such that they may avoid consuming those species. Extensive advertising on Dutch television by the World Wildlife Fund promotes the informative brochure.² However, the product itself (i.e., the fish) does not display this kind of information, which means that this campaign is clearly distinct from (eco)labeling. Another example falls under a somewhat broader definition of advertising, namely Greenpeace's attempt to inform consumers about toxic ingredient in Disney branded pajamas.³ The strategy that this EG used was to (illegally) attach labels to pajamas, warning the consumers of the toxins present. This kind of guerilla advertising is clearly distinct from the conventional definition of labeling, in that the action is neither sanctioned by Disney nor are there any controls over the accuracy of the label. In both these examples, the visible activities by the EG require time, effort, and money. Therefore, the EG can use such instruments to signal the severity of the situation.

To model this phenomenon, I extend the standard industrial organization framework of vertical product differentiation to environmental damage and asymmetric information about that damage (i.e., consumers are unaware of the level of damage). I first consider a case in which the firm cannot signal the level of damage through price or

² This information, written in Dutch, can be found at www.wnf.nl, under *Natuurbescherming/Campagnes/Leven in de zee 2006/Viswijzer/*. All available fish guides can be found at http://wwf.panda.org/what_we_do/how_we_work/conservation/marine/sustainable_fishing/sustainable_seafood/seafood_guides/; South Africa and Switzerland offer English language versions.

³ See www.greenpeace.org.uk.

advertising (by assuming that marginal costs are independent of the damage level). I then introduce an EG that might choose to advertise with a certain intensity. The analysis reveals that the EG can signal the level of damage through its advertising. The EG chooses to advertise if a firm is producing a good with a high level of damage. Advertising harms “bad” firms instead of praising “good” ones.⁴ *Ex ante* advertising benefits both the firm and the consumers. Although in this respect information transmission by the EG is useful, I show that from an *ex ante* perspective, the EG itself does not necessarily benefit. Even if I drop the assumption that marginal cost is independent of damage level, the conclusions do not qualitatively change, though this extension does identify when the EG is more efficient at signaling than the firm, namely, when environmental problems are not severe. Finally, I consider a variant of this model, in which environmental damage is chosen by the firm, and information acquisition by the EG is endogenized. Under certain conditions, there exists an equilibrium where the firm randomizes between low and high damage and the EG acquires information (with positive probability). The EG advertises if it acquires information and the firm chooses high damage. Therefore in all three specifications I investigate, advertising campaigns by the EG occur in equilibrium.

My approach is related to several strands of the literature. First, the informative advertising model I propose relates to the auditing model presented by [Feddersen and Gilligan \(2001\)](#), who study a duopoly situation in which firms are price-takers. In their model, the EG can investigate a firm and discover how damaging the production process is, and then transmit this information to consumers. However, Feddersen and Gilligan suppose that information must be gathered, whereas I propose that information primarily needs to be disseminated. Both approaches emphasize the difficulty the EG faces when it tries to convince the public. Because Feddersen and Gilligan study costless investigations, their model relies on the concept of cheap talk to transmit information. The possibility of cheap talk represents an artifact of their model. [Crawford and Sobel \(1982\)](#) investigate the conditions under which an informative equilibrium arises in general cheap talk games and conclude that though such an equilibrium can exist, it is only possible if the disagreement between the receiver and the sender is not too large. Applied to the situation here, though the EG and consumer may agree that low environmental damage is preferable, consumers also want to keep price as low as possible. Feddersen and Gilligan eliminate this factor by making the firms price-takers.

Second, [Baron \(2005\)](#) examines a situation in which EGs (and firms) cannot directly communicate with consumers: All information is dispersed through the news media. The results indicate that the media are biased and the EG has an incentive to conceal information that is unfavorable to its interests. I study direct communication between

⁴ [Diermeier and Baron \(2007\)](#) assume that EGs hurt bad firms. In the model presented here this follows directly from the incentive constraints of the EG. But because there are examples of positive advertising by EGs, this focus tells only part of the story. For example, the ‘Viswijzer’ not only indicates which fish to avoid, but also suggests alternatives. Greenpeace has advertised in favor of the CFC-free refrigerator Greenfreeze, originally produced by a small German company and later adopted by mainstream manufacturers. However, because a viable alternative or better technology often is not available, this article focuses mainly on negative advertising.

EGs and consumers instead and show that there are situations in which the EG would like to conceal information.

Third, the model I propose is similar to the model used in [Heijnen and Schoonbeek \(2008\)](#), but they study persuasive advertising. In their model, consumers are aware of the level of environmental damage, and the EG uses advertising to increase the marginal disutility it causes to consumers. In equilibrium, firms preempt the EG's advertising campaign by producing a cleaner good. In the final variant of my model in which firms can choose the level of environmental damage, I observe preemption (to some degree) as well.

Fourth, my approach relates to the theoretical industrial organization literature where prices and advertising signal product quality ([Bagwell and Ramey 1988](#); [Milgrom and Roberts 1986](#)). [Bagwell and Ramey \(1991\)](#) and [Hertendorf and Overgaard \(2001\)](#) also address the case of two competing senders of information, similar to the model presented here.⁵ However, their senders are two firms, whereas I investigate a firm and an EG as senders. In turn, this research contributes to the signaling literature by investigating signaling by an outsider, namely, the EG that becomes informed and may benefit from passing this information on to consumers.

The remainder of this article is organized as follows: Sect. 2 introduces the model. Then in Sect. 3, I present the equilibria of the model, before outlining the results in Sect. 4, including the effect of advertising on the EG, the firm, and consumers. Section 5 introduces differences in marginal costs, and then I consider endogenous information acquisition in Sect. 6. Section 7 concludes.

2 Model

Consider a market in which a single, profit-maximizing firm sells one good. This good is characterized by the damage d it causes to the environment per product sold. *Ex ante* with probability ρ , damage is low, $d = d_L$, and with probability $1 - \rho$, damage is high, $d = d_H$. Assume that $0 < d_L < d_H$ and $0 < \rho < 1$. The firm and the EG know the true value of d , but consumers, who care about the damage the product causes, only know the distribution of d . The price of the good is p .

In the short run, the production technology is fixed. Irrespective of d , it costs $c > 0$ to produce one unit of the good. Usually, marginal cost is assumed to be decreasing in damage level. By abstracting from this, I focus purely on the communication problem. As a consequence of this assumption, the firm's profit level does not depend on the true value of d . Therefore the single crossing property, a necessary condition for the existence of an equilibrium in which the firm signals its type with the price of the product, cannot hold. The firm cannot convince consumers that it is a low damage firm (cf. [Mahenc 2008](#)).

⁵ [Kihlstrom and Riordan \(1984\)](#) consider a case with an infinite number of senders of information and find that signaling through advertising is possible in this case, but only if the high quality good has lower marginal costs 'on average'. Although the Porter hypothesis ([Porter and van den Linde 1995](#)), which states that environmentally friendly goods are cheaper to produce, has received a lot of attention, practically all studies assume that environmentally friendly goods are more costly ([Cremer and Thisse 1999](#); [Mahenc 2008](#); [Moraga-González and Padrón-Fumero 2002](#)). The tentative conclusion is that more competitive markets hamper the signaling of environmental damage by firms.

Consumers are indexed by θ , where θ is uniformly distributed on the interval $[0, \bar{\theta}]$. Without loss of generality, let $\bar{\theta} = 1$. Consumer θ has the following utility function:

$$U(\theta, p) = \begin{cases} V - \theta\mu - p & \text{if one unit of the good is bought and} \\ 0 & \text{if zero units are bought,} \end{cases} \tag{1}$$

where $V > c$ is a positive constant, θ measures the disutility of environmental damage, and μ is the expected (perceived) damage level associated with the consumption of one unit of the good. This is the standard vertical differentiation model used for products that differ in environmental quality (e.g. Cremer and Thisse 1999, Moraga-González and Padrón-Fumero 2002).

Consumers individually have no impact on total environmental damage (i.e., damage per product consumed times demand), but they have a lower utility if they buy a good with higher perceived damage, due to the guilt associated with buying a dirtier good. All consumers have the same information, which they process in the same manner. After they have observed the actions of the firm and the EG, consumers have the *ex post* belief that damage is low with a probability ϕ and $\mu = \phi d_L + (1 - \phi)d_H$. If they receive no new information, then $\phi = \rho$. The relation between ϕ and μ is strictly decreasing, and utility is linear in expected damage. Therefore, there is no loss of generality to consider only μ .

Only consumers with a sufficiently low θ will buy the good. The indifferent consumer $\hat{\theta}$ can be found by equating utility to zero:

$$\hat{\theta} = \frac{V - p}{\mu}. \tag{2}$$

The demand for the good $q(p, \mu)$ is $\hat{\theta}$. I assume an interior solution: there are always some consumers who do not buy the good because they care about the environment, but there are also always some consumers who do buy the good. Therefore, demand ranges between 0 and 1.

The profit of the firm is $\Pi(p, \mu) = q(p, \mu)(p - c)$. It then can be easily verified that the optimal price is given by:

$$p^* = \frac{1}{2}V + \frac{1}{2}c. \tag{3}$$

If consumers could observe the damage of the product, the price chosen by the firm would be the optimal price, which does not depend on μ .⁶ Finally, observe that:

$$q(p^*, \mu) = \frac{(V - c)}{2\mu} \quad \text{and} \tag{4}$$

⁶ This lack of dependency is a result of two assumptions. First, the marginal cost of producing the good does not depend on d . Second, the lower boundary of the distribution of θ is 0. Therefore, even if the EG signals environmental damage, the price that the firm sets will be completely uninformative in that both the low damage firm and the high damage firm choose p^* as the price.

$$\Pi(p^*, \mu) = \frac{(V - c)^2}{4\mu}. \quad (5)$$

Therefore, demand and profit are decreasing in μ . Furthermore, both functions are convex in μ .

Assume that an environmental group knows the true damage per product and cares about the total environmental damage caused by the production of the good. The cost of the environmental damage to the EG is $q(p, \mu)g$, where $g = g_H$ if $d = d_H$ and $g = g_L$ if $d = d_L$. I assume $\frac{g_H}{g_L} \geq \frac{d_H}{d_L}$.⁷ The EG and consumers agree that damage is either high or low, but the EG reacts more strongly, because its dislike of high damage is stronger than that of the consumers. Simultaneously with the firm's price decision, the EG can launch an advertising campaign, that consists of a message ("The firm is of type d ") and has cost x . When advertising is observed by consumers, they are able to infer the cost of the advertising campaign. For example, if consumers see an advertisement run on prime-time television, they note that it must have cost the EG a great deal of money. The cost of the advertising campaign is common knowledge.

The EG tries to minimize the sum of expected total damage and expected advertising cost: $q(p, \mu)g + x$. Note that demand q may depend indirectly on x through the beliefs of consumers. For example, consumers might think that a costly advertising campaign signals a dirty product and therefore update their expectations about the damage level upward. The members of the EG are activists who never buy the product and are willing to exert effort to determine how damaging the good really is. I assume that the EG has ample funds to launch an advertising campaign, but also wants to spend money as efficiently as possible. Therefore, the EG aims to minimize costs.

It is important to highlight that the actual message is not important, whereas the cost of advertising is. It is *not* the objective of the EG to maximize consumers' surplus. The public's skeptical view of the EG stems from non-alignment of the public interest and the EG's interest. The EG always benefits if consumers believe that the product is more damaging, so consumers might not trust the EG. To simplify the analysis (and without loss of generality), I assume that the advertising content is cheap talk, but that by showing how much it is willing to spend, the EG can try to convince the public that it is telling the truth.

The timing of the game is as follows:

- Period 0: Nature draws the damage level per unit sold d .
- Period 1: The firm and the EG observe d . The EG chooses the intensity of advertising x . The firm chooses its price p .
- Period 2: Consumers observe x and p . They update their beliefs μ (following Bayes' rule whenever possible) and choose to buy the good or not, which results in demand $q(p, \mu)$.

The equilibrium concept used is perfect Bayesian equilibrium, which has two main requirements: The firm and the EG must choose optimal actions, given the beliefs of consumers, and the consumers' beliefs must be consistent with these actions. Whereas

⁷ I thank an anonymous referee for this suggestion.

most studies consider only one sender of information (cf. [Milgrom and Roberts 1986](#)), here there are two senders (the firm and the EG).

2.1 Discussion of assumptions

Admittedly, the model presented here is simple. Specific functional forms simplify the analysis, but the existence of separating and pooling equilibria can be derived under more general conditions. Even if I allowed for a more general specification, the crucial assumption in this model is that the marginal costs do not depend on the actual damage of the firm. This assumption is not totally at odds with reality: the additional cost of producing an environmentally friendly good may be largely sunk, such as the investment in plants and infrastructure or R&D costs.⁸ The merit of an assumption of equal marginal cost is that in a full information setting, the price of the clean good and the price of the dirty good will be the same. In the absence of price effects, the EG's advertising can then be discussed purely in terms of its informational effect. However, in Sect. 5, I also consider the case in which marginal costs depend on the level of damage.

3 Equilibria

The forces that shape the set of equilibria are as follows. Intuitively, the EG wants to convince consumers that damage per unit sold is high, but the firm wants the opposite. If, in equilibrium, consumers believe that damage is low, then the cost of convincing consumers that damage is high should be great enough to discourage the EG from choosing this action. One way of 'punishing' a deviation by the EG is to assume that the damage expected by consumers is sufficiently low. In contrast, the firm's profit is at its highest when consumers believe that it sells a low damage product, and deviations by the firm are averted by assigning a sufficiently high expected damage level to these deviations. Therefore, the beliefs that sustain an equilibrium assign low expected damage levels to deviations from the EG and high expected damage levels to deviations from the firm.

Two types of equilibria can be distinguished: separating and pooling. A separating equilibrium is one in which consumers can distinguish between each state of the world, i.e., $\mu = d_L$ if $d = d_L$ and $\mu = d_H$ if $d = d_H$. Because consumers must be able to differentiate the two situations, $(x_L, p_L) \neq (x_H, p_H)$, where $x_L = x(d_L)$, $x_H = x(d_H)$, $p_L = p(d_L)$, and $p_H = p(d_H)$ denote advertising levels and prices for both states of the world.

In a game with two senders, the EG and the firm, there are two-sided separating equilibria (TSE) and one-sided separating equilibria (OSE). [Bagwell and Ramey \(1991\)](#) define these concepts as follows: In a TSE, both senders emit state-contingent signals, that is, $x_L \neq x_H$ and $p_L \neq p_H$. In an OSE, only one sender emits a state-contingent signal. There are two distinct OSE: one in which the firm signals ($p_L \neq p_H$ and $x_L = x_H$) and one in which the EG signals ($x_L \neq x_H$ and $p_L = p_H$).

⁸ As an example, consider the development of a fuel-efficient engine. Designing the engine is costly, but the production costs are unlikely to differ much.

I am interested in the effect of advertising by EGs, so the OSE in which the EG informs consumers is the obvious one to explore. Because the firm’s profit does not depend on actual damage, the single crossing property is not satisfied for the firm, and equilibria in which the firm signals do not exist. As a benchmark I use the case in which it is impossible to advertise (i.e., $x_L = x_H = 0$), and the firm does not inform consumers. Formally, this benchmark equilibrium coincides with a pooling equilibrium.

In the remainder of this article I use the notation $\mu(x, p)$ to denote the belief about the expected value of d if consumers observe an advertising level of x and a price of p .

3.1 The benchmark

The benchmark of no advertising is a pooling equilibrium in which neither the firm nor the EG provides the consumer with information. Consequently, in each state of the world, they choose the same action. For the EG, I choose out of equilibrium beliefs, such that the EG does not want to advertise, $\forall x > 0 : \mu(x, p) = d_L$. Hence $p_L = p_H = \hat{p}$ and $x_L = x_H = 0$. Then, $\mu(0, \hat{p}) = \rho d_L + (1 - \rho)d_H \equiv \bar{\mu}$. Consider the incentives of the firm when profit does not depend on the true value of d . In the pooling equilibrium, the constraint on out-of-equilibrium beliefs to rationalize the firm’s action is:

$$\frac{(V - \hat{p})(\hat{p} - c)}{\bar{\mu}} \geq \frac{(V - p)(p - c)}{\mu(0, p)} \quad \text{for all } p \neq \hat{p}, \tag{6}$$

which can be rewritten as:

$$\mu(0, p) \geq \frac{(V - p)(p - c)}{(V - \hat{p})(\hat{p} - c)} \bar{\mu}. \tag{7}$$

If the RHS of (7) exceeds d_H , then \hat{p} cannot be part of an equilibrium. If $\hat{p} = p^*$, then the RHS is always smaller than or equal to $\bar{\mu} < d_H$. Thus, there is a pooling equilibrium in which $\hat{p} = p^*$.

There are also other pooling equilibria. Suppose $\hat{p} \neq p^*$ and the deviation is p^* . For \hat{p} to be part of an equilibrium, it must be that

$$\frac{(V - p^*)(p^* - c)}{(V - \hat{p})(\hat{p} - c)} \bar{\mu} \leq d_H, \tag{8}$$

i.e., the RHS of (7) should not exceed d_H for the (most tempting) deviation p^* . Let I be the interval for \hat{p} implied by (8). If $\hat{p} \in I \subset [c, V]$, it can be part of a pooling equilibrium. Note that $p^* \in I$ and the equilibrium in which $p = p^*$ is the natural one to select.⁹ Summarizing:

⁹ A formal refinement that selects this equilibrium is the notion of undefeated equilibrium. Both types of firm prefer the equilibrium in which $p_L = p_H = p^*$. To quote Riley (2001, p. 451) “if an out-of-equilibrium action is observed, the revised beliefs should be fully consistent with some other equilibrium.” The notion of undefeated equilibrium requires $\mu(0, p^*) = \bar{\mu}$. Suppose the price in the pooling equilibrium is $\hat{p} \neq p^*$. In equilibrium, beliefs must be consistent, so we get $\mu(0, \hat{p}) = \bar{\mu}$. Note that

Proposition 1 (Benchmark) *Given out-of-equilibrium beliefs $\mu(x, p) = d_L$ for all $x > 0$, there is a perfect Bayesian equilibrium in which the firm sets its price equal to p^* in both states of the world and the EG never advertises.*

3.2 Equilibrium with advertising

In an OSE in which the EG informs and the firm sets the same price in each state of the world, beliefs must be such that $\mu(x_L, p_L) = d_L$ and $\mu(x_H, p_H) = d_H$. Moreover, in this kind of equilibrium $p_L = p_H$. Specifically, the price that the firm will set must be p^* in both states of the world. Suppose, to the contrary, that $p_L = p_H = \hat{p} \neq p^*$. For $d = d_H$, the deviation to p^* must not be profitable, and thus

$$\frac{(V - \hat{p})(\hat{p} - c)}{d_H} \geq \frac{(V - p^*)(p^* - c)}{\mu(x_H, p^*)}. \tag{9}$$

Because $\hat{p} \neq p^*$, we know that $(V - \hat{p})(\hat{p} - c) < (V - p^*)(p^* - c)$, in which case it must be that $d_H < \mu(x_H, p^*)$, which is impossible. Hence $\hat{p} = p^*$. Take $\mu(x_H, p) = d_H$ for all $p \neq p^*$. These out-of-equilibrium beliefs ensure that it is never profitable for the firm to deviate from the equilibrium, because $\frac{(V-p^*)(p^*-c)}{d_H}$ is the minimal profit the firm receives in equilibrium as well as the supremum of profit it could receive out-of-equilibrium. Similarly, $\mu(x_L, p) = d_L$ for all $p \neq p^*$ ensures that it is never profitable for the firm to deviate if $d = d_L$. Of course, these are extreme out-of-equilibrium beliefs, and the equilibrium can be supported by more moderate ones.

In equilibrium, the EG determines x_L from

$$x_L = \arg \min_x \frac{V - p^*}{\mu(x, p^*)} g_L + x \tag{10}$$

if $d = d_L$, and x_H from

$$x_H = \arg \min_x \frac{V - p^*}{\mu(x, p^*)} g_H + x \tag{11}$$

if $d = d_H$. The aim is to find all x_L and x_H , with $x_L \neq x_H$, that for some $\mu(x, p^*)$ satisfy (10) and (11).

Suppose $d = d_L$ and the EG deviates. What out-of-equilibrium beliefs are necessary to sustain the equilibrium amount of advertising x_L ? It must be that

Footnote 9 continued

$$\frac{(V - \hat{p})(\hat{p} - c)}{\mu(0, \hat{p})} < \frac{(V - p^*)(p^* - c)}{\mu(0, p^*)}$$

because p^* maximizes $(V - p)(p - c)$ and $\mu(0, \hat{p}) = \mu(0, p^*) = \bar{\mu}$. Therefore both types deviate to p^* and the only equilibrium that survives this refinement is the one in which $p_L = p_H = p^*$. See Mailath et al. (1993) for details.

$$\frac{V - p^*}{\mu(x, p^*)} g_L + x \geq \frac{V - p^*}{d_L} g_L + x_L \quad \text{for all } x \neq x_L. \tag{12}$$

For $x > x_L$, this inequality always holds if $\mu(x, p^*) = d_L$ (i.e. $\mu(x, p^*)$ remains low). This restriction implies that extra money spent on advertising is useless, because consumers do not increase their expected damage levels. For $x < x_L$, a necessary condition is $\mu(x, p^*) < d_L$, which is impossible. To make this deviation unattractive, consumers must believe that the product is cleaner than they know is possible. Therefore, if $x_L > 0$, the incentive compatibility constraint (12) cannot hold, and I conclude that $x_L = 0$. If the product is clean, the EG does not advertise.

Next, suppose that $d = d_H$. Here the out-of-equilibrium beliefs must satisfy:

$$\frac{V - p^*}{\mu(x, p^*)} g_H + x \geq \frac{V - p^*}{d_H} g_H + x_H \quad \text{for all } x \neq x_H. \tag{13}$$

If $x > x_H$, this inequality holds if $\mu(x, p^*) = d_H$. Again, the interpretation is that if the extra cost does not increase the expected damage level, then for the EG, such a deviation is not worthwhile. Let $x < x_H$ and rewrite (13) as:

$$\mu(x, p^*) \leq \frac{(V - p^*)g_H}{(V - p^*)g_H + (x_H - x)d_H} d_H. \tag{14}$$

Note that the RHS is smaller than d_H . However, without further restrictions, it might also be smaller than d_L , which should be impossible. The RHS is minimal at $x = 0$. If we take $0 < x_H \leq (V - p^*) \left(\frac{1}{d_L} - \frac{1}{d_H} \right) g_H \equiv R$, and we set $\mu(x, p^*) = d_L$ for $x < x_H$ (the worst belief for the EG), then the inequality in (14) is true. Thus, the out-of-equilibrium beliefs that support the equilibrium under consideration are: $\mu(x, p^*) = d_L$ if $x < x_H$ and $\mu(x, p^*) = d_H$ if $x > x_H$. The EG is punished severely if it advertises less than x_H , and the beliefs ensure that it is never optimal to advertise more than x_H .

In equilibrium, $x_L = 0$ and $0 < x_H \leq R$, but there also is an additional restriction: The EG must not be tempted to choose x_H if $d = d_L$ and x_L if $d = d_H$. For the EG, these incentive constraints imply:

$$\frac{V - p^*}{d_L} g_L \leq \frac{V - p^*}{d_H} g_L + x_H \quad \text{if } d = d_L \text{ and} \tag{15}$$

$$\frac{V - p^*}{d_H} g_H + x_H \leq \frac{V - p^*}{d_L} g_H \quad \text{if } d = d_H, \tag{16}$$

which implies that

$$x_H \in \left[(V - p^*) \left(\frac{1}{d_L} - \frac{1}{d_H} \right) g_L, (V - p^*) \left(\frac{1}{d_L} - \frac{1}{d_H} \right) g_H \right] \equiv [L, R]. \tag{17}$$

Observe that $R > L > 0$. In these OSE, the EG always advertises if $d = d_H$. The cost of advertising if the product is dirty must be sufficiently large; if the product is clean, the benefits of campaigning are too small.

In summary, in these OSE, the firm no longer needs to inform consumers and chooses its optimal price p^* . The EG now has to advertise with at least an intensity of L if damage is high. So there are OSE in which $x_L = 0, x_H \in [L, R]$, and $p_L = p_H = p^*$. These equilibria are supported by out-of-equilibrium beliefs in which, in the case of deviations by the firm, perceived damage is high enough, and, in case of deviations by the EG, perceived damage is low enough. The OSE in which $x_L = 0, x_H = L$, and $p_L = p_H = p^*$ Pareto-dominates any other equilibrium of this type. Therefore I will use this equilibrium with advertising as a comparison with the benchmark. Formally,

Proposition 2 (Equilibrium with advertising) *There is a perfect Bayesian equilibrium in which the EG does not advertise if damage is low and advertises with intensity $L > 0$ if damage is high. The firm sets a price p^* in both states of the world.*

4 Comparison

In this section, I will compare the OSE with advertising with the benchmark of no advertising. In particular, I will compare the *ex ante* cost of the EG and profit of the firm. For the equilibrium with advertising, the *ex ante* expected total cost and profit are:

$$\Delta_{ads} = \rho \frac{V - p^*}{d_L} g_L + (1 - \rho) \frac{V - p^*}{d_H} g_H + (1 - \rho)L, \tag{18}$$

$$\Pi_{ads} = \rho \frac{(V - p^*)(p^* - c)}{d_L} + (1 - \rho) \frac{(V - p^*)(p^* - c)}{d_H}. \tag{19}$$

In the benchmark equilibrium, we have:

$$\Delta_{bench} = \frac{V - p^*}{\rho d_L + (1 - \rho)d_H} (\rho g_L + (1 - \rho)g_H) \tag{20}$$

and

$$\Pi_{bench} = \frac{(V - p^*)(p^* - c)}{\rho d_L + (1 - \rho)d_H}. \tag{21}$$

Now we can state the following Proposition:

Proposition 3 *Compared with the benchmark, the firm’s profit and consumers’ surplus increase when the EG is allowed to advertise. However, the cost of the EG only decreases if $\frac{g_L}{g_H} < \frac{d_L}{d_L + d_H}$ and ρ is sufficiently close to 1.*

Proof The firm’s *ex ante* profit is always higher in the equilibrium with advertising than in the benchmark equilibrium. Because profit is convex in μ , applying Jensen’s inequality to (19) and (21) yields the desired result. Regarding consumers’ surplus, in the case of a linear demand function and a monopolistic firm, consumers’ surplus is exactly half the monopoly profit. *Ex ante*, consumers also prefer the advertising equilibrium.

Observe that $\Delta_{ads} \leq \Delta_{bench}$ can be rewritten as:

$$\frac{g_L}{d_L} + (1 - \rho) \frac{g_H - g_L}{d_H} \leq \frac{\rho g_L + (1 - \rho)g_H}{\rho d_L + (1 - \rho)d_H}$$

Divide both sides by g_H and multiply by d_H to obtain:

$$\frac{\hat{g}}{\hat{d}} + (1 - \rho)(1 - \hat{g}) \leq \frac{\rho \hat{g} + (1 - \rho)}{\rho \hat{d} + (1 - \rho)},$$

where $\hat{g} = g_L/g_H$ and $\hat{d} = d_L/d_H$. Using the assumption that $0 \leq \hat{g} \leq \hat{d}$, solving for ρ yields:

$$\rho \geq \frac{\hat{g}}{\hat{d}(1 - \hat{g})}.$$

Because $\rho \in [0, 1]$, the RHS of this inequality must be smaller than 1. It is easy to check that this is true if and only if $\frac{g_L}{g_H} < \frac{d_L}{d_L+d_H}$. □

The equilibrium with advertising could unambiguously raise welfare, but there must be a discrepancy between the EG’s assessment of the environmental damage and the consumers’. If this assessment is the same ($g_H/g_L = d_H/d_L$), it is never beneficial for the EG to engage in advertising. This claim is not surprising, because in this case, on average and net of advertising cost, the EG is indifferent between having informed or uninformed consumers. As the discrepancy increases, the cost to the EG of misinformed consumers increases as well. At some point, that cost offsets the cost of advertising.

5 Cost differences

For the reasons outlined previously, I began with the case in which both goods have equal marginal costs. However, a common assumption in the literature is that the marginal cost of the low damage good exceeds the marginal cost of the high damage good. In this section, I therefore explore the consequences of introducing cost differences. This exercise provides a robustness check; to avoid notational cluttering, I focus solely on the case in which $g_H = d_H$ and $g_L = d_L$. Recall that in this case, the EG prefers the benchmark over the equilibrium with advertising.

Let c_i be the marginal cost of producing a good with damage level $i = L, H$. Assume that $c_L > c_H$. With full information, the prices and profit levels, for the low- and high damage producer respectively are

$$p_L^* = \frac{1}{2}V + \frac{1}{2}c_L \quad \text{and} \quad \pi_L^* = \frac{(V - c_L)^2}{4d_L}, \tag{22}$$

$$p_H^* = \frac{1}{2}V + \frac{1}{2}c_H \quad \text{and} \quad \pi_H^* = \frac{(V - c_H)^2}{4d_H}. \tag{23}$$

In the absence of an EG, this version of the model is similar to the model of Mahenc (2008). To obtain comparable results, I impose the following assumption:

Assumption 1 With full information, the firm prefers to sell the low damage good: $\pi_L^* > \pi_H^*$.

I heuristically derive three equilibria for this game. The detailed derivations are analogous to those in the previous sections and are available on request. The three equilibria are

1. an OSE in which the firm signals,
2. a TSE in which the firm’s strategy is to use the full information prices,
3. the benchmark pooling equilibrium.

The choice of these equilibria is motivated by several considerations. First, I want to determine if the welfare conclusions of Sect. 4 change qualitatively when cost differences are introduced. A potential problem arises, because the equilibrium with advertising was an OSE in which the firm used undistorted prices. These prices were the same for low and high damage firms, whereas in this extension, the prices differ, and the equilibrium with advertising necessarily becomes a TSE. Therefore, the derivation is a bit more involved. Second, because there are OSE in which the firm signals (cf. Mahenc 2008), I can now determine who is more efficient at signaling environmental quality; the EG or the firm.

5.1 Derivation of equilibria

5.1.1 OSE in which the firm signals

Because this is a separating equilibrium, there is no need for the EG to spend money on advertising, and $x_L = x_H = 0$. The firm’s strategy is as follows: The bad type, which is the high damage firm, plays an undistorted strategy with $p_H = p_H^*$.¹⁰ The good type, which is the low damage firm, sets a price p_L that makes mimicking unprofitable for the high damage firm:

$$\pi_H^* \geq \frac{(V - p_L)(p_L - c_H)}{d_L}. \tag{24}$$

It is immediately evident that $p_L \neq p_L^*$, because substituting p_L^* for p_L yields, after some manipulation,

$$\frac{-(c_L - c_H)(V - c_L)}{2d_L} \geq \pi_L^* - \pi_H^*, \tag{25}$$

where the LHS is negative and the RHS is positive by assumption. Hence we have a contradiction and $p_L \neq p_L^*$. I take p_L to be the smallest price, larger than p_L^* , that satisfies (24). This equilibrium is the one that any reasonable refinement would

¹⁰ This is the well-known “no distortion at the bottom” result.

select, namely, the least distorting signal with the property that the product with higher marginal cost has a higher price. In summary,

Proposition 4 (OSE in which the firm signals) *There is an OSE in which the firm signals, such that $p_H = p_H^*$, and*

$$p_L = \frac{(V + c_H) + (V - c_H)\sqrt{1 - \frac{d_L}{d_H}}}{2} \in (p_L^*, V).$$

The EG does not advertise.

Proof The equilibrium has already been derived in the text. Appendix A contains the derivation of p_L . □

5.1.2 TSE in which the firm’s strategy is to use full information prices

I explicitly look for TSE in which $p_L = p_L^*$ and $p_H = p_H^*$, so that I can eventually compare two extreme scenarios: one in which the EG bears all the cost of signaling and one in which the firm bears all the cost (as derived in Sect. 5.1.1). Intuitively, the EG bears all the cost if the firm uses undistorted prices. From the discussion in Sect. 3, it is also obvious that $x_L = 0$ and $x_H = \hat{x} > 0$. The level of advertising, when damage is high, is the only undetermined quantity.

The restrictions on \hat{x} derive from $\mu_{LH} := \mu(x_L, p_H)$ and $\mu_{HL} := \mu(x_H, p_L)$. Note that these are the beliefs at $(x_L, p_H) = (0, p_H^*)$ and $(x_H, p_L) = (\hat{x}, p_L^*)$, respectively, where the consumer does not know whether the firm or the EG has deviated from the equilibrium strategy. Therefore, the beliefs at these points must simultaneously deter the EG and the firm from deviating.

With regard to μ_{LH} , for the EG, the actual state of the world is high damage, and the deviation is to zero advertising. The EG does not want to deviate if

$$\frac{V - p_H^*}{d_H} d_H + \hat{x} \leq \frac{V - p_H^*}{\mu_{LH}} d_H \tag{26}$$

For the firm, the actual state of the world is low damage, and the deviation is to p_H^* . This situation is never profitable, because the equilibrium strategy gives the firm the highest possible profit. This restriction is as loose as possible if $\mu_{LH} = d_L$.

With regard to μ_{HL} , for the EG, the actual state of the world is low damage and the deviation is to an advertising level of \hat{x} . Again, this should not be profitable:

$$\frac{V - p_L^*}{d_L} d_L \leq \hat{x} + \frac{V - p_L^*}{\mu_{HL}} d_L. \tag{27}$$

For the firm the actual state of the world is high damage and the deviation is to p_L^* . This deviation could be profitable if it were accompanied by the belief that damage is sufficiently low. Setting $\mu_{HL} = d_H$ solves this problem. Note that (27) is satisfied as well for this belief. Therefore, this belief simultaneously deters the EG and the firm from deviating.

From (26), it is possible to derive an upper bound on \hat{x} :

$$\hat{x} \leq (V - p_H^*) \left(\frac{d_H}{d_L} - 1 \right). \tag{28}$$

From (27) we obtain a lower bound on \hat{x} :

$$\hat{x} \geq (V - p_L^*) \left(1 - \frac{d_L}{d_H} \right). \tag{29}$$

The upper bound should be larger than the lower bound. This implies

$$(V - p_H^*) \left(\frac{d_H}{d_L} - 1 \right) \geq (V - p_L^*) \left(1 - \frac{d_L}{d_H} \right). \tag{30}$$

Because $(V - p_H^*) > (V - p_L^*)$, it is sufficient to show that $\left(\frac{d_H}{d_L} - 1 \right) > \left(1 - \frac{d_L}{d_H} \right)$, which is clearly true. The TSE with the smallest advertising level is the one in which \hat{x} equals the RHS of (29).

Proposition 5 (TSE with advertising) *There is a perfect Bayesian equilibrium in which the EG does not advertise when the state of the world is low and advertises with intensity $(V - p_L^*) \left(1 - \frac{d_L}{d_H} \right)$ when the state of the world is high. The firm uses undistorted prices: $p_L = p_L^*$ and $p_H = p_H^*$.*

5.1.3 The benchmark pooling equilibrium

In the pooling equilibrium, $x_L = x_H = 0$ and $p_L = p_H = \hat{p}$. The price the firm charges in the pooling equilibrium is such that for the best deviation (p_i^*) and the most punishing belief ($\mu = d_H$), the firm prefers \hat{p} . In particular, I examine the pooling equilibrium in which the high damage firm mimics the low damage firm, that is, $\hat{p} = p_L^*$.

In this case, the high damage firm can choose between p_L^* and the belief $\mu = \bar{\mu}$ or p_H^* and $\mu = d_H$. The high damage firm should prefer to play p_L^* :

$$\frac{(V - p_L^*)(p_L^* - c_H)}{\bar{\mu}} \geq \frac{(V - c_H)^2}{4d_H}, \tag{31}$$

which leads to the following Proposition.

Proposition 6 *If $(1 - \beta)c_H + \beta V \geq c_L$, where $\beta = \sqrt{1 - \frac{\bar{\mu}}{d_H}}$, then there is a pooling equilibrium in which $p_L = p_H = p_L^*$ and $x_L = x_H = 0$.*

Proof The equilibrium has been derived above. Appendix A contains the derivation of the condition under which this pooling equilibrium exists. □

5.2 Welfare comparison

5.2.1 Comparing the TSE with the pooling equilibrium

It is possible now to answer the first question, regarding whether cost differences change the basic welfare conclusions of Sect. 4 for the case in which $g_L = d_L$ and $g_H = d_H$. Summarizing Proposition 3 succinctly: the EG prefers the benchmark pooling equilibrium, whereas the firm prefers the separating equilibrium in which the EG advertises. In the case with cost differences, the *ex ante* cost for the EG in the pooling equilibrium is given by:

$$\Delta_P = \rho \frac{V - c_L}{2\bar{\mu}} d_L + (1 - \rho) \frac{V - c_L}{2\bar{\mu}} d_H = \frac{1}{2}(V - c_L), \tag{32}$$

and the *ex ante* cost in the TSE is given by:

$$\begin{aligned} \Delta_T &= \rho \frac{V - c_L}{2d_L} d_L + (1 - \rho) \left[\frac{V - c_H}{2d_H} d_H + x_H \right] \\ &= \frac{1}{2} \left(V - c_L + (1 - \rho)(V - c_H) - (1 - \rho)(V - c_L) \frac{d_L}{d_H} \right). \end{aligned} \tag{33}$$

Note that $\Delta_P \leq \Delta_T$ if

$$0 \leq (1 - \rho)(V - c_H) - (1 - \rho)(V - c_L) \frac{d_L}{d_H} \iff (V - c_L)d_L \leq (V - c_H)d_H. \tag{34}$$

Under the assumption that $c_L > c_H$ and $d_L < d_H$, this inequality holds, and the EG prefers the pooling equilibrium over the TSE. That is, the ability to advertise hurts the EG.

For the firm, the *ex ante* profit in the pooling equilibrium is:

$$\Pi_P = \rho \frac{(V - c_L)^2}{4\bar{\mu}} + (1 - \rho) \frac{(V - c_L)^2 + 2(c_L - c_H)(V - c_L)}{4\bar{\mu}}, \tag{35}$$

and the *ex ante* profit in the TSE is:

$$\Pi_T = \rho \frac{(V - c_L)^2}{4d_L} + (1 - \rho) \frac{(V - c_H)^2}{4d_H}. \tag{36}$$

As the derivation in Appendix B shows $\Pi_T \geq \Pi_P$. The welfare conclusions of Proposition 3 are robust to the introduction of cost differences.

5.2.2 Comparing TSE (EG advertises) with OSE (firm advertises)

It is obvious that the firm prefers the TSE over the OSE. Both equilibria are separating, but in the OSE, the profit of the firm diminishes because of costly signaling. For opposite reasons, the EG prefers the OSE over the TSE. In which equilibrium is welfare higher though? That is, who signals most efficiently?

Welfare if the firm signals is:¹¹

$$\rho \left(\frac{(V - p_L)(p_L - c_L)}{d_L} + \frac{1}{2} \frac{(V - p_L)^2}{d_L} \right) + (1 - \rho) \frac{3(V - c_H)^2}{8d_H} - \frac{1}{2} \rho(V - c_L) - \frac{1}{2} (1 - \rho)(V - c_H), \tag{37}$$

where p_L is as defined in Proposition 4. In contrast, welfare if the EG signals is:

$$\frac{3}{2} \left(\rho \frac{(V - c_L)^2}{4d_L} + (1 - \rho) \frac{(V - c_H)^2}{4d_H} \right) - \frac{1}{2} \rho(V - c_L) - \frac{1}{2} (1 - \rho)(V - c_H) - \frac{1}{2} (1 - \rho)(V - c_L) \left(1 - \frac{d_L}{d_H} \right). \tag{38}$$

Through subtraction, we obtain:

$$\frac{\rho}{d_L} \left((V - p_L)(p_L - c_L) + \frac{1}{2} (V - p_L)^2 - \frac{3(V - c_L)^2}{8} \right) + \frac{1}{2} (1 - \rho)(V - c_L) \left(1 - \frac{d_L}{d_H} \right) \tag{39}$$

which is positive if the firm is more efficient at signaling. Note that the first term is negative,¹² and the second term is positive.

Proposition 7 *The EG signals more efficiently than the firm if $\rho > \bar{\rho} \in (0, 1)$.*

Proof Note that p_L is independent of ρ in the OSE and (39) is linear in ρ , positive if $\rho = 0$ and negative if $\rho = 1$. This establishes the existence of $\bar{\rho}$. \square

Therefore, the EG is more efficient only when it is likely that the product is clean, i.e., when the probability of advertising is low. One interpretation is that the EG is a good channel when the problems due to the lack of information are minimal.

¹¹ From the demand function (cf. Eq. 4) we see that consumers' surplus equals $(V - p)^2/d$. If the firm charges undistorted prices, consumers' surplus is half of profit, and welfare is $\frac{3}{2}$ times profit minus the cost of the EG. This is the case when $d = d_H$.

¹² Note that $(V - p_L)(p_L - c_L) + \frac{1}{2} (V - p_L)^2$ equals $\frac{3(V - c_L)^2}{8}$ at $p_L = p_L^*$. However $p_L > p_L^*$, and therefore, the first term is negative.

6 Endogenizing the choice of damage and the acquisition of information

Thus far, this study has offered little reason for an EG to provide information to the public, whether from a social welfare viewpoint or from the viewpoint of the EG. In this section, I investigate when the EG collects information. Therefore, the game includes two additional prior stages. First, the firm decides to produce a low damage good or a high damage good. I return to the assumption that there is no difference in marginal costs between the two, but there is a (small) fixed cost K of starting to produce the low damage good. Second, the EG chooses whether to inform itself about the production process of the firm (at cost F). I again allow for the possibility that the EG's preferences for a low damage good are stronger than those of the consumers: $g_H/g_L \geq d_H/d_L$. After these choices, the game is identical to the one analyzed in Sect. 3, but now the probability that $d = d_L$ must be consistent with the firm's strategy in the first period. Moreover, it is common knowledge whether the EG is informed or not.¹³ An informed EG leads to an equilibrium with advertising, whereas an uninformed EG yields the benchmark of no advertising.

There is always an equilibrium in which the firm decides to produce a high damage good and the EG does not get informed, because in the absence of an informed EG, the classic market for lemons emerges, and the consumer is unable to distinguish between low and high damage goods. Because of the fixed cost of producing the low damage good, the firm's profit is highest if it chooses $d = d_H$. Because there is no uncertainty about the firm's damage level, there are no gains for the EG from collecting and disseminating information among consumers. Its best response is not to acquire information.

It is more interesting to investigate whether there are equilibria in which the EG does acquire information. From Proposition 3, it is obvious that two cases must be distinguished, namely, $\frac{g_L}{g_H} \geq \frac{d_L}{d_L+d_H}$ and $\frac{g_L}{g_H} < \frac{d_L}{d_L+d_H}$. In the first case, the cost of the EG is never lower when it is informed, so there is no incentive to acquire information.

Suppose $\frac{g_L}{g_H} < \frac{d_L}{d_L+d_H}$ and the EG believes that the firm produces a low damage good with probability ρ . If the EG chooses to become informed, its cost will be $\Delta_{ads} + F$; otherwise, its cost is Δ_{bench} . According to Proposition 3, there exists a $\hat{\rho} \in (0, 1)$ such that $\Delta_{ads} = \Delta_{bench}$. Moreover at $\rho = 1$, the EG is indifferent, because $\bar{\mu} = d_L$ and there is no advertising. For $\rho \in (\hat{\rho}, 1)$, the cost of the EG strictly decreases if it is informed; therefore, the EG gathers information if the cost of acquiring that information is small. In equilibrium, the belief of the EG must be consistent with the strategy of the firm, so the firm chooses low damage with probability ρ and high damage with probability $1 - \rho$. To ensure that it is an equilibrium strategy, the firm must remain indifferent between its two options, which can be achieved by letting the EG acquire information with a certain probability. By the logic of mixed strategy equilibria, the EG also must be indifferent between its two options, which means that the cost of information acquisition equals the benefit of committing to advertise (when damage is high). Because advertising is most beneficial if uncertainty about the quality of the product is greatest,

¹³ This assumption is common (e.g. Feddersen and Gilligan 2001), and helps avoid many technical issues. If consumers can observe that the EG starts an investigation into the practices of the firm and the investigation always leads the EG to conclude that a firm has either low or high damage, the assumption of common knowledge is justified.

but least beneficial if the product is almost certainly low or high damage, there are two such equilibria; one with a low probability of low damage and one with a high probability. Formally,

Proposition 8 *For $F > 0$ and $K > 0$ sufficiently small, in addition to the equilibrium in which the firm always sets high damage and the EG does not acquire information, there exist two mixed equilibria in which*

1. *the EG acquires information with probability $\sigma^*(K) = \frac{4K}{(v-c)^2(d_L^{-1}-d_H^{-1})}$,*
2. *the firm sets low damage with probability $\rho^*(F) \in \{\underline{\rho}(F), \bar{\rho}(F)\}$, where $\underline{\rho}(F)$ and $\bar{\rho}(F)$ are the two solutions to $\Delta_{ads} + F = \Delta_{bench}$,*
3. *$\rho(0) = \hat{\rho}$ and ρ is increasing in F ,*
4. *$\bar{\rho}(0) = 1$ and $\bar{\rho}$ is decreasing in F .*

The Pareto-superior equilibrium is the one in which $\rho^(F) = \bar{\rho}(F)$. As $F \rightarrow 0$, the Pareto-superior equilibrium approaches the efficient outcome, in the sense that the firm always chooses low damage and there is no advertising.*

Proof See Appendix C. □

This proposition features several caveats. Most important, the equilibria only exist when the cost of information acquisition and the fixed cost of producing the low damage good are sufficiently small, and the discrepancy between the EG’s assessment of the environmental damage and the consumers is sufficiently high. Nevertheless, this section shows that the threat of advertising might induce the firm to choose low damage with high probability.

7 Conclusion

This article has investigated informative advertising by an EG when a monopolist sells a good that causes a certain amount of damage per unit sold to the environment. Consumers care about this damage but are unable to discern its level. Therefore, the EG uses an advertising campaign to disseminate information about product quality.

When the firm’s level of damage is given and the EG possesses information, advertising by the EG unambiguously increases the firm’s profit and consumer surplus. However, the EG itself is unlikely to benefit. This finding naturally leads to the question of why the EG would acquire information in the first place. If its assessment of environmental damage is stronger than the public’s and the cost of acquiring information is sufficiently low, there exists an equilibrium where the EG acquires information (with some probability) and the threat of advertising causes the firm to produce a low damage good (with some probability).

However, if the cost of acquiring information is too high, then the only equilibrium is one in which the firm always produces a high damage good. One way to offset the cost of acquiring information is to lower the cost of campaigning. This cost stems from the EG’s lack of credibility, so a possible solution might be reputation building. Further research should investigate the ways in which the EG can build and maintain a trustworthy reputation.

Acknowledgments The author thanks Bert Schoonbeek, José-Luis Moraga-González, John Maxwell, Peter Kooreman, an anonymous referee, and participants of the 3rd World Congress of Environmental and Resource Economists in Kyoto 2006, the 6th IDEI-LERNA Conference on Environmental and Resource Economics in Toulouse 2007, and seminars at the University of Groningen and the University of Amsterdam, for their helpful comments and discussion. Financial support by the Netherlands Organization for Scientific Research (NWO) is gratefully acknowledged.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

Appendix A: Details Sect. 5

Preliminary remarks The objects of interest are the intervals determined by inequalities of the following form:

$$\frac{(V-p)(p-c)}{\mu} \geq \frac{(V-c)^2}{4d_H}, \quad (40)$$

where c is either c_L or c_H , and $\mu \in (d_L, d_H)$. The lower and upper bound are solutions of:

$$(V-p)(p-c) - \alpha \frac{(V-c)^2}{4} = 0, \quad (41)$$

where $\alpha = \frac{\mu}{d_H} \in (0, 1)$. This quadratic equation in p has two real positive solutions, namely,

$$p = \frac{(V+c) \pm (V-c)\sqrt{1-\alpha}}{2}. \quad (42)$$

OSE (firm signals) The price that the low damage firm charges is the upper bound of the interval determined by

$$\frac{(V-p)(p-c_H)}{d_L} \geq \frac{(V-c_H)^2}{4d_H}, \quad (43)$$

which is equal to:

$$p_L = \frac{(V+c_H) + (V-c_H)\sqrt{1-\frac{d_L}{d_H}}}{2}. \quad (44)$$

Pooling equilibrium As noted in the main text, the price the firm charges in the pooling equilibrium is such that for the best deviation (p_i^*) and the most punishing

belief ($\mu = d_H$), both types prefer to play \hat{p} . For the low damage firm, this implies

$$\frac{(V - \hat{p})(\hat{p} - c_L)}{\bar{\mu}} \geq \frac{(V - c_L)^2}{4d_H}, \tag{45}$$

and for the high damage firm,

$$\frac{(V - \hat{p})(\hat{p} - c_H)}{\bar{\mu}} \geq \frac{(V - c_H)^2}{4d_H}. \tag{46}$$

The main text focuses on the particular pooling equilibrium in which $\hat{p} = p_L^*$. Because $p_L^* = \frac{V+c_L}{2}$, it is always in the interval determined by (45) and not necessarily in the interval determined by (46). Because $p_L^* > \frac{V+c_H}{2}$, it is only necessary to check if the upper bound implied by (46) is larger than p_L^* :

$$\frac{(V + c_H) + (V - c_H)\sqrt{1 - \alpha}}{2} \geq \frac{V + c_L}{2}, \tag{47}$$

which implies

$$(1 - \sqrt{1 - \alpha})c_H + (\sqrt{1 - \alpha})V \geq c_L. \tag{48}$$

This is a weighted average of c_H and V , and a pooling equilibrium in which $\hat{p} = p_L^*$ exists when the valuation for the product is high compared with the marginal cost.¹⁴

Appendix B: Details $\Pi_T \geq \Pi_P$

It is possible to rewrite $\Pi_T - \Pi_P$ as

$$\frac{\rho(1 - \rho)(d_H - d_L)(\pi_L^* - \pi_H^*)}{\bar{\mu}} + \frac{(1 - \rho)(c_L - c_H)^2}{4\bar{\mu}},$$

which is clearly positive under the assumption that $\pi_L^* > \pi_H^*$ and $\rho < 1$.¹⁵ It has to be shown that it is indeed possible to rewrite $\Pi_T - \Pi_P$ in this form, which boils down to showing that

$$\begin{aligned} & \frac{\rho(1 - \rho)(d_H - d_L)(\pi_L^* - \pi_H^*)}{\bar{\mu}} + \frac{(1 - \rho)(c_L - c_H)^2}{4\bar{\mu}} \\ &= \rho\pi_L^* + (1 - \rho)\pi_H^* - \frac{(V - c_L)^2}{4\bar{\mu}} - \frac{(1 - \rho)(V - c_L)(c_L - c_H)}{2\bar{\mu}}. \end{aligned}$$

¹⁴ This proof generalizes to pooling equilibria, where $\hat{p} \neq p_L^*$.

¹⁵ I thank an anonymous referee for this suggestion.

Multiplying by $4\bar{\mu}$ and rearranging yields:

$$\begin{aligned} (V - c_L)^2 + (1 - \rho)(c_L - c_H)(2V - c_L - c_H) \\ = 4\rho[\bar{\mu} - (1 - \rho)(d_H - d_L)]\pi_L^* + 4(1 - \rho)[\bar{\mu} + \rho(d_H - d_L)]\pi_H^*. \end{aligned}$$

Using the definitions of $\bar{\mu}$, π_L^* , and π_H^* , this equation simplifies to

$$(V - c_L)^2 + (1 - \rho)(c_L - c_H)(2V - c_L - c_H) = \rho(V - c_L)^2 + (1 - \rho)(V - c_H)^2.$$

Collecting similar terms and dividing by $(1 - \rho)$, we obtain:

$$(V - c_L)^2 - (V - c_H)^2 + (c_L - c_H)(2V - c_L - c_H) = 0.$$

Finally, using the identity $(V - c_L)^2 - (V - c_H)^2 = (V - c_L + V - c_H)(V - c_L - V + c_H)$ shows that this equality always holds.

Appendix C: Proof of Proposition 8

Let ρ be the probability that the firm chooses low damage and σ be the probability that the EG chooses to acquire information. As the discussion in the main text shows, ρ should be such that the EG is indifferent between acquiring information and not acquiring information, and σ should be such that the firm is indifferent between low and high damage.

The expected profit of choosing $d = d_L$ is:

$$\sigma \frac{(V - c)^2}{4d_L} + (1 - \sigma) \frac{(V - c)^2}{4\bar{\mu}} - K. \tag{49}$$

Similarly, for $d = d_H$:

$$\sigma \frac{(V - c)^2}{4d_H} + (1 - \sigma) \frac{(V - c)^2}{4\bar{\mu}}. \tag{50}$$

To make the firm indifferent, equate the two, which yields:

$$\sigma^*(K) = \frac{4K}{(V - c)^2(d_L^{-1} - d_H^{-1})}, \tag{51}$$

which falls between 0 and 1 for a sufficiently small K .

To find ρ , note that the EG is indifferent between acquiring information and not acquiring information if

$$\Delta_{ads} + F = \Delta_{bench}. \tag{52}$$

Inspecting (18) and (20), we see that this equation can be rewritten as a second-degree polynomial in ρ . For $F = 0$, the two solutions are $\rho = \hat{\rho}$ and $\rho = 1$. When F is sufficiently high, there is no solution. Then there exists $\bar{F} > 0$ such that for $F \in [0, \bar{F}]$, the EG is indifferent if $\rho = \underline{\rho}(F)$ and $\rho = \bar{\rho}(F)$, where $\hat{\rho} \leq \underline{\rho}(F) \leq \bar{\rho}(F) \leq 1$. It can be easily established that $\underline{\rho}(0) = \hat{\rho}$, $\underline{\rho}(F)$ is increasing in F , $\bar{\rho}(0) = 1$, and $\bar{\rho}(F)$ is decreasing in F .

To establish that the equilibrium in which $\rho = \bar{\rho}(F)$ is the Pareto-superior outcome, it suffices to show that the EG's expected cost is decreasing in ρ and the firm's expected profit is increasing in ρ . Note that the EG's expected cost is equal to $\sigma(\Delta_{ads} + F) + (1 - \sigma)\Delta_{bench}$, which equals Δ_{bench} in equilibrium, because (52) holds. It also is straightforward to verify that under the assumption that $g_H/g_L > d_H/d_L$, Δ_{bench} is a decreasing function of ρ . The firm's expected profit is

$$\sigma \Pi_{ads} + (1 - \sigma)\Pi_{bench} - \rho K = \sigma \frac{(V - c)^2}{4d_H} + (1 - \sigma)\Pi_{bench}$$

where the equality follows from combining (49) and (50). Only Π_{bench} depends on ρ . It is trivial to verify that Π_{bench} is increasing in ρ . Therefore, the firm's expected profit is increasing in ρ .

As $F \rightarrow 0$, $\bar{\rho}(F) \rightarrow 1$ and the firm chooses low damage almost with certainty. Because the EG only advertises when it is informed and damage is low, this happens with probability $(1 - \rho)\sigma$, and the probability of advertising goes to 0, which is the same as in the efficient outcome.

References

- Amacher G, Koskela E, Ollikainen M (2004) Environmental quality competition and eco-labeling. *J Environ Econ Manag* 47:284–306
- Bagwell K, Ramey G (1988) Advertising and limit pricing. *RAND J Econ* 19:59–71
- Bagwell K, Ramey G (1991) Oligopoly limit pricing. *RAND J Econ* 22:155–172
- Baron D (2003) Private politics. *J Econ Manag Strategy* 12:31–66
- Baron D (2005) Competing for the public through the news media. *J Econ Manag Strategy* 14:339–376
- Crawford V, Sobel J (1982) Strategic information transmission. *Econometrica* 50:1431–1451
- Cremer H, Thisse JF (1999) On the taxation of polluted products in a differentiated industry. *Eur Econ Rev* 43:575–594
- Darby M, Karni E (1973) Free competition and the optimal amount of fraud. *J Law Econ* 16:67–88
- Diermeier D, Baron D (2007) Strategic activism and non-market strategy. *J Econ Manag Strategy* 16:599–634
- Emons W (1997) Credence goods and fraudulent experts. *RAND J Econ* 28:107–119
- Emons W (2001) Credence goods monopolists. *Int J Ind Organ* 19:375–389
- Feddersen T, Gilligan T (2001) Saints and markets; activists and the supply of credence goods. *J Econ Manag Strategy* 10:149–171
- Heijnen P, Schoonbeek L (2008) Environmental groups in monopolistic markets. *Environ Resour Econ* 39:379–396
- Hertzenndorf M, Overgaard P (2001) Price competition and advertising signals: signaling by competing senders. *J Econ Manag Strategy* 10:621–662
- Ibanez L, Grolleau G (2008) Can ecolabeling schemes preserve the environment. *Environ Resour Econ* 40:233–249
- Innes R (2006) A theory of consumer boycotts under symmetric information and imperfect competition. *Econ J* 116:355–381

- Kihlstrom R, Riordan M (1984) Advertising as a signal. *J Polit Econ* 92:427–450
- Kuhn M (2005) The greening of markets: product competition. Pollution and policy making in a duopoly. Edward Elgar, Cheltenham
- Mahenc P (2007) Are green products over-priced? *Environ Resour Econ* 38:461–473
- Mahenc P (2008) Signaling the environmental performance of polluting products to green consumers. *Int J Ind Organ* 26:59–68
- Mailath G, Okuno-Fujiwara M, Postlewaite A (1993) Belief based refinements in signaling games. *J Econ Theory* 60:241–276
- Maxwell J, Lyon T, Hackett S (2000) Self-regulation and social welfare: the political economy of corporate environmentalism. *J Law Econ* 43:583–617
- Milgrom P, Roberts J (1986) Price and advertising signals of product quality. *J Polit Econ* 94:796–821
- Moraga-González JL, Padrón-Fumero N (2002) Environmental policy in a green market. *Environ Resour Econ* 22:419–447
- Petrakis E, Sartzetakis E, Xepapadeas A (2005) Environmental information provision as a public policy instrument. *Contrib Econ Anal Policy* 4:1–31
- Porter M, van den Linde C (1995) Toward a new conception of the environment-competitiveness relationship. *J Econ Perspect* 9:97–118
- Riley J (2001) Silver signals: twenty-five years of screening and signaling. *J Econ Lit* 39:432–478