Honest versus Misleading Certification

By Philippe Mahenc*

March 9, 2015

Abstract

This paper questions the honesty of third-party certification in the market for a good whose environmental quality is not observable by consumers. The certifier maximizes a weighted sum of its own revenue and social welfare. The higher the relative weight placed on revenue, the stronger the certifier's incentive to mislead consumers. Certification is analyzed as a costly signaling mechanism that, besides displaying labels, transmits information through market prices. Honest certification requires that prices credibly signal environmental quality to prevent cheating. I show that certification can only be honest when the certifier is driven more by social welfare than by profit. In the reverse case, the certifier cannot help jamming the price signal, thereby granting unreliable labels.

*LAMETA, Montpellier I University, Avenue Raymond Dugrand, Site de Richter, C. S. 79606, 34960 Montpellier Cedex 2 - France

Keywords: Bayesian inference, Certification, Credence good, Signaling.

JEL Code: D82, Q28.

Honest versus Misleading Certification

1 Introduction

Like any good with quality characteristics that are hard for consumers to observe, organic food needs third-party certification to be efficiently traded. But also, and above all, this certification must be honest. Recently in the press, industry lobbying has been accused of capturing the organic certification process and distracting the certifier from the objective of disclosing accurate information[1]. If the certifier really gives in to corporation pressure, this may have two consequences: first, the certifier tends to neglect consumer well-being and focus instead on the desire of the industry to grow, and second, the certifier may falsely claim that the food is organic to let more products in.

[1] This corporate pressure is acknowledged in recent newspaper articles. One can read for instance that "major food companies ... has added to pressure on the (US) government to expand the definition of what is organic, in part because processed foods offered by big industry often require ingredients, additives or processing agents that either do not exist in organic form or are not available in large enough quantities for mass production. (Kimberly Kindy and Lyndsey Layton, "Purity of Federal 'Organic' Label Is Questioned", *Washington Post*, July 3, 2009). Or else, "over the last decade, since federal organic standards have come to the fore, giant agri-food corporations ... — Coca-Cola, Cargill, ConAgra, General Mills, Kraft and M&M Mars among them — have gobbled up most of the nation's organic food industry. Between the time the Agriculture Department came up with its proposed regulations for the organic industry in 1997 and the time those rules became law in 2002, myriad small, independent organic companies — businesses like Cascadian Farm — were snapped up by corporate titans ... Major corporations have come to dominate the board that sets the standards for organic foods ... Corporate interests are behind the increase in nonorganic materials deemed acceptable in "organic" food." (Stephanie Strom, "Has 'Organic' Been Oversized?", *New York Times*, July 7, 2012).

Beyond organic food, misleading certification is an issue that more broadly affects all the goods with a credence attribute, whether environmental or ethical. For instance, Hamilton and Zilberman (2006) point out the recent emergence of fraudulent labelling in markets for environmentally-friendly products. Furthermore, misleading practices are not specific to the industry: Feddersen and Gilligan (2001) show that a certifier biased toward environmental protection may also have incentives that depart from honesty. Nonetheless, certification bodies are usually under the oversight of a board that consists of representatives from different sectors with a significant divergence of interests, including the government, the industry, and environmental non-governmental organizations. A balanced composition of the oversight board can reasonably be expected to provide consumers with sufficient protection against misleading information.

In this paper, I examine whether honest third-party certification hinges on some mix between two objectives: on the one hand, the profit-maximization objective that represents the joint interest of the certifier and the industry, and on the other hand, the welfare-maximization objective that represents the general interest. Third-party certification is analyzed as a costly signaling mechanism that, besides displaying "informative" labels, transmits information on a credence attribute of the good (namely, environmental quality) through market prices. As defined here, honest certification requires that price signaling be credible in the sense that jamming the price signal sent by the market is too costly for the certifier. The main insight of the analysis is that third-party certification can only be honest if the certifier is driven more by social welfare than by profit. If, on the other hand, the certifier values its own interest more than social efficiency, then the certifier grants misleading labels and prevents the market prices from disclosing truthful information.

The context is the following. In a market where consumers do not know precisely to what extent the good they purchase pollutes the environment, a third-party certifier grants a label, either brown or green, which is meant to disclose information on the environmental quality of the good. Firms must pay a fee for certification and they can choose to have their product certified or not. Consumers also seek to infer information from the price signal sent by the market. Understanding that the green good is more costly to produce than its brown variant, consumers expect the market price to reflect at least the cost discrepancy. Hence, the certifier must take care that the market clearing price be consistent with the label granted to the firm. Specifically, the goods labelled differently as brown or green should not be sold at the same "pooling" price, for the sake of label reliability. The certifier can influence the market signal via setting a fee and this influence is all the stronger as the certifier values its own revenue more than social welfare. On the one hand, the weight attributed to the converging interests of the certifier and the industry can also be read as an index of market power in the certification sector. On the other hand, the weight on social welfare measures social efficiency. In this context, I develop a signaling model and investigate the existence of separating equilibria which ensure honest certification through different market prices for the brown and green goods. As there exists no separating equilibrium when the certifier's interest is skewed in favor of the industry, I further characterize pooling equilibria in the limit case where the certifier is purely self-interested.

On the demand side, consumers are heterogeneous in their dislike of pollution. The pleasure from consuming "goods" may be spoilt by the awareness that a public "bad" possibly arises from the act of consumption. As recognized by the environmental literature, the trouble experienced by environmentally friendly consumers from buying polluting goods mitigates their willingness-to-pay for these goods, thereby influencing their purchasing behavior. So, for example, the growing care for the environment dissuades consumers from purchasing food that allows a high percentage of chemicals in the preparation, processing and packaging of the product. Similarly, the burning of fossil fuels involved in the manufacturing of goods may cause significant concern for consumers, since it makes consumption decisions responsible for greenhouse gases emissions. Another worrisome externality is the exhaustion of the resources due to consumption of an endangered species of animal. To fulfill their desire for environmentally friendly consumption, consumers will wish to switch from conventional goods to substitutes that claim that they are "organic", "green" or "sustainable", whenever they exist.

This decision, however, is made difficult by any information asymmetry between consumers and producers about the environmental quality of the product. Two essential channels through which information is disclosed, namely, market price and certification, do not always work properly. Just like the uninformative prices in the market for lemons analyzed by Akerlof (1970), the prices of polluting goods may conceal information about their environmental performance (see Mahenc, 2007, 2008). Furthermore, the proliferation in recent years of untrustworthy claims that a product is "green"—the phenomenon is known as greenwashing —is bringing discredit on the use of environmental labels[2]. Few are the certifiers who possess the skill and knowledge needed to specify whether some goods are environmentally friendly or not, and the certifier's incentive to extract consumer surplus on the basis of privileged information does not necessarily guarantee full information disclosure, as shown by Lizzeri (1999).

In the present paper, consumers prefer the good to be gentle to the environment but they cannot ascertain whether the good is brown or green. As they are willing to pay more for the green than for the brown good, the market prices account for vertical differentiation between the two goods provided that the certifier honestly distinguishes brown from green firms. In turn, honest certification relies on credible price signaling. The difficulty in ensuring credible signaling is that the certifier has an incentive to label the brown good as green and prevent market prices from disclosing full in-

[2] Greenwashing encompasses all practices that range from vague claims to misleading advertising about the environmental performance of firms. Some evidence that greenwashing is becoming widespread in the U.S. can be found in the growing number of complaints about green ads received by the Advertising Standard Authority. According to Lyon and Maxwell (2011), greenwash aims at hiding negative information about a company's environmental or social performance, rather than misleading consumers.

formation about the actual color of the good. The certifier's aim is then to extract more surplus from uninformed consumers and, possibly, let more brown products in, which finally increases the certifier's revenue. Clearly, the larger the relative weight given to the certifier's revenue in the objective function, the stronger the certifier's incentive to cheat consumers. To counteract this incentive, signaling through its market price that the good is green must be costly. The signaling cost is purely strategic but it cannot be avoided to ensure that certification is honest. This requirement takes the form of incentive compatibility constraints in the signaling model. In separating equilibria, market prices credibly signal the environmental quality of the good, hence certification is honest. On the other hand, in pooling equilibria, the brown and the green goods sell at the same price, regardless of the label. As consumers infer no further information from the price signal, pooling equilibria preclude honest certification.

It turns out that there exist separating equilibria only if the certifier is driven more by social welfare than by profit. Moreover, the signaling cost decreases as the certifier attributes a larger relative weight to social welfare. The intuition underlying this result is that the cost of signaling the green good must be lower than that of signaling the brown good not to yield to the temptation of cheating. This only holds true if the certifier puts enough emphasis on social welfare because the certifier then internalizes the extra cost involved in producing the green good whatever the type of good signaled by market price. If, on the other hand, the certifier's objective is skewed in favour of the industry, then the certifier cannot resist the temptation to jam the price signal and grant brown firms the wrong label. The resulting market price misleads consumers into believing that the certified product is more gentle to the environment than it actually is, so that they are more willing to pay for it. In the limit case where the certifier is purely self-interested, I show that pooling occurs when consumers' prior beliefs of environmental quality are sufficiently optimistic.

This paper is organized as follows. Section 2 presents a detailed review of the related literature. Section 3 sets out the structure of the model and presents the equilibrium concept. Section 4 proposes a five-finger exercise in the limit case where the certifier is purely self-interested, in order to illustrate the failure of honest certification. Section 5 investigates the existence of separating and pooling equilibria in the general setting. Section 6 offers some conclusions.

2 Related literature

Crespi and Marette (2001), Kuhn (2005), Dranove and Jin (2011), and Fischer and Lyon (2014) provide comprehensive overviews of the literature on certification. This literature generally assumes that certification is honest, although not always perfect. A prolific line of research has built on the assumption of perfect certification, so that the mere presence of third-party certification solves the problem of asymmetric information between consumers and firms about some characteristics of the good, whether environmental or not. The central role of certification is then to create vertically differentiated market segments. From this perspective, most of the papers use standard vertical differentiation models under perfect information to analyze the strategic interplay between firms and certification. Representative of this line of research, Amacher et al. (2004) conclude that it is important to relax the assumption of perfect information. This indeed opens up a new research path that helps further our understanding of the failure of certification processes to disclose truthful information. However, research along this line often ignores the possibility of dishonest third-party certification. If certification happens to be misleading, these models place the blame on firms that bear "disguise costs" (Hamilton and Zilberman, 2006, Ibanez and Grolleau, 2008) to cheat consumers, without questioning the honesty of third-party certification. One exception is Feddersen and Gilligan (2001) in which an environmental activist is responsible for sending misleading messages to consumers. The idea of dishonest third-party certification seems much easier to accept in the economic literature on financial markets after the subprime crisis of 2007. For instance, Mathis et al. (2009) examine the possibility that credit rating agencies have been deliberately too lax in the ratings of some complex financial products. Besides these papers, the literature on imperfect certification also insists on the role of "efficacy costs" due to auditing and inspection procedures in collecting and disclosing truthful information. It is easy to understand that the gap in efficacy costs somehow explains the heterogeneity in certification accuracy, as observed for instance by Duflo et al. (2013) in the Gujarat market for third-party environmental auditors.

The present paper departs from most of the papers cited in the emphasis put on signaling costs rather than efficacy costs. I argue that honest third-party certification requires costly signaling, therefore misleading certification stems from the failure to bear signaling costs. From this perspective, the most closely related paper is Mathis et al. (2009). They show that a monopolistic credit rating agency always inflates the ratings of financial products when a too large fraction of the agency's revenue comes from the issuers of theses products rather than from the investors. This has the same flavor as the present finding that the certifier always cheats when it is driven more by profit than by social welfare.

Regarding the literature on environmental certification, Fischer and Lyon (2014) investigate a model of strategic differentiation between two eco-labels developed by an environmental nongovernmental organization and by an industry association. Assuming that certification provides consumers with perfectly reliable information about the firms' performance in green production, these authors discuss the impact of label rivalry on industry profits and the environment. In the present model, the possibility of the certifier's exercising market power is more roughly captured by the relative weight attached to the certifier's revenue in the objective function. The key findings that a separating equilibrium exists or not depending on this weight calls into question the honesty of imperfectly competitive certifiers. Strong enough competitive pressure is needed to reach the threshold of social concern above which certification discloses truthful information, otherwise the certifier's work simply consists in extracting consumer surplus.

Furthermore, the demonstration that pooling may occur in the present setting provides a rationale for fraudulent labelling as postulated by Hamilton and Zilberman (2006). In a market for a green product vertically differentiated from brown production, those authors investigate second-best optimal regulation of the green industry through the use of a two-part tax scheme involving a unit fee levied on output and a lump-sum tax, combined or not with fraud control. Their work, unlike Fischer and Lyon (2014), deals with imperfect competition within the green industry rather than between certification bodies. My analysis completes Hamilton and Zilberman (2006) in several respects. Most importantly, the focus is on the emergence of endogenous costs of signaling the environmental attribute of the good. These costs add to the overall costs of labelling the good as green because they are needed to prevent fraudulent labelling, which Hamilton and Zilberman (2006) overlook. In addition, they assume the existence of exogenous costs of displaying wrong labels. Here, instead of paying for cheating, the certifier always has incentives to engage in fraud and label the brown good as green, which makes honest certification difficult. So, the disguise costs are nil and the "crime" is a constant temptation.

Ibanez and Grolleau (2008) restrict attention to self-certification in an imperfectly competitive market. Self-certification allows two rival producers to differentiate their products in the eyes of consumers who trust the displayed eco-label. They show that truthful information is disclosed in equilibrium if misleading certification is sufficiently costly for the brown producer. The present setting differs from that paper in two respects: first, it deals with third-party certification rather than firms' self-certification, and second, it shows that the role of market prices in transmitting information about environmental quality is no less important than the role of labels alone.

Feddersen and Gilligan (2001) emphasize some reasons why certifiers sharing the view of environmental activists may be dishonest and send misleading messages to uninformed consumers. Specifically, the activist may falsely signal a green good as brown to shrink overall product sales, hence the damage to the environment. In the present setting, it may also happen that the green good is labelled as brown at pooling equilibria: indeed, it makes no difference which label is then displayed since no label is viewed as reliable. The difference from Feddersen and Gilligan (2001) is that pooling occurs for a wide range of certifier's objective functions that place a larger relative weight on the certifier's revenue. Another difference is that market prices here play a crucial role in signaling environmental quality to consumers.

Mason (2006) initiates a signaling approach to labeling in which Bayesian consumers form inferences about the severity of environmental harm upon seeing firms' decision to seek certification or not. In addition, Mason (2011) investigates the welfare effects of resorting to noisy certification. The signaling model presented here relates to both of these papers in that consumers are assumed to be Bayesian. However, Mason's models do not allow for misleading certification, hence there is no need for costly signaling to achieve honest certification. Mahenc (2009) addresses the issue of the manipulation of consumers' beliefs by a welfare-maximizing certifier in a signaling model. As a result, certification always proves to be honest and the equilibrium outcome highlights the need for costly signaling to prevent the certifier from cheating. The present work investigates further how the price signaling mechanism interplays with labelling in a setting which allows the certifier's objective to depart from welfare maximization. This results in a richer set of findings that provides more insight into both honest and misleading certification.

The signaling model of certification builds on the industrial literature on quality signaling through price, initiated by Milgrom and Roberts (1986) and Bagwell and Riordan (1991). Those authors develop signaling models in which the price set by a monopolist serves as a direct signal of product quality. In the present model, the price also plays a signaling role, however the market is competitive and the certifier is the only one to have the strategic power to disclose information. While the certification fee is not observed by consumers, it determines the level of the market price upon which consumers revise their beliefs regarding environmental quality. The extent to which information is disclosed to consumers only depends on the certifier's behavior.

In the limit case where the certifier pursues the objective of profit, the certifier closely resembles the "certification intermediary" in Lizzeri (1999), in that both certifiers manipulate information while maximizing their own revenue. However, unlike Lizzeri (1999), the traded good has some value for the certifier here, as long as consumers are willing to pay for this good. In Lizzeri (1999), there is no signaling mechanism and the seller of a product with a characteristic unknown to buyers pays a certification intermediary who is proved to disclose no information. In the present signaling model, a similar result obtains with the existence of pooling equilibria in which the market price conceals information.

Daughety and Reinganum (2008) analyze the interplay between quality certification and price signaling in a monopoly setting. Essentially, the present model departs from that previous research in two respects: first, the certifier here behaves strategically, and second, the good with an unverifiable characteristic is supplied by competitive producers who can not, on their own, use prices to disclose information.

The finding that a self-interested certification intermediary plays a purely parasitic role is consistent with real-world observations that certification of environmental quality happens to fail in its promise of truthfully signaling the best environmental choice. For instance, a recent study by Jacquet et al. (2010) shows concern about potential conflicts of interest raised by the Marine Stewardship Council (MSC) certification for sustainable fisheries. The authors draw the conclusion that some of the fisheries certified as sustainable by the non-profit organization MSC may not be environmentally sound and that a tightening of the green certification process might be needed. Dranove and Jin (2011) provide further real-world examples in which the certification system creates the following potential financial conflict of interest: certifiers that leniently interpret existing criteria might expect to receive more work and profit from ongoing annual audits. A slight difference here is that the certifier may conceal information about environmental quality because it raises more revenue from consumers' over-optimism.

In the present framework, the certifier enjoys market power and charges a fee which is not observable to consumers. These features are largely inspired by organic certification and the "Appellation d'Origine" system in the European Union. Crespi and Marette (2001) offer a detailed discussion of how different governmental agencies finance inspection and grading services regarding food quality.

Moreover, the model assumes that consumers' willingness-to-pay for the good is adversely affected by the dislike of the negative externality generated by the good, and consumers' preferences differ in that respect. A general interpretation is that taste heterogeneity reflects the degree of social conscience of consumers. More specifically, if the good is an overfished species, the taste parameter represents consumer dislike of eating endangered species and degrading healthy marine ecosystems. If the good is fossil energy, consumers may differ in their dislike of the negative impact on global warming, and if it is nuclear energy, they may differ in their dislike of the potential risks imposed on future generations by nuclear repositories. The idea that consumers feel moral obligations intimately tied with their consumption desire may stem from Kant's (1785) concept of the "categorical imperative". Early in the economic literature, Laffont (1975) introduced the Kantian imperative into economic modeling. In the present model, the bad feeling caused by the consumption of polluting goods strongly resembles preferences in Becker (1991) or Akerlof (1997) in that adverse consequences turn the economic act of consumption into a social decision. Consumers have heterogeneous tastes for polluting goods due to differences in education, social class, or simply the goose bumps of helping to create a negative externality: "cold-prickle" is the phrase used by Andreoni (1995) to recognize the existence of a feeling of guilt through public goods experiments. Symmetrically, there is evidence of a warm glow experienced by individuals who contribute to the environment. Popp (2001) shows that concern for the environment may obey altruistic motives, such as those emphasized by Adam Smith in *The Theory of Moral Sentiments* (1759):

"How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others..."

3 The model

Consider a market in which a continuum of mass 1 of competitive firms supply a product whose environmental quality is unknown to consumers. The good may be either "brown" or "green" depending on whether it harms the environment or not. Risk-neutral consumers derive a common gross surplus of value u for the good, regardless of whether the good is brown or green.

The total number of consumers is normalized to unity. Consumers purchase at most one unit of the good and, otherwise, they receive a net surplus u_0 from consuming outside goods.

The brown good (indexed by $i = b$) generates an external damage recognized as worrisome (pollution from chemicals, fish species exhaustion, etc.)[3] or even catastrophic (global warming)[4]. Purchasing the brown good makes consumers feel uncomfortable either because they experience health problems when exposed to environmental pollution, or because they feel guilty: every time they buy, consumers are aware that they generate some negative by-products they are not paying for. Consumers differ in their dislike of the negative externality. If the supplied good is brown, consumer taste is represented by a utility loss x which is uniformly distributed over the interval $[0, 1]$, hence a consumer with taste x has a reservation price $u - x$ for the brown good. This heterogeneity in taste may reflect cultural differences. For instance, a French wine certified as organic by France's national logo for organic products AB (Agriculture Biologique) may not be recognized as organic by the National Organic Program of the United States, and so will be seen as simply conventional by American consumers. As will be seen below,

[3] Some fisheries are publicly accused by non-governmental environmental organizations of endangering species and degrading healthy marine ecosystems. In this spirit, Greenpeace International posts a seafood red list of fish that are commonly sold in supermarkets around the world, and which have a very high risk of being sourced from unsustainable fisheries.

[4] According to Weitzman (2009), the best available climate models state that the future holds a 5 percent chance of a terrible-case scenario, that is, our current path will lead temperatures to rise more than eighteen degrees Fahrenheit (ten degrees Celsius).

the uniform distribution of x together with consumer risk-neutrality generate a linear uninformed demand for the good, which will greatly simplify the analysis without loss of intuition.

The green" good (indexed by $i = g$) has no adverse social consequences. If, for example, "green" stands for "organic" in the case of food, the green certification imposes both a percentage and a list of synthetic substances in the preparation, processing and packaging of the food. The brown food departs from a strict interpretation of "organic", e. g. by allowing a larger number of pesticides and other chemicals in the food[5]. I assume that consumer dislike of pollution vanishes when the good is truthfully certified to be green, hence u is the common reservation price for the green good. In other terms, the green good is vertically differentiated from its brown alternative.

Each firm sells one unit of the good at price p. To produce the good $i \in \{b, g\}$, firms incur a cost $c_i y$, where c_i is a positive parameter related to the environmental quality of the good, and y is an efficiency index uniformly distributed on the segment $[0, 1]$. In addition, firms must pay a fee φ to pursue certification. Thus any firm with efficiency y enters the market if and only if $p - \varphi - c_i y \geq 0$, which yields the supply function $S_i(p) = \frac{p-\varphi}{c_i}$. Let normalize c_b to 1 and c_g to $c > 1$, so that the technology (production plus pollution abatement) required to produce the green good is more costly than

[5] In the United States, a product is certified organic by the Department of Agriculture if no more than 5 percent of the product incorporates non-organic substances, provided they are approved by the National Organic Standards Board. That list has grown from 77 to 245 substances since it was created in 2002. (see Stephanie Strom, "Has 'Organic' Been Oversized?", *New York Times*, July 7, 2012).

that used for the brown good[6].

The good is a "credence" good in the sense of Darby and Carny (1973): consumers never perfectly observe the social damage generated by the brown good, hence they cannot ascertain its actual environmental quality in the absence of certification[7]. For instance, consumers are unable to recognize to what extent a product includes synthetic chemicals, antibiotics, or hormones, unless it is certified "organic" by the national program. To simplify, I assume that consumers are unsure whether the available good is brown or green before purchase. Initially, without any certification, the good is believed to be green with probability $\mu^{\circ} \in (0, 1)$. Unlike consumers, firms perfectly know their environmental practices, hence the actual damage caused to the environment by the good they supply.

A certifier is responsible for delivering the labels "brown" or "green". The certifier acquires perfect knowledge of environmental quality at negligible costs, and infers from this the actual cost of producing the good. Therefore, the certifier and the firms share the same information about the production costs and the environmental damage. The certifier charges firms a fee φ, which is not observed by consumers. This is an important assumption. Indeed, if the fee were publicly observable, it might become a direct

[6]The statement that there is a trade-off between environmental improvements and firm efficiency is consistent with the conclusions of Palmer, Oates and Portney (1995) or Jorgenson and Wilcoxen (1990) for the U.S. economy.

[7]Such an assumption is consistent with the observation made by Karl and Orwat (2000) that the individual's cost of ascertaining the environmental characteristics of goods is likely to be prohibitive for consumers.

signal regarding the environmental damage and the market price would not provide further information. If, on the contrary, consumers cannot observe the fee, they make their purchase decision after observing only the market price, which therefore becomes the sole relevant source of information. One simple interpretation of this assumption is that the certification fee is part of a contract to which consumers have no access. I deliberately abstract from efficacy costs of certification such as monitoring or inspection costs, to highlight the emergence of endogenous costs of signaling environmental quality.

Observing the market price p as a whole, consumers have no way of isolating the producer price. However, they understand that the green good is more costly to produce. Consumers' posterior beliefs will be denoted by $\mu(p) : R^+ \to [0, 1]$, giving the probability consumers attach to the possibility that the good is green after observing a price p.

Timing.—The timing of the game is as follows. In stage 1, Nature chooses the good $i \in \{b, g\}$ and the corresponding production cost c_i; the certifier and the firms perfectly observe c_i. In stage 2, the certifier charges firms φ and awards them a label; firms decide whether to pay the fee and have the good certified or stay out. In stage 3, consumers observe the market price and the label color; they revise their beliefs and make their purchase decision.

Demand for the credence good and market equilibrium.—The market size is determined by the taste of the consumer indifferent between buying or not. If consumers' inference process yields posterior beliefs $\mu = \mu(p)$, a consumer with taste x derives an expected surplus $v-p-(1-\mu)\,x$ from purchasing the unknown good at price p, where $v \equiv u - u_0$ denotes the effective reservation price. This generates the following piecewise-linear demand function:

$$X\left(p,\mu\right)=\begin{cases} 0 \text{ if } p \geq v, \\ \frac{v-p}{1-\mu} \text{ if } v-(1-\mu) \leq p \leq v, \\ 1 \text{ if } p \leq v-(1-\mu)\,. \end{cases} \tag{1}$$

When the good is brown for sure, demand will be defined by $X\left(p,0\right) = v-p$ for all $p \in [v-1, v]$, and when the good is green for sure, demand is given by $X\left(p,1\right) = 1$ for all $p \leq v$. I will assume that all consumers purchase the green good sold at marginal cost, i. e., $v > c$.

The market clears at that price $p_i\left(\varphi,\mu\right)$ that equates the supply to the demand, that is,

$$S_i(p) = X\left(p,\mu\right), \tag{2}$$

which yields the equilibrium price

$$p_i\left(\varphi,\mu\right)=\begin{cases} \frac{vc_i+(1-\mu)\varphi}{1-\mu+c_i} \text{ if } \varphi \geq v-c_i-(1-\mu)\,, \\ c_i+\varphi \text{ otherwise,} \end{cases} \tag{3}$$

and the equilibrium allocation of the good

$$X\left(p_i\left(\varphi,\mu\right),\mu\right)=\begin{cases} \frac{v-\varphi}{1-\mu+c_i} \text{ if } \varphi \geq v-c_i-(1-\mu)\,, \\ 1 \text{ otherwise.} \end{cases} \tag{4}$$

Note that there exists a threshold fee $\widetilde{\varphi}_i(\mu) \equiv v - c_i - (1-\mu)$ below which the market is fully covered in equilibrium, and raising the certification fee only results in pushing up the equilibrium price of the good. Above this threshold, an increase in φ reduces the equilibrium sales volume, in addition to raising the market price. Moreover, when the market is not fully covered in equilibrium ($\varphi > \widetilde{\varphi}_i(\mu)$), both the market price and the market size increase as consumers' beliefs as to environmental quality are more optimistic (higher μ). In contrast, consumers' perceptions of environmental quality do not affect the market equilibrium when the market is fully covered. I will assume that

$$\widetilde{\varphi}_g(0) = v - c - 1 > 0, \tag{5}$$

so that the certifier will always have the option of covering the whole market or not via the choice of fees.

Social welfare.—The social welfare is the consumer surplus (net of the bad feeling about pollution) less the aggregate production cost. Given that consumers' beliefs are given by μ, the expected social welfare in the market equilibrium, when the actual good is i, has the following expression with respect to φ:

$$W^i(\varphi,\mu) = \int_0^{X(p_i(\varphi,\mu),\mu)} [v - (1-\mu)\,x - c_i x - \varphi]\,dx = \begin{cases} \frac{(v-\varphi)^2}{2(1-\mu+c_i)} \text{ if } \varphi \geq \widetilde{\varphi}_i(\mu), \\ v - \varphi - \frac{1-\mu+c_i}{2} \text{ otherwise.} \end{cases} \tag{6}$$

Not surprisingly, social welfare decreases with the certification fee. As the certification process is assumed to be costless, charging a fee is undesirable from the social standpoint. Hence, the truth about the environmental quality of the good should be revealed at no cost if the certifier were socially responsible and honest, or, which is equivalent here, if there were perfect competition between certifiers.

The certifier's objective.—The certifier raises $R^i(\varphi, \mu) = \varphi X(p_i(\varphi, \mu), \mu)$ in expected revenue. Maximizing this revenue would be the only objective, were the certifier self-interested. Nevertheless, the objective assumed for the certifier is intended to be descriptive of the goal of a certification body submitted to an oversight board that includes representatives from the government and the industry. Though concerned with the revenue resulting from the firms' output, the certifier is likely to have social concern to at least some extent. Therefore, I will write the certifier's objective function U as a convex combination of social welfare and the certifier's revenue:

$$U \equiv \alpha W + (1 - \alpha) R, \tag{7}$$

where $\alpha \in [0, 1]$ represents the certifier's social concern. Another interpretation is that α reflects the degree of competition in the certification sector. As α decreases, the certifier's objective moves from the welfare-maximizing benchmark ($\alpha = 1$) to that of a monopolistic for-profit certifier ($\alpha = 0$). In market equilibrium, the certifier's objective takes the following functional forms

$$U^i(\varphi,\mu)=\begin{cases}\frac{v-\varphi}{2(c_i+1-\mu)}\left(\alpha v+(2-3\alpha)\varphi\right) \text{ if } \varphi\geq\widetilde{\varphi}_i(\mu),\\ \frac{\alpha}{2}(2v-(c_i+1-\mu))+(1-2\alpha)\varphi \text{ otherwise.}\end{cases} \tag{8}$$

These functions are depicted in Figure 1 for different values of α. Two cases can be distinguished depending on α. First, when $\alpha<\frac{1}{2}$, the certifier is relatively more driven by the need for profit. The certifier's utility increases as long as the market is fully covered, that is, $\varphi<\widetilde{\varphi}_i(\mu)$, and otherwise reaches a maximum at $\max\left\{\widetilde{\varphi}_i(\mu),\frac{(1-2\alpha)v}{2-3\alpha}\right\}$. This is because a higher φ raises the certifier's revenue, unless the market size shrinks due to excessive fee levels. Under (5), the optimal fee is positive regardless of the actual environmental quality and consumers' beliefs, meaning that the certifier exploits its market power, using the fee to extract consumer surplus for its own benefit. Second, when $\alpha>\frac{1}{2}$, the certifier's utility is strictly decreasing with φ because the certification fee is worthless from the social standpoint. In other terms, were certification honest, the optimal fee would be nil, so the market would implement by itself the first-best optimal allocation. Generally, I will denote the optimal fee by $\varphi^i(\mu)$ and write $U^i(\mu)\equiv U^i(\varphi^i(\mu),\mu)$.

Note that the certifier's objective is computed from the consumers' beliefs rather than the actual environmental quality[8]. The worse the consumers feel

[8] To some extent, the certifier fits into the category of the so-called "populist" organization in that it takes people's beliefs into account, as opposed to the "paternalist" organization which ignores these beliefs. This distinction can be found in Salanié and Treich (2006) for instance. More generally, the assumption of a biased bureaucracy is reminiscent of Niskanen (1971) and consistent with the work of Prendergast (2007) on the

(low values of μ), the lower the certifier's utility, regardless of its degree of social concern.

The certifier knows that the choice of certification fee will influence the market price from which consumers make their inference about the environmental quality of the good. As the market size $X\left(p_i\left(\varphi,\mu\right),\mu\right)$ is increasing in μ, the certifier has an incentive to misrepresent its information when the actual good is brown. If such is the case, consumers feel bad for the negative externality they help to generate, and the certifier would like consumers to mistake the brown good for the green one in order to boost consumption when the market is not fully covered in equilibrium. In an attempt to fool consumers, the certifier might award firms the green label while influencing the market price $p_i\left(\varphi,\mu\right)$ through φ so that consumers are more willing to purchase the good than they would be were the truth known to them. But consumers should know that this is how the certifier will reason. If consumers want the truth from the certifier, they must prevent the certifier from jamming the market signal.

In particular, the green label must be reliable in the sense that it is consistent with the information on environmental quality transmitted by the market price. As consumers can try to infer the true environmental quality of the good from observing $p_i\left(\varphi,\mu\right)$, this price plays the role of a signal. The

potential differences in preference between bureaucrats and the public Another instance of biased bureaucracy can be found in Viscusi and Hamilton (1999). These authors provide convincing evidence that risk regulators often reason on the basis of people's perception about risks rather than the actual risks.

central question then is whether the market price can credibly signal the actual environmental quality to consumers.

Note that competitive firms are unable by themselves to use price as a direct signal for environmental quality. As argued by Spence (1977), raising price is costless to competitive firms: if consumers' perceptions of environmental quality rose with price, every firm would raise the price to signal higher environmental quality at no cost. Therefore, the market price under price competition cannot usually serve the same signaling role as if the price were set by imperfectly competitive firms[9]. Here, in contrast, the market price $p_i(\varphi, \mu)$ embodies the certification fee which can be used to influence the signal of environmental quality.

Equilibrium concept.—The model has a signaling structure *à la* Spence (1973). While charging a fee for each type i, the certifier sends a signal to consumers through the market price $p_i(\varphi_i)$. Restricting attention to pure strategies in perfect Bayesian equilibrium, I will distinguish between separating and pooling equilibria. In separating equilibria, the market prices truthfully signal whether the good is brown or green, which prevents the certifier from cheating by granting the wrong label. For this, the certifier ought to satisfy incentive-compatibility (IC) constraints in addition to individual-rationality (IR) constraints. The IC constraints guarantee the credibility of the signaling behavior in the sense that the price signal sent by the market

[9] See, for instance, Mahenc (2008) for an analysis of the monopoly price as a signal of a firm's environmental performance.

is consistent with the label granted by the certifier. In pooling equilibria, the market prices are the same for the brown and the green good, so that no information is disclosed to consumers. Pooling tarnishes the honesty of certification. A perfect Bayesian equilibrium is a set of strategies $\{(\varphi_i^*)_{i=c,g}\}$ and a probability distribution $\mu^*(p_i(\varphi_i))$ such that strategies must be optimal given the consumers' beliefs, and consumers form posterior beliefs from their prior beliefs μ° by using Bayes's rule. (See Appendix 1 for a formal definition).

4 The issue of misleading certification and a five-finger exercise with $\alpha = 0$

As a five-finger exercise, I apply the signaling model to the issue mentioned in the Introduction about the reliability of organic certification under the pressure of corporate interests. Following recent newspaper articles (see Footnote 2), I assume that this lobbying tends to align the certifier's objective with the desire for the industry to grow, which, in the present model, amounts to setting $\alpha = 0$. The aim is to show that the certifier can in fact falsely claim that the food is organic in order to jam the market signal and let more products in. This is for illustration purpose only, a thorough analysis of the general model is conducted in Section 5. The self-interested certifier maximizes the following revenue

$$R^i(\varphi_i,\mu)=\begin{cases}0 \text{ if } \varphi_i\geq v,\\ \varphi_i(\frac{v-\varphi_i}{c_i+1-\mu}) \text{ if } v-c_i-(1-\mu)\leq\varphi_i\leq v,\\ \varphi_i \text{ if } \varphi_i\leq v-c_i-(1-\mu),\end{cases} \tag{9}$$

This function is concave, with a maximum at $\varphi^i(\mu)=\max\left\{\frac{v}{2},v-c_i-(1-\mu)\right\}$. Using the notation $R^i(\mu)\equiv R^i(\varphi^i(\mu),\mu)$, the maximum revenue is

$$R^i(\mu)=\begin{cases}\frac{v^2}{4(c_i+1-\mu)} \text{ if } v\leq 2(c_i+1-\mu),\\ v-c_i-(1-\mu) \text{ otherwise.}\end{cases} \tag{10}$$

The supply is $S_i(p)=\frac{p-\varphi}{c_i}$, and the demand is given by (1). These functions are depicted in Figure 2 for $\mu=0$ and $\mu=1$.

Let us restrict attention to the case where, certification being trustworthy, the market for the organic food is fully covered ($c_i=c,\mu=1$ and $v\geq 2c$ in (10)) and the market for the conventional food is not ($c_i=1,\mu=0$ and $v\leq 4$ in (10)).

Assume that the actual food is conventional rather than organic. Knowing this, the certifier charges firms the fee $\varphi^b(0)=\frac{v}{2}$ and grants the trustworthy label "brown" if the certifier is honest. It may seem surprising to pay a positive fee for the brown label. The reason is that the certifier exploits its market power to extract consumer surplus through the fee charged to firms. Were α higher than $\frac{1}{2}$, $\varphi^b(0)$ would be nil. From (3), the market clears at equilibrium price $p_b\left(\varphi^b(0)\right)=\frac{3v}{2}$, which results in the equilibrium allocation of the good $X\left(p_b\left(\varphi^b(0)\right),0\right)=\frac{v}{4}$ being lower than 1. Setting $\varphi^b(0)=\frac{v}{2}$ yields the revenue $R^b(0)=\frac{v^2}{8}$.

Were the food organic and perceived as such by consumers thanks to the label "green", the market would clear at $p_g(\varphi) = \min\{v, c + \varphi\}$and, setting $\varphi^g(1) = v - c$, the certifier would allow the full coverage of the market at the equilibrium price $p_g(\varphi^g(1)) = v$.

However, when the food is conventional, the certifier may be tempted to grant it the green label in order to boost demand and raise more revenue. How can this be made to happen? First, displaying the green label must be consistent with the price signal sent by the market equilibrium. The certifier can manipulate this price through the fee. To persuade consumers that the food is organic, the certifier can switch to a fee δ that mimics the price $p_g(\varphi^g(1)) = v$ (see Figure 2). Consumers infer from this price that the good is organic with probability $\mu = 1$. As any firm $y \in [0, 1]$ with marginal cost $y + \delta$ lower than v enters the market, the deviant fee $\delta = v - 1$ makes it possible to fully cover the market and provide the certifier with a revenue $R^b(\delta, 1) = v - 1$ that exceeds $R^b(0) = \frac{v^2}{8}$ for all v inside $(2c, 4]$. It turns out that price manipulation is profitable to the certifier: misleading consumers by charging firms $\delta = v - 1$ achieves full coverage of the market with conventional food.

Nevertheless, consumers should suspect that the certifier may want to cheat. Besides certification, they should moreover consider the market price as a further transmission channel from which to extract information. Price signaling imposes additional requirements on the certifier's behavior relative to that of the honest certifier. The next section investigates the existence of separating and pooling equilibria in the price signaling game presented in Section 2 and shows that, in fact, there is no separating equilibrium when $\alpha \leq \frac{1}{2}$.

5 Separating and pooling equilibria

Suppose now that the objective of the certifier is to maximize revenue $U^i(\varphi_i, \mu)$ formally given in (8). I first examine separating equilibria in which market prices fully reveal whether the good is brown or green. The label granted to the good is reliable in a separating equilibrium, because the signal sent by the market price is credible. As previously mentioned, equilibrium fees $\left(\varphi_b^*, \varphi_g^*\right)$ such that $p_b\left(\varphi_b^*\right) \neq p_g\left(\varphi_g^*\right)$ must satisfy IR and IC constraints.

The two IR constraints are:

$$U^b(\varphi_b^*, 0) \geq U^b(0), \tag{11}$$

$$U^g(\varphi_g^*, 1) \geq U^g(0). \tag{12}$$

These constraints guarantee that the certifier will find it worthwhile to reveal full information via market prices and grant reliable labels. Indeed, the fees consistent with a truthful signal sent by the market price are more beneficial to the certifier than the maximum utility value obtained with the worst belief, namely $\mu = 0$, that consumers can have about environmental quality. The next lemma determines the unique fee that can signal a brown good through the market price.

Lemma 1: *In any separating equilibrium, the certifier charges* $\varphi_b^* = \varphi^b(0)$ *when the good is brown, which results in the market price* $p_b(\varphi_b^*) = \min\left\{\varphi^b(0) + 1, \frac{v+\varphi^b(0)}{2}\right\}$.

Proof: see Appendix 2.

I can henceforth write $\varphi_b^* = \varphi^b(0)$ and consider the IC constraints. They prevent the certifier from manipulating the market price to mislead consumers, thereby ensuring honesty in certification. They require that the certifier should not defect to the equilibrium fee associated with the market price signaling a different color of the good, which would make the label unreliable.

More precisely, when the good is brown, the certifier should not deviate from $\varphi^b(0)$ to δ_b that results in the market price $p_b(\delta_b, 1)$, which might falsely signal that the brown good is green. Putting $\mu = 1, c_i = 1$ and $\delta_b = \varphi$ in (3), we get $p_b(\delta_b, 1) = \min\{v, 1 + \delta_b\}$, so that any fee δ_b lower than $\widetilde{\varphi}_b(1) = v - 1$ generates full coverage of the market for the good labeled as green. The price $p_b(\delta_b, 1)$ is actually misleading if it coincides with the price $p_g\left(\varphi_g^*\right)$ intended to credibly signal the green good, that is, $p_g\left(\varphi_g^*\right) = p_g\left(\varphi_g^*, 1\right) = \min\left\{v, \varphi_g^* + c\right\}$. Specifically, at the price $p_g\left(\varphi_g^*\right)$, the market for the good labeled as green is either fully covered when $\varphi_g^* \leq \widetilde{\varphi}_g(1) = v - c$, or not when $\varphi_g^* > \widetilde{\varphi}_g(1)$.

If, on the one hand, φ_g^* results in the full coverage of the market, then, $p_g\left(\varphi_g^*\right) = c + \varphi_g^*$, and, from $p_b\left(\delta_b, 1\right) = p_g\left(\varphi_g^*\right)$, the only possible deviation is $\delta_b = \varphi_g^* + c - 1$. If, on the other hand, $\varphi_g^* > \widetilde{\varphi}_g(1)$ so that the market is not fully covered, then consumer surplus is fully extracted at price $p_g\left(\varphi_g^*\right) = v$, and any deviation $\delta_b > \widetilde{\varphi}_b(1)$ results in $p_b\left(\delta_b, 1\right) = p_g\left(\varphi_g^*\right)$.

Furthermore, when the good is green, the certifier should not be tempted by a switch from φ_g^* to δ_g such that the market posts the price $p_g\left(\delta_g, 0\right)$. This price first replicates $p_b\left(\varphi_b^*\right)$, thereby falsely signaling that the green good is brown, and second, it allows supply $S(p_g\left(\delta_g, 0\right)) = \frac{p_g(\delta_g,0)-\delta_g}{c}$ to come into line with demand $X\left(p_b\left(\varphi_b^*\right), 0\right)$, i. e., $p_g\left(\delta_g, 0\right) = \min\left\{\frac{cv+\delta_g}{c+1}, c + \delta_g\right\}$.

Formally, the IC constraints can be written as follows: for all $\delta_b, \delta_g \in [0, v]$,

$$U^b(\delta_b, 1) \quad < \quad U^b(0), \tag{13}$$

$$U^g(\varphi_g^*, 1) \quad > \quad U^g(\delta_g, 0). \tag{14}$$

Notice that both inequalities are strong because equality would result in the same market price for the brown and green goods, thereby contradicting the separation requirement. Clearly, condition (14) can be eliminated as being non-binding. By definition of $\varphi^g\left(0\right)$, we know that $U^g(\delta_g, 0) \leq U^g(0)$, which implies that (12) is more demanding than (14). This exempts the certifier from taking (14) into consideration.

Finally, the certifier is left with both incentive constraints (13) and (12). When the good is green, the certifier has two options for truthtelling: the certifier can either choose a fee $\varphi_g^* \leq \widetilde{\varphi}_g(1)$ that credibly signals the green good through the market price $p_g\left(\varphi_g^*\right) = c + \varphi_g^*$ which results in the full coverage of the market, or the certifier can choose a fee $\varphi_g^* > \widetilde{\varphi}_g(1)$ that credibly signals the green good through the market price $p_g\left(\varphi_g^*\right) = v$ which only allows partial coverage of the market. As previously seen, the first option gives the certifier the incentive to deviate from $\varphi_b^* = \varphi^b(0)$ to $\delta_b = \varphi_g^* + c - 1$ when the good is brown. Taking this into account, the IC constraint (13) can be rewritten

$$U^b(\varphi_g^* + c - 1, 1) < U^b(0). \tag{15}$$

If the certifier chooses the alternative option of setting $\varphi_g^* > \widetilde{\varphi}_g(1)$, then the IC constraint (13) must hold for any deviation $\delta_b > \widetilde{\varphi}_b(1)$ which would result in the misleading price $p_b\left(\delta_b, 1\right) = v$.

The next proposition summarizes the previous discussion.

Proposition 1: *There exists a separating equilibrium if and only if*

- $\varphi_b^* = \varphi^b(0)$,
- φ_g^* *satisfies* (12) *and:*
- *either* φ_g^* *satisfies* (15) *when* $\varphi_g^* \leq \widetilde{\varphi}_g(1)$ *so that* $p_g\left(\varphi_g^*\right) = c + \varphi_g^*$ *credibly signals the green good,*
- *or, for all* $\delta_b > \widetilde{\varphi}_g(1)$, (13) *holds when* $\varphi_g^* > \widetilde{\varphi}_g(1)$ *so that* $p_g\left(\varphi_g^*\right) = v$

credibly signals the green good.

It is useful to examine separately the two cases $\alpha \leq \frac{1}{2}$ and $\alpha > \frac{1}{2}$, since the certifier may be better off raising the fee in the first case, while this is always worse in the second case.

Non-existence of separating equilibria when the certifier is driven more by profit than by social welfare $\left(\alpha \leq \frac{1}{2}\right)$. Under the assumption that $\alpha \leq \frac{1}{2}$, the certifier's objective function given in (8) is concave with respect to φ for all $\mu \in [0,1]$, with a maximum at $\varphi^i(\mu) = \max\left\{\widetilde{\varphi}_i(\mu), \frac{(1-2\alpha)v}{2-3\alpha}\right\}$. From Proposition 1, we know that $\varphi_b^* = \varphi^b(0)$, which is always strictly positive by Assumption (5). As previously seen, the fee is used to extract consumer surplus, thereby satisfying the relative need for revenue in the certifier's objective function. This is the reason why the certifier should charge firms a positive fee, so that the market price can credibly signal that the good is brown. From (3), $\varphi^b(0)$ raises the market price up to the level $p_b(\varphi_b^*) = \min\left\{\varphi^b(0) + 1, \frac{v+\varphi^b(0)}{2}\right\}$. This is the only way of making the label "brown" reliable in the sense that consumers can correctly infer from the market price that the good is brown. As demonstrated in Appendix 3, this strategy fails in equilibrium because it is not possible to find a fee φ_g^* that results in a distinct market price for the green good.

To state the next proposition, I now give a sketch of the proof detailed in Appendix 3. The IR constraint (12) defines a set $F^g = \left[\underline{\varphi}_g, \overline{\varphi}_g\right]$ of putative equilibrium fees φ_g^* which guarantee that the certifier will tell the truth about the green good. In particular, setting the lowest fee $\underline{\varphi}_g$ is the most costly way for the certifier to credibly signal the green good through market price and grant it the right label. Then, the certifier might use two different strategies to achieve separation. One strategy consists in pushing down the market price of the green good while maintaining the full market coverage. For this, the certifier charges firms a fee φ_g^* no higher than $\widetilde{\varphi}_g(1)$. The alternative strategy is to shrink overall product sales for the green good below the full market coverage while maintaining the market price at the level $p_g\left(\varphi_g^*\right) = v$, which extracts all consumer surplus. For this, the fee φ_g^* must exceed $\widetilde{\varphi}_g(1)$.

If, on the one hand, $\varphi_g^* \leq \widetilde{\varphi}_g(1)$, Proposition 1 says that φ_g^* must in addition satisfy (15) to prevent the certifier from cheating consumers with a misleading market price. Appendix 3 shows that this boils down to $\varphi_g^* < \widehat{\varphi}_b$, where $\widehat{\varphi}_b$ solves the equality version of (15). Appendix 3 also shows that $\widehat{\varphi}_b$ is weakly lower than the lowest possible informative fee $\underline{\varphi}_g$, thus no separating equilibrium exists where $\varphi_g^* \leq \widetilde{\varphi}_g(1)$.

If, on the other hand, $\varphi_g^* > \widetilde{\varphi}_g(1)$, Appendix 3 shows that the certifier can always find a deviation δ_b which does not satisfy (13). In other words, when the good is brown, the certifier always has the temptation to enhance the consumers' willingness-to-pay for this good by influencing their beliefs in a direction that contradicts the actual environmental quality. From Proposition 1, there is no way of separating the green good from the brown one in this case either.

Proposition 2: *When $\alpha \leq \frac{1}{2}$, there is no separating equilibrium.*

Detailed Proof: see Appendix 3.

The crucial assumption in signaling models is that the type of signal sender who is willing to lie finds signaling more costly than the type affected by the lie. As a result, the signal is costly, and credible because it is too costly for the liar, which is why separation of the types can be achieved in equilibrium. This assumption is not satisfied here when $\alpha \leq \frac{1}{2}$. In fact, the certifier finds it more beneficial to raise the fee for the brown label than for the green one when the market is not fully covered $\left(\frac{\partial U^b}{\partial \varphi} > \frac{\partial U^g}{\partial \varphi}\right)$, otherwise the cost of signaling is not related to the environmental quality of the good $\left(\frac{\partial U^b}{\partial \varphi} = \frac{\partial U^g}{\partial \varphi}\right)$. In other words, sending the wrong signal that the good is brown is never more expensive than sending the truthful signal that the good is green. Therefore, price signaling fails in equilibrium and the label cannot be reliable. Honest certification is not possible when the certifier is driven more by profit than by social welfare.

Existence of separating equilibria when the certifier has more social concern than self-interest $\left(\alpha > \frac{1}{2}\right)$. Under the assumption that $\alpha > \frac{1}{2}$, the certifier's objective function given in (8) is now strictly decreasing in φ and $\varphi^i(\mu) = 0$ is the maximum regardless of μ. From Proposition 1, we know that $\varphi_b^* = \varphi^b(0)$, thus $\varphi_b^* = 0$. Hence, the certifier charges no fee to label the good as brown. This results in the market price $p_b(\varphi_b^*) = 1$ from (3) together with Assumption (5). This market price in turn credibly signals that the good is brown, which makes the label reliable.

To characterize the equilibrium fees φ_g^* charged to the green firms, I build on Proposition 1. The equality version of (12) has a unique solution $\widehat{\varphi}_g$ which sends the most costly signal that the good is green through the market price $p_g\left(\varphi_g^*\right)$. In addition to the constraint $\varphi_g^* \leq \widehat{\varphi}_g$, the certifier must not be tempted to send a misleading signal when the good is brown, i. e., we must have that (13) holds. Following Proposition 1, this IC constraint is expressed differently according to whether the market is fully covered or not (see Appendix 4): either $\varphi_g^* \leq \widetilde{\varphi}_g(1)$ and, when $\alpha > \frac{1}{2}$, (13) boils down to $\varphi_g^* > \widehat{\varphi}_b$; or, $\varphi_g^* > \widetilde{\varphi}_g(1)$ and, when $\alpha > \frac{1}{2}$, (13) requires that all δ_b higher than $\widetilde{\varphi}_b(1)$ be also higher than $\widehat{\delta}_b$, where $\widehat{\delta}_b$ is defined in the Appendix.

Assume first that the certifier chooses $\varphi_g^* \leq \widetilde{\varphi}_g(1)$. The fee $\widehat{\varphi}_b$ is the minimum distortion needed to signal the green good through a market price below v, and grant it the right label. It turns out that $\widehat{\varphi}_b > \widetilde{\varphi}_g(1)$ for all $v < \frac{5\alpha-2}{2(2\alpha-1)}$. In this event, separation is not possible. If now $v \geq \frac{5\alpha-2}{2(2\alpha-1)}$, Appendix 4 shows that $\widetilde{\varphi}_g(1) > \widehat{\varphi}_g$ for all $v > c + \frac{\alpha}{2(2\alpha-1)}$, where $c + \frac{\alpha}{2(2\alpha-1)} > \frac{5\alpha-2}{2(2\alpha-1)}$. Therefore, for all $v \geq \frac{5\alpha-2}{2(2\alpha-1)}$, separation can be achieved with a fee φ_g^* inside $(\widehat{\varphi}_b, \min\left\{\widetilde{\varphi}_g(1), \widehat{\varphi}_g\right\}]$.

Assume now that the certifier sets $\varphi_g^* > \widetilde{\varphi}_g(1)$, which can possibly result in a price that credibly signals the green good provided that $\widetilde{\varphi}_g(1) < \widehat{\varphi}_g$, or, equivalently, $v < c + \frac{\alpha}{2(2\alpha-1)}$. In this case, the deviation $\widehat{\delta}_b$ is the maximum fee that misleads consumers about the green good with a market price equal to v. Thus, $\widehat{\delta}_b < \widetilde{\varphi}_b(1)$ precludes any misleading deviation which would result in the market price $p_b\left(\delta_b, 1\right) = v$. Appendix 4 shows that this inequality holds for all $v > \frac{5\alpha-2}{2(2\alpha-1)}$. Therefore, a separating equilibrium exists in the case where $\varphi_g^* > \widetilde{\varphi}_g(1)$ if and only if $\frac{5\alpha-2}{2(2\alpha-1)} < v < c + \frac{\alpha}{2(2\alpha-1)}$, with a fee φ_g^* chosen inside $(\widetilde{\varphi}_g(1), \widehat{\varphi}_g]$.

Proposition 3: *When $\alpha > \frac{1}{2}$, there exists a pair of separating equilibrium fees $\left(\varphi_c^*, \varphi_g^*\right)$ if and only if*

$$v > \frac{2-5\alpha}{1-2\alpha}. \tag{16}$$

Every pair of fees such that $\varphi_b^* = 0$ and $\varphi_g^* \in (\widehat{\varphi}_b, \widehat{\varphi}_g]$ can be supported as a separating equilibrium, where

$$\widehat{\varphi}_b = \frac{U^b(0) - \alpha(2v-1)/2}{1-2\alpha} + 1 - c \tag{17}$$

and

$$\widehat{\varphi}_g = \frac{U^g(0) - \alpha(2v-c)/2}{1-2\alpha}. \tag{18}$$

Detailed Proof: see Appendix 4.

When the certifier has more social concern than self-interest, the certifier can credibly signal environmental quality with reliable labels in the following way. In the market for the brown good, the price signal is costless, therefore the certifier charges no fee to deliver the brown label. In contrast, signaling the green good is costly because the market price, necessarily different from that available for the brown good, must discourage any masquerade. Observing this price, consumers can make successful inferences only if they can be sure that the certifier is not cheating them. This requires the certifier, either to distort the market price downward relative to what would prevail were consumers perfectly informed about environmental quality, or to shrink the total sales volume for the green good. It turns out that the certifier can prove its honesty by choosing a fee φ_g^* inside the set $(\widehat{\varphi}_b, \widehat{\varphi}_g]$. This is the only way of turning market prices into a credible signal that the good is green and of making the label reliable. Compared to the previous case where $\alpha \leq \frac{1}{2}$, the certifier has now a weaker incentive to mislead consumers into believing that the brown good is green, because the certifier places a larger relative weight on social welfare. As a result, the certifier can afford the signaling cost and counteract the incentive to cheat consumers.

When $\alpha > \frac{1}{2}$, the assumption that the cost of signaling is inversely related to the environmental quality of the good is satisfied at least when the market is not fully covered. Indeed, the certifier finds it more detrimental to raise the fee when the good is brown than when it is green $\left(\frac{\partial U^b}{\partial \varphi} < \frac{\partial U^g}{\partial \varphi}\right)$. As a result, the certifier is able to resist the temptation of cheating, and so separation of the types does succeed in equilibrium for sufficiently high values of v.

Unlike φ_b^*, any equilibrium fee φ_g^* distorts the allocation of the green good relative to the socially efficient outcome. The resulting welfare loss $U^g(1) - U^g(\varphi_g^*, 1)$ represents the cost of signaling the green good or, equivalently, the cost needed to ensure that certification is honest. This is a purely strategic cost whose only reason for existing is to prevent cheating. The minimum of this cost is at $\varphi_g^* = \widehat{\varphi}_b$.

Furthermore, it is not really surprising that there may be no separating equilibrium for some parameter values since there is no difference in the cost of signaling between the two goods when the market is fully covered. Nevertheless, one can remark that the lower limit imposed on v by (16) is decreasing in α, meaning that the greater the certifier's social concern, the less restrictive is the condition on the reservation price for the existence of separating equilibria. Another consequence of the partial lack of monotonicity of the signaling cost relative to the types is that the standard selection criteria of the signaling literature may have no bite, so that a plethora of separating equilibria cannot be excluded[10]. However, it is instructive to focus on the least-cost separating equilibrium supported by $\varphi_g^* = \widehat{\varphi}_b$. Putting $U^b(0) = \alpha(v-1)$ in (17), one can check that $\frac{\partial \widehat{\varphi}_b}{\partial \alpha} < 0$. Hence, the cost of signaling the green good decreases as the certifier has more social concern.

Corollary 1: *Honest certification entails a signaling cost which decreases as the certifier places a larger relative weight on social welfare.*

The intuition underlying this result is straightforward. As social concern increases, the certifier is simultaneously less greedy, which weakens its incentive to raise revenue by cheating. This reduces the need to distort the allocation of the green good to send a credible signal. Therefore, the certifier with a higher concern for social efficiency finds it easier to prove its honesty.

Pooling equilibria. Let us now turn to the analysis of pooling equilibrium strategies in which the certifier jams the price signal sent by the market.

[10] In fact, I have tried to adapt and apply the logic of the Cho and Kreps' (1987) intuitive criterion to this game. The result is that the selection process reduces the interval of φ_g^* but the multiplicity of equilibria remains.

Pooling occurs when consumers observe the same market price for the brown and green goods, that is, $p_g\left(\varphi_g^*\right) = p_b\left(\varphi_b^*\right)$, thereby keeping prior beliefs μ_0 unchanged by virtue of Bayes's rule. As no further information is disclosed through price, no label should be reliable. Indeed, certification is not honest in the eyes of consumers who know that there is no way of preventing the certifier from pretending that the brown good is green.

To be part of a pooling equilibrium, the pair of fees $\left(\varphi_b^*, \varphi_g^*\right)$ must satisfy conditions

$$U^b(\varphi_b^*, \mu_0) \geq U^b(0) \tag{19}$$

$$U^g(\varphi_g^*, \mu_0) \geq U^g(0) \tag{20}$$

At market equilibrium, sales volumes are the same for both goods. Moreover, to equalize the market prices for the brown and green goods, the equilibrium fees must be related to each other by $\varphi_b^* = f\left(\varphi_g^*, \mu_0\right)$, as follows from (3), where

$$f\left(\varphi_g^*, \mu_0\right) = \begin{cases} \varphi_g^* + c - 1 \text{ if } \varphi_g^* \leq \widetilde{\varphi}_g(\mu_0) = v - c - (1 - \mu_0)\,, \\ \frac{(c-1)v + (2-\mu_0)\varphi_g^*}{c+1-\mu_0} \text{ otherwise.} \end{cases} \tag{21}$$

Substituting for $\varphi_b^* = f\left(\varphi_g^*, \mu_0\right)$, (19) can be rewritten as $U^b(f\left(\varphi_g^*, \mu_0\right), \mu_0) \geq$

$U^b(0)$. Setting $\mu_0 = 1$, assume now that $U^b(f\left(\varphi_g^*, 1\right), 1) > U^b(0)$, that is,

$$U^b(0) < \begin{cases} \frac{\alpha}{2}(2v-1) + (1-2\alpha)\left(\varphi_g^* + c - 1\right) \text{ if } \varphi_g^* \leq \widetilde{\varphi}_g(1), \\ \frac{\left(v-\varphi_g^*\right)\left[(2-3\alpha)\varphi_g^* - v(2-3\alpha+2c(\alpha-1))\right]}{2c^2} \text{ otherwise.} \end{cases} \tag{22}$$

Under this assumption, there exists a solution $\mu_0 \equiv \mu_0^b \in [0,1)$ to the equality version of (19) because $U^b(., \mu_0)$ and $f\left(., \mu_0\right)$ are both increasing in μ_0. Formally,

$$U^b(f\left(\varphi_g^*, \mu_0^b\right), \mu_0^b) = U^b(0). \tag{23}$$

Similarly, assume $U^g(\varphi_g^*, 1) > U^g(0)$, that is,

$$U^g(0) < \begin{cases} \frac{\alpha}{2}(2v-c) + (1-2\alpha)\varphi_g^* \text{ if } \varphi_g^* \leq \widetilde{\varphi}_g(1), \\ \frac{v-\varphi_g^*}{2c}\left(\alpha v + (2-3\alpha)\varphi_g^*\right) \text{ otherwise.} \end{cases} \tag{24}$$

Then, there exists a solution $\mu_0^g \in [0,1)$ to the equality version of (20). The next proposition provides necessary and sufficient conditions for the existence of a pooling equilibrium.

Proposition 4: *For all $\mu_0 \geq \max\left\{\mu_0^b, \mu_0^g\right\}$, there exists a pair of pooling equilibrium fees $\left(\varphi_b^*, \varphi_g^*\right)$ such that $\varphi_b^* = f\left(\varphi_g^*, \mu_0\right)$ if and only if φ_g^* satisfies both* (22) *and* (24)*.*

An exhaustive examination of pooling equilibria in the general setting would involve tedious calculations. In what follows, I restrict attention to the case $\alpha = 0$ and characterize the pooling equilibria that result in full coverage of the market. Then, assuming (22) and (24) simply boils down to

$$\max\left\{U^b(0)+1-c, U^g(0)\right\} \leq \varphi_g^* < \widetilde{\varphi}_g(1). \tag{25}$$

The next proposition states the equilibrium outcome in this specific case.

Proposition 5: *Assume $\alpha = 0$. For all $\mu_0 \geq U^g(0)+c+1-v$, there exists a pair of pooling equilibrium fees $\left(\varphi_b^*, \varphi_g^*\right)$ such that $\varphi_g^* \in \left[U^g(0), \widetilde{\varphi}_g(\mu_0)\right]$ and $\varphi_b^* = \varphi_g^*+c-1$, for which the market is fully covered at price $p_g\left(\varphi_g^*\right) = \varphi_g^*+c$.*

Proof: see Appendix 5.

In the limit case where $\alpha = 0$, the certifier's objective function is perfectly aligned with the firms' objectives. The certifier behaves as a profit-maximizing monopolist without giving any consideration to social welfare. In that case, we know from Proposition 2 that separation is not possible in equilibrium, therefore certification fails to be honest. The certifier cannot resist the temptation to send wrong signals to consumers about the firms' performance in green production because it is more profitable to raise the fee when the good is brown than when it is green. As a result, only pooling equilibria exist, provided that consumers' prior beliefs of environmental quality are sufficiently optimistic. In these equilibria, labelling is not reliable since market prices conceal information. The reason why the pooling strategy is worthwhile is that the certifier extracts more surplus from consumer optimism when the good is brown. The resulting increase in consumers' willingness to pay for the good may also boost the production of brown goods under some circumstances.

6 Conclusion

The recent literature on certification has unfolded along two complementary paths: one considers that imperfect certification provides an information whose accuracy depends on the efficacy costs of auditing and inspection procedures; the other path considers that misleading certification stems from firms' fraudulent practices. Along these lines, there have been few papers that question the honesty of third-party certification. The present paper has argued that third-party certification bears some responsibility for cheating uninformed consumers when the certifier is more concerned with its own revenue than social welfare. The analysis has emphasized the emergence of endogenous signaling costs that, besides the efficacy costs or the costs of fraud, are needed to ensure honest certification and prevent cheating.

In the proposed model, consumers infer information on environmental quality both from market prices and third-party certification. The analysis investigates a signaling model in which honest certification displays reliable labels in separating equilibria. These equilibria are supported by distinct market prices that credibly signal whether the good is brown or green. In contrast, pooling equilibria preclude honest certification in that they involve the same misleading price for the two types of good.

As a result, separating equilibria exist— and hence certification is honest

—only if the certifier is driven more by social welfare than by profit. Moreover, the signaling cost increases as the certifier attributes a larger relative weight to its own revenue. In the reverse case, where the certifier has more self-interest than social concern, only pooling occurs, hence certification is misleading, provided that consumers sufficiently believe the good to be green before purchase. In this case, the cost of credibly signaling the green good through price is too high to counteract the certifier's incentive to jam the price signal and cheat consumers.

One limitation of the model involves the exogenous combination of profit and welfare in the certifier's objective function. The relative weights given to the joint interests of the certifier and the firms on one side, and to the social interest on the other, roughly reflect the composition of the certifier's oversight board. The weighting of the diverging objectives could be endogenized with a lobbying pressure model in the spirit of Grossman and Helpman (1994) so as to explain which one dominates among the representatives of the different sectors. In the present setting, a power imbalance in favour of the industry interest results in misleading certification. There are likely to be a number of reasons for this imbalance. First, the industry representatives have more bargaining power than the other representatives within the oversight board because they are more efficient at capturing the certification process. Second, the assumption that the certifier's revenue is derived entirely from the fees charged to the firms generates a conflict of interest similar to that involving credit rating agencies (CRA). On this issue, Mathis et al. (2009)

conclude that it may take a long time to detect an opportunistic CRA and the conflict of interest is not solved by reputation concerns. Extending their model to the case where the screening technology of the CRA is imperfect, they also find that the opportunistic CRA is never disciplined at equilibrium. Similarly, the present model could be extended to the case where the certifier imperfectly observes whether the good is brown or green, so that the failure to disclose truthful information is no longer only due to the excessive weight given to profit maximization, but can also come from a socially concerned certifier having observed a noisy signal. However, this would not suppress the need for costly signaling to prevent cheating and guarantee honest certification. As shown in the present paper, honest certification is more likely to occur when the cost of signaling the green good decreases. So the main concern is how to reduce this signaling cost. If, for one reason or another, the monitoring scheme increases more the signaling cost for the brown good than for the green good, then cheating will be more difficult, though not fully removed.

Another interesting extension would be to endogenize competition between multiple certifiers in the same way as Fischer and Lyon (2014). Indeed, the present findings suggest that there is a competition threshold in the certification sector above which certification is honest, and moreover, the cost of signaling the green good decreases as competition between certifiers is more intense. Further investigation of the strategic interaction between certifiers is needed to examine the following issues. First, to what extent are dishonest certifiers excluded by increased competition between certifiers? Second, competition between certifiers may allow firms as well as environmentalists to seek and reward the most sympathetic to their interest among the certifiers regardless of honest certification, giving rise to undesirable shopping for labels.

Lastly, the signaling model of the type analyzed here assumes that privately informed firms are competitive, which prevents them from using price as a signaling device. Further research should examine imperfectly competitive firms instead. When a firm has some leeway in selecting price, the firm may have an incentive to convey some amount of its private information to consumers through price. This alternative signal adds to the signal sent by the certifier, thereby raising a two-sided self-selection problem. This is a complex issue because the firm and the certifier must coordinate the signals they send. The appropriate signaling game would then have a structure similar to that investigated by Bagwell and Ramey (1991) in the context of limit pricing allowing for multiple incumbents with common, private information.

7 Appendix

7.1 Appendix 1: Formal definition of the perfect Bayesian equilibrium

Formally, the perfect Bayesian equilibrium of this game requires that, for each $i = c, g$,

$$\varphi_i^* \in \arg\max_{\varphi_i} U^i(\varphi_i, \mu^* (p_i (\varphi_i))) \tag{26}$$

Consumers form posterior beliefs from their prior beliefs by using Bayes's rule:

- If $p_g \left(\varphi_g^*\right) \neq p_b (\varphi_b^*)$, then $\mu^* \left(p_g \left(\varphi_g^*\right)\right) = 1$ and $\mu^* (p_b (\varphi_b^*)) = 0$.

  From (3), the resulting market prices are

$$\begin{aligned} p_b (\varphi_b^*) &= p_b (\varphi_b^*, 0) = \min \left\{ \varphi_b^* + 1, \frac{v + \varphi_b^*}{2} \right\} && (27) \\ p_g \left(\varphi_g^*\right) &= p_g \left(\varphi_g^*, 1\right) = \min \left\{ \varphi_g^* + c, v \right\} && (28) \end{aligned}$$

- If $p_g \left(\varphi_g^*\right) = p_b (\varphi_b^*)$, then $\mu^* \left(p_g \left(\varphi_g^*\right)\right) = \mu^* (p_b (\varphi_b^*)) = \mu^\circ$.

7.2 Appendix 2: Proof of Lemma 1

Putting $\mu = 0, c_i = 1$ and $\varphi = \varphi^b(0)$ in (3) yields the market price $p_b \left(\varphi^b(0), 0\right) = \min \left\{ \varphi^b(0) + 1, \frac{v + \varphi^b(0)}{2} \right\}$. We will argue by contradiction, and suppose that there exists a separating equilibrium in which $\varphi_b^* \neq \varphi^b(0)$. As the resulting market price $p_b(\varphi_b^*)$ also signals in equilibrium that the good is brown, consumers perfectly identify environmental quality, and the certifier's revenue is $U^b(\varphi_b^*, 0)$, which is strictly lower than $U^b(0)$. Consequently, the certifier has an incentive to deviate to $\varphi^b(0)$ whatever the consumers' inference μ from observing $\varphi^b(0)$ since, for any $\mu \in (0, 1]$, we have $U^b(0) < U^b(\varphi^b, \mu)$. Thus, no fee φ_b^* different from $\varphi^b(0)$ can be part of a separating equilibrium.

7.3 Appendix 3: Proof of Proposition 2

Assume that $\alpha \leq \frac{1}{2}$. From (8), the certifier's objective function is concave with respect to φ, for all $\mu \in [0, 1]$, with a maximum at $\varphi^i(\mu) = \max\left\{\widetilde{\varphi}_i(\mu), \frac{(1-2\alpha)v}{2-3\alpha}\right\}$. Given the most pessimistic belief $\mu = 0$, the maximum utilities the certifier can get with the green and brown goods are, respectively,

$$U^g(0) = \begin{cases} \frac{[(\alpha-1)v]^2}{2(2-3\alpha)(c+1)} \text{ if } \alpha \leq \widetilde{\alpha}_g, \\ (1-\alpha)\,v + (3\alpha - 2)\,\frac{c+1}{2} \text{ if } \widetilde{\alpha}_g \leq \alpha \leq \frac{1}{2}, \\ \text{where } \widetilde{\alpha}_g = \frac{2(c+1)-v}{3(c+1)-v}, \end{cases} \tag{29}$$

and

$$U^b(0) = \begin{cases} \frac{[(\alpha-1)v]^2}{4(2-3\alpha)} \text{ if } \alpha \leq \widetilde{\alpha}_b, \\ (1-\alpha)\,v + 3\alpha - 2 \text{ if } \widetilde{\alpha}_b \leq \alpha \leq \frac{1}{2}, \\ \text{where } \widetilde{\alpha}_b = \frac{4-v}{6-v} < \widetilde{\alpha}_g. \end{cases} \tag{30}$$

Given the most optimistic belief $\mu = 1$ when the good is i, the certifier's objective function is

$$U^i(\varphi, 1) = \begin{cases} \frac{v-\varphi}{2c_i}\left(\alpha v + (2-3\alpha)\varphi\right) \text{ if } \varphi \geq \widetilde{\varphi}_i(1), \\ \frac{\alpha}{2}(2v - c_i) + (1-2\alpha)\varphi \text{ otherwise,} \end{cases} \tag{31}$$

where $\widetilde{\varphi}_i(1) = v - c_i$.

From Proposition 1, we know that the certifier must set $\varphi_b^* = \varphi^b(0)$ to credibly signal the brown good through market price. In addition, the fee φ_g^* charged for granting the green label to the green good must satisfy the IR constraint(12). In the case where $\alpha \leq \frac{1}{2}$, the equality version of (12) gives the upper and lower bounds $\overline{\varphi}_g$ and $\underline{\varphi}_g$ of the interval F^g of putative equilibrium fees φ_g^*. Given this constraint, the certifier chooses between two options for credibly signaling the green good: the certifier can set φ_g^* either below or above $\widetilde{\varphi}_g(1) = v - c$.

Consider first that $\varphi_g^* \leq v - c$.

Then, $\underline{\varphi}_g$ solves the following equation in φ:

$$\frac{\alpha}{2}(2v - c) + (1-2\alpha)\varphi = U^g(0), \tag{32}$$

and $\varphi_g^* \geq \underline{\varphi}_g$ is a further requirement for the existence of a separating equilibrium. Moreover, from Proposition 1, φ_g^* must also satisfy the IC constraint (15). Putting $i = b, \mu = 1, c_i = 1$ and $\varphi = \varphi_g^* + c - 1$ in (31) helps rewrite inequality(15) as:

$$\frac{\alpha}{2}(2v-1)+(1-2\alpha)(\varphi_g^*+c-1)<U^b(0). \tag{33}$$

Straightforward computations yield that (33) holds if $\varphi_g^* < \widehat{\varphi}_b$ where $\widehat{\varphi}_b$ denotes the solution in φ_g^* of the equality version of (33).

Thus, a necessary condition for the existence of a separating equilibrium fee φ_g^* is that $\underline{\varphi}_g < \widehat{\varphi}_b$, where $\underline{\varphi}_g$ and $\widehat{\varphi}_b$ are respectively defined by (32) and the equality version of (33).

To check whether this inequality holds, three cases must be distinguished depending on the value of α:

1. When $\alpha \leq \widetilde{\alpha}_b$, we have $U^b(0) = \frac{[(\alpha-1)v]^2}{4(2-3\alpha)}$ and $U^g(0) = \frac{[(\alpha-1)v]^2}{2(2-3\alpha)(c+1)}$.

   Solving equation $\underline{\varphi}_g - \widehat{\varphi}_b = 0$ for v gives an upper and lower root $\overline{v} = \frac{\sqrt{2(c+1)}(2-3\alpha)}{1-2\alpha}$ and $\underline{v} = -\frac{\sqrt{2(c+1)}(2-3\alpha)}{1-2\alpha}$ such that, for all $\alpha \leq \widetilde{\alpha}_b = \frac{4-v}{6-v}$, we have $v \in (\underline{v}, \overline{v})$. It follows that $\underline{\varphi}_g - \widehat{\varphi}_b > 0$ for all $\alpha \leq \widetilde{\alpha}_b = \frac{4-v}{6-v}$, thus no separation can be achieved with $\varphi_g^* \leq v - c$ in this parameter configuration.

2. When $\widetilde{\alpha}_b \leq \alpha \leq \widetilde{\alpha}_g$, we have $U^b(0) = (1-\alpha)\,v + 3\alpha - 2$ and $U^g(0) = \frac{[(\alpha-1)v]^2}{2(2-3\alpha)(c+1)}$.

   After some calculations, we obtain that $\underline{\varphi}_g - \widehat{\varphi}_b = \frac{[(3\alpha-2)(c+1)+(1-\alpha)v]^2}{2(1-2\alpha)(2-3\alpha)(c+1)} > 0$. Again, no separating equilibrium exists with $\varphi_g^* \leq v - c$.

3. When $\widetilde{\alpha}_g \leq \alpha \leq \frac{1}{2}$, we have $U^b(0) = (1-\alpha)\,v + 3\alpha - 2$ and $U^g(0) =$

$(1-\alpha)\, v + (3\alpha - 2)\, \frac{c+1}{2}$.

Straightforward calculations give that $\underline{\varphi}_g = \widehat{\varphi}_b$, and so there is no separating equilibrium such that $\varphi_g^* \leq v - c$ in that case either.

Consider now that $\varphi_g^* > v - c$.

From Proposition 1, a necessary condition to achieve separation with a fee φ_g^* that signals the green good through the market price $p_g\left(\varphi_g^*\right) = v$ is that (13) holds for all $\delta_b > \widetilde{\varphi}_b(1) = v - 1$. In the case where $\alpha \leq \frac{1}{2}$, the equality version of (13) admits an upper and lower root denoted by $\overline{\delta}_b$ and $\underline{\delta}_b$, respectively. After putting $i = b, \mu = 1, c_i = 1$ and $\varphi = \delta_b$ in (31), the equality version of (13) can be rewritten

$$\frac{v - \delta_b}{2}\left(\alpha v + (2 - 3\alpha)\delta_b\right) = U^b(0). \tag{34}$$

Two cases must now be distinguished, depending on the value of α:

1. When $\alpha \leq \widetilde{\alpha}_b$, we have $U^b(0) = \frac{[(\alpha-1)v]^2}{4(2-3\alpha)}$ and $\varphi^b\,(0) = \frac{(1-2\alpha)v}{2-3\alpha} > \widetilde{\varphi}_b(0) = v - 2$.

   It is always possible to find a deviation δ_b inside $\left[\max\left\{v-1, \underline{\delta}_b\right\}, \overline{\delta}_b\right]$ for which (13) does not hold. Thus, in that case, no separating equilibrium exists with $\varphi_g^* > v - c$.

2. When $\widetilde{\alpha}_b \leq \alpha \leq \frac{1}{2}$, we have $U^b(0) = (1-\alpha)\, v + 3\alpha - 2$ and $\varphi^b\,(0) = \widetilde{\varphi}_b(0) = v - 2$.

In that case, (13) does not hold for any δ_b inside $\left[v-1,\overline{\delta}_b\right]$. Again, there is no separating equilibrium such that $\varphi_g^* > v - c$.

7.4 Appendix 4: Proof of Proposition 3

Assume that $\alpha > \frac{1}{2}$. From (8), the certifier's objective function is strictly decreasing in φ for all $\mu \in [0,1]$, hence the maximum is at $\varphi^i(\mu) = 0$. Given the most pessimistic belief $\mu = 0$, the maximum utilities the certifier can get with the green and brown goods are, respectively,

$$U^g(0) = \alpha(v - \frac{c+1}{2}) \tag{35}$$

and

$$U^b(0) = \alpha(v-1) \tag{36}$$

From Proposition 1, the fee φ_g^* chosen to signal the green good must satisfy the IR constraint (12). In the case where $\alpha > \frac{1}{2}$, the equality version of (12) has a unique solution $\widehat{\varphi}_g$ in φ. Thus, any $\varphi_g^* \leq \widehat{\varphi}_g$ is a putative equilibrium fee. Moreover, from (31), we know that

$$U^g(\varphi, 1) = \begin{cases} \frac{v-\varphi}{2c}\left(\alpha v + (2-3\alpha)\varphi\right) \text{ if } \varphi \geq \widetilde{\varphi}_g(1), \\ \frac{\alpha}{2}(2v-c) + (1-2\alpha)\varphi \text{ otherwise,} \end{cases} \tag{37}$$

where $\widetilde{\varphi}_g(1) = v - c$. Hence, signaling the green good with $\varphi_g^* \leq v - c$ results in the full coverage of the market, while the industry output is reduced below

1 with a fee $\varphi_g^* > v - c$.

Consider first that $\varphi_g^* \leq v - c$.

Then, truthtelling requires $\varphi_g^* \leq \min\left\{v - c < \widehat{\varphi}_g\right\}$, where $\widehat{\varphi}_g$ solves (32). Straightforward calculations show that $v - c < \widehat{\varphi}_g$ for all $v < c + \frac{\alpha}{2(2\alpha-1)}$.

Moreover, we have previously seen that the IC constraint (15) boils down to (33). When $\alpha > \frac{1}{2}$, (33) holds if $\varphi_g^* > \widehat{\varphi}_b$, where $\widehat{\varphi}_b$ is defined in Appendix 3. Thus, a necessary condition for the existence of a separating equilibrium fee φ_g^* is that $\widehat{\varphi}_b < \min\left\{v - c < \widehat{\varphi}_g\right\}$.

It turns out that $\widehat{\varphi}_b > v - c$ for all $v < \frac{5\alpha-2}{2(2\alpha-1)}$, and $\frac{5\alpha-2}{2(2\alpha-1)} < c + \frac{\alpha}{2(2\alpha-1)}$ when $c > 1$. We can conclude:

- If $v \leq \frac{5\alpha-2}{2(2\alpha-1)}$, then no separating equilibrium exists with $\varphi_g^* \leq v - c$.
- If $\frac{5\alpha-2}{2(2\alpha-1)} < v \leq c + \frac{\alpha}{2(2\alpha-1)}$, every fee φ_g^* inside $(\widehat{\varphi}_b, v - c]$ can be supported as a separating equilibrium.
- If $c + \frac{\alpha}{2(2\alpha-1)} < v$, every fee φ_g^* inside $(\widehat{\varphi}_b, \widehat{\varphi}_g]$ can be supported as a separating equilibrium.

Consider now that $\varphi_g^* > v - c$.

As φ_g^* must be lower than $\widehat{\varphi}_g$ to satisfy (12), a necessary condition for the existence of a separating equilibrium is $v - c < \widehat{\varphi}_g$, which holds for all $v < c + \frac{\alpha}{2(2\alpha-1)}$.

Moreover, in this case, the IC constraint (13) must hold for all $\delta_b > \widetilde{\varphi}_b(1) = v - 1$. In the case where $\alpha > \frac{1}{2}$, the equality version of (13) given by (34) has a unique solution $\widehat{\delta}_b$ lower than v. To compute $\widehat{\delta}_b$, one can put $U^b(0) = \alpha(v-1)$ in (34), which yields

$$\frac{v-\delta_b}{2}\left(\alpha v + (2-3\alpha)\delta_b\right) = \alpha(v-1). \tag{38}$$

Then, (13) boils down to $\delta_b > \widehat{\delta}_b$. Hence, the existence of a separating equilibrium requires $v - 1 > \widehat{\delta}_b$. One can check that this inequality holds for all $v > \frac{5\alpha-2}{2(2\alpha-1)}$.

Hence:

- If $v \leq \frac{5\alpha-2}{2(2\alpha-1)}$, then no separating equilibrium exists with $\varphi_g^* > v - c$.
- If $\frac{5\alpha-2}{2(2\alpha-1)} < v \leq c + \frac{\alpha}{2(2\alpha-1)}$, every fee φ_g^* inside $(v-c, \widehat{\varphi}_g]$ can be supported as a separating equilibrium.
- If $c + \frac{\alpha}{2(2\alpha-1)} < v$, then no separating equilibrium exists with $\varphi_g^* > v - c$.

7.5 Appendix 5: Proof of Proposition 5

When $\alpha = 0$, any fee $\varphi_g^* \leq \widetilde{\varphi}_g(\mu_0)$ such that $\varphi_b^* = f\left(\varphi_g^*, \mu_0\right) = \varphi_g^* + c - 1$ establishes a pooling price that results in the full coverage of the market. In these circumstances, (19) boils down to

$$U^b(0) \leq \varphi_g^* + c - 1, \tag{39}$$

which holds as an equality for all $\mu_0 \geq U^b(0) + 2 - v$. Moreover, (20) is

$$U^g(0) \leq \varphi_g^*, \tag{40}$$

which holds as an equality for all $\mu_0 \geq U^g(0)+c+1-v$. The calculations below will show that $U^b(0)+1-c \leq U^g(0)$, or, equivalently, $U^g(0)+c+1-v \geq U^b(0) + 2 - v$. Consequently, for all $\mu_0 \geq \mu_0^g = U^g(0) + c + 1 - v$, any φ_g^* inside $\left[U^g(0), \widetilde{\varphi}_g(\mu_0)\right]$ is a candidate for a pooling equilibrium. Furthermore, conditions (22) and (24) simplify to $U^g(0) < v - c$ and $U^b(0) + 1 - c < v - c$, respectively. Hence, to demonstrate the existence of pooling equilibria, it suffices to check that $U^g(0) \leq v - c$ in every parameter configuration.

Three cases should be considered to show that $U^b(0) + 1 - c \leq U^g(0)$:

1. When $c < v \leq 4$, we have $U^b(0) = \frac{v^2}{8}$, $U^g(0) = \frac{v^2}{4(c+1)}$, and one can check that $U^b(0) + 1 - c \leq U^g(0)$. Moreover, $U^g(0) \leq v - c$ for all v inside $\left[2\left(1 + c - \sqrt{1+c}\right), 4\right]$. Thus, for all $\mu_0 \geq \mu_0^g$, any φ_g^* inside $\left[U^g(0), \widetilde{\varphi}_g(\mu_0)\right]$ supports a pooling equilibrium.

2. When $4 \leq v \leq 2(1+c)$, we have $U^b(0) = v - 2$, $U^g(0) = \frac{v^2}{4(c+1)}$, and again $U^b(0)+1-c \leq U^g(0)$. As, moreover, $U^g(0) \leq v-c$ for all v inside $\left[2\left(1 + c - \sqrt{1+c}\right), 2(1+c)\right]$, any φ_g^* inside $\left[U^g(0), \widetilde{\varphi}_g(\mu_0)\right]$ supports again a pooling equilibrium for all $\mu_0 \geq \mu_0^g$.

3. When $2(1+c) \leq v$, we have $U^b(0) = v - 2$, $U^g(0) = v - 1 - c$, and $U^b(0) + 1 - c = U^g(0)$. Thus, for all $\mu_0 \in (0,1)$, any φ_g^* inside

$[U^g(0), v - 1 - c + \mu_0]$ supports a pooling equilibrium.

References

[1] **Akerlof, G. A.** 1970. The Market for 'Lemons': Qualitative Uncertainty and the Market Mechanism, *Quarterly Journal of Economics* 84, 488-500.

[2] **Akerlof, G. A.** 1997. Social Distance and Social Decisions, *Econometrica* 65, 1055–1027.

[3] **Amacher, G. S., E. Koskela, and M. Ollikainen**. 2004. Environmental Quality Competition and Eco-Labeling, *Journal of Environmental Economics and Management* 47, 284–306.

[4] **Andreoni, J.** 1995. Warm-Glow Versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments, *Quarterly Journal of Economics* 110, 1–21.

[5] **Bagwell, K., and G. Ramey.** 1991. Oligopoly Limit Pricing, *RAND Journal of Economics* 22, 155–172.

[6] **Bagwell, K., and M. H. Riordan.** 1991. High and Declining Prices Signal Product Quality, *American Economic Review* 81, 224–239.

[7] **Becker, G. S.** 1991. A Note on Restaurant Pricing and Other Examples of Social Influences on Price, *Journal of Political Economy* 99, 1109–1116.

[8] **Cason, T. N., and L. Gangadharan.** 2002. Environmental Labelling and Incomplete Consumer Information in Laboratory Experiments, *Journal of Environmental Economics and Management* 43, 113–134.

[9] **Cho, I-K., and D. Kreps.** 1987. Signalling Games and Stable Equilibria, *Quarterly Journal of Economics* 102, 179–221.

[10] **Crespi, J. M., and S. Marette.** 2001. How Should Food Safety Certification Be Financed? *American Journal of Agricultural Economics* 83, 852–861.

[11] **Darby, M., and E. Karny.** 1973. Free Competition and the Optimal Amount of Fraud, *Journal of Law and Economics* 16, 77–88.

[12] **Daughety, A. F., and J. F. Reinganum.** 2008. Communicating Quality: A Unified Model of Disclosure and Signalling, *RAND Journal of Economics* 39, 973–989.

[13] **Dranove, D., and G. Z. Jin.** 2010. Quality Disclosure and Certification: Theory and Practice, *Journal of Economic Literature* 48, 935–963.

[14] **Duflo, E., M. Greenstone, R. Pande, and N. Ryan**. 2013. What Does Reputation Buy? Differentiation in a Market for Third-Party Auditors, *American Economic Review* 103, 314–319.

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

[16] **Fischer, C., and T. P. Lyon**. 2014. Competing Environmental Labels. *Journal of Economics & Management Strategy* 23, 692–716.

[17] **Grossman, G., and E. Helpman.** 1994. Protection for Sale, *American Economic Review* 84, 833–850.

[18] **Hamilton, S. F., and D. Zilberman**. 2006. Green Markets, Eco-Certification, and Equilibrium Fraud. *Journal of Environmental Economics and Management* 52, 627–644.

[19] **Heyes, A. G., and J. W. Maxwell.** 2004. Private vs. Public Regulation: Political Economy of the International Environment, *Journal of Environmental Economics and Management* 48, 978–996.

[20] **Ibanez, L., and G. Grolleau.** 2008. Can Ecolabeling Schemes Preserve the Environment? *Environmental and Resource Economics* 40, 233–249.

[21] **Jacquet, J., D. Pauly, D. Ainley, S. Holt, and J. Jackson.** 2010. Seafood Stewardship in Crisis, *Nature* 467, 28–29, doi:10.1038/467028a.

[22] **Kant, E.** 1785. *Grundlegung zur Metaphysik der Sitten.*

[23] **Karl, H., and C. Orwat.** 1999. Economic Aspects of Environmental Labelling, in: H. Folmer and T. Tietenberg (eds), *Yearbook of Environmental and Resource Economics 1999/2000*, Cheltenham, UK: Elgar, pp. 107–170.

[24] **Kuhn, M.** 2005. *The Greening of Markets: Product Competition, Pollution and Policy Making in a Duopoly*. Cheltenham, UK: Elgar.

[25] **Laffont, J.-J.** 1975. Macroeconomic Constraints, Economic Efficiency and Ethics: An Introduction to Kantian Economics, *Economica* 42, 430–437.

[26] **Lizzeri, A.** 1999. Information Revelation and Certification Intermediaries, *RAND Journal of Economics* 30, 214–231.

[27] **Lyon, T. P., and J. W. Maxwell.** 2011. Greenwash: Corporate Environmental Disclosure under Threat of Audit, *Journal of Economics & Management Strategy* 20, 3–41.

[28] **Mahenc, P.** 2007. Are Green Products Over-Priced? *Environmental and Resource Economics* 38, 461–473.

[29] **Mahenc, P.** 2008. Signalling the Environmental Performance of Polluting Products to Green Consumers, *International Journal of Industrial Organization* 26, 59–68.

[30] **Mahenc, P.** 2009. Wasteful Labeling, *Journal of Agricultural & Food Industrial Organization* 7, 1–20.

[31] **Mason, C. F.** 2006. An Economic Model of Eco-Labeling, *Environmental Modelling and Assessment* 11, 131–143.

[32] **Mason, C. F.** 2011. Eco-Labeling and Market Equilibria with Noisy Certification Tests, *Environmental and Resource Economics* 48, 537–560.

[33] **Mathis, J., J. McAndrews, and J-C. Rochet.** 2009. Rating the Raters: Are Reputation Concerns Powerful Enough to Discipline Rating Agencies?, *Journal of Monetary Economics* 56, 675–677.

[34] **Milgrom, P., and J. Roberts.** 1982. Limit Pricing and Entry under Incomplete Information: An Equilibrium Analysis, *Econometrica* 50, 443–459.

[35] **Milgrom, P., and J. Roberts.** 1986. Price and Advertising Signals of Product Quality, *Journal of Political Economy* 94, 796–821.

[36] **Niskanen, W. A. Jr.** 1971. *Bureaucracy and Representative Government*. Chicago: Aldine-Atherton.

[37] **Popp, D.** 2001. Altruism and the Demand for Environmental Quality, *Land Economics* 77, 339–349.

[38] **Prendergast, C.** 2007. The Motivation and Bias of Bureaucrats, *American Economic Review* 97, 180–196.

[39] **Riley J.** 1979. Informational Equilibrium, *Econometrica* 47, 331–360.

[40] **Salanie, F. and N. Treich.** 2009. Regulation in Happyville, *The Economic Journal* 119, 665–679.

[41] **Smith, A.** 1759. *The Theory of Moral Sentiments*, in Liberty classics, edited by D. D. Raphael and A. L. Mactie. Indianapolis: Liberty Press (1976).

[42] **Spence, A. M.** 1973. Job Market Signalling, *Quarterly Journal of Economics* 87, 355–374.

[43] **Spence, A. M.** 1977. Consumer Misperceptions, Product Failure and Producer Liability, *Review of Economic Studies* 44, 561–572.

[44] **Viscusi, W. K., and J. T. Hamilton.** 1999. Are Risk Regulators Reliable? Evidence from Hazardous Waste Cleanup Decisions, *American Economic Review* 89, 1010–1027.

[45] **Weitzman M. L.** 2009. On Modeling and Interpreting the Economics of Catastrophic Climate Change, *The Review of Economics and Statistics* 91, 1–19.

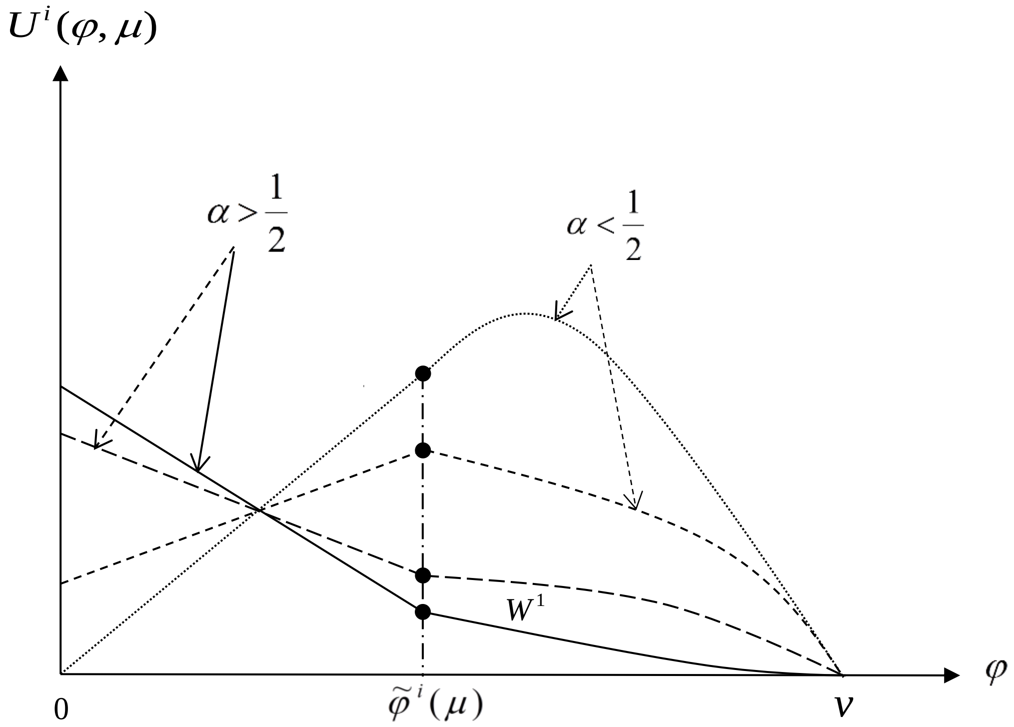

Figure 1
The certifier’s utility functions