Incentives for Quality in Friendly and Hostile Informational Environments

PRELIMINARY AND INCOMPLETE

Pierre Fleckinger[*], Matthieu Glachant[†], Gabrielle Moineville[‡]

February 24, 2012

Abstract

We develop a simple lemons model with endogenous quality where the amount of information available is quality-dependent. The distinctive feature of the analysis is to contrast friendly informational environments, in which the buyer has better information when quality is high than when it is low, and hostile environments, in which information is better when quality is low. Differences are clear-cut: Hostile environments give rise to a bandwagon effect across sellers, which can lead to multiple equilibria. In contrast, friendly environments create free riding among sellers, which always induces a unique equilibrium. Comparative statics results are also contrasted. A key notion is that incentive provision is relatively better when the informational environment targets less expected evidence. The results shed new light on several insights of the literature on statistical discrimination, collective reputation and quality certification.

Keywords: Moral Hazard, Asymmetric Information, Multiple Equilibria, Monitoring, Lemons.

JEL classification: D82.

[*]Paris School of Economics - University of Paris 1 & Cerna, Mines ParisTech. Centre d'Economie de la Sorbonne, 106-112 boulevard de l'Hôpital, 75013 Paris. Email: pierre.fleckinger@univ-paris1.fr

[†]Cerna, Mines ParisTech, 60 bd St Michel, 75006 Paris. Email: glachant@ensmp.fr.

[‡]Cerna, Mines ParisTech, 60 bd St Michel, 75006 Paris. Email: gabrielle.moineville@ensmp.fr.

Contents

1 Introduction

It is widely recognized that buyers may have less information about certain product attributes than the seller, implying severe inefficiencies. Since the publication of Akerlof's paper in 1970, this lemon problem has been explored in the economics literature from many different perspectives. A crucial yet widely overlooked aspect is that, in reality, the asymmetry of information between buyer and seller may be more or less severe depending on the level of product quality, and that this has important consequences for the sellers' incentives to provide high quality.

Consider the example of the reliability of a vehicle. Reliability basically means the absence of breakdowns. Hence, when a car actually breaks down, its quality is disclosed.[1] But if nothing occurs, it does not mean that the vehicle will not break down in the future. In this case, experience leads consumers to be more informed about the quality of poorly reliable products than about that of reliable ones.

In a similar vein, anti-doping tests in sport can only identify a limited set of performance-enhancing drugs and it is common knowledge that certain doping substances are undetectable. As a result, tests are able to provide evidence of doping, but they fail to ascertain its absence. That is, they can only uncover low quality.

In other circumstances, consumers receive more informative signals when quality is high rather than low. For example, a movie award selectively signals a high-quality product. But the consumers are left uncertain about the quality of non-awarded movies: they could be good or bad, given that not all good movies receive a prize. Academic publication is another illustration: prestigious journals (mostly) include good papers, but a fraction of the unpublished manuscripts are excellent–notably those under review/revision at good journals.

Those examples suggest that the quantity of information about a good or service available to a buyer or, more generally, to a user, frequently varies with the level of product

[1] Reliability is to a large extent an experience good attribute in that it is revealed over time after the purchase. But this does not prevent quality disclosure to influence demand either through repeated transactions or through reputation effects. Alternatively, the potential buyer can test the vehicle for a short period and obtain some information–mostly in case of a breakdown.

quality, and can be biased in either direction. In some cases, the informational environment is *friendly* - more information is available on high quality - while in others, the environment is *hostile* - more information is available on low quality. Note that these differences could be either related to the very nature of the evaluation technology (reliability, doping) or to the selective disclosure of evaluation results (movie awards, academic publications). To clarify things on the notions of friendly and hostile environments, one can think of the monitoring technology as a generic supervisor, whose attitude is either friendly, in which case it has a tendency to put more emphasis on good news (perhaps by withholding bad news), or hostile, in which case more emphasis is put on bad news. An unbiased supervisor is referred to as neutral.

In this paper, we develop a simple model that accommodates all those types of environments *at the same time* and we seek to identify their impacts on quality provision. It depicts a continuum of agents who are willing to produce and sell a good whose quality can either be low, at no cost, or high, at a cost varying across agents. Each agent first chooses the level of quality, which is imperfectly observable by the potential buyer. Once the agent has selected its level of quality, some monitoring occurs, which discloses quality with a probability g if quality is high ("good news") and probability b if quality is low ("bad news"). After monitoring, the buyer updates its beliefs about quality and decides to purchase or not the good.

It is easy to see that differences between g and b are all but neutral. In particular, when the buyer receives no news after monitoring, belief updating is totally different: when the informational environment is friendly ($g > b$), the buyer knows that monitoring filters out high quality. Therefore receiving no news about a product leads him to become more pessimistic about its quality. When the informational environment is hostile ($g < b$), no news conversely improves its belief because he is aware that monitoring filters out low quality.[2]

The model describes how agents' quality investments interact with these belief revisions. We find clear-cut differences between hostile and friendly environments. Hostile environments give rise to a bandwagon effect among agents, which leads beliefs to be

[2] Our model has the minimal information structure to capture those aspects in a tractable way, but it is still richer than most standard models of disclosure or imperfect monitoring with endogenous incentives.

self-fulfilling in some cases, so that there may be multiple equilibria. The intuition is the following: when the prior is very optimistic, supplying high quality is an equilibrium because hostile monitoring would easily reveal the agents supplying low quality. But when the buyer's prior belief is pessimistic, incentives to increase quality are limited because it is relatively difficult to ascertain high quality, and there is few hope for the seller to prove wrong a pessimistic belief.

In contrast, friendly environments create a form of free riding across agents, which always induces a unique equilibrium. Comparative statics results are also contrasted. In particular, we show that moving from a hostile to a friendly environment increases the average equilibrium quality when the resources allocated to monitoring–reflected in the sum $g + b$ in our model–is low. The reverse is true for a higher $g + b$. The idea is that little information induces little reward to high quality firms and thus little incentive to produce high quality. Knowing that, the buyer is pessimistic about a good's quality on which he did not receive any feedback. It is more effective to go against the buyer's belief by increasing the probability to get the full reward when producing high quality. Also, reducing the cost of quality (in a precise sense which is defined in the paper) increases quality in any case, but less when the environment is friendly.

This paper is obviously far from being the first to explore the impact of informational environments on quality provision.[3] The distinct feature of our analysis is to build a model that highlights the key differences between hostile and friendly environments. In previous works, this dimension, albeit crucial as we point out, is either left implicit as in works on statistical discrimination (Coate and Loury , 1993) and collective reputation (Tirole , 1996); or the focus is on friendly environments as in the literature on quality disclosure which looks at the incentives for firms to voluntarily disclose quality and for certifiers[4] to provide unbiased certification about quality (for a recent survey, see Dranove

[3]Recent important works on the information structures of lemons problems include Sarath (1996), Levin (2001) and Kessler (2001). However they assume exogenous quality, hence the incentive dimension at the heart of our analysis is absent.

[4]There exists an important literature on information intermediaries, Lizzeri (1999) is among the seminal papers. Those analyses underline that certifiers somehow bias the informational environment towards friendliness. While we do abstract from strategic auditors here, it is an interesting connection to make in future works. The role of NGOs and firm monitoring is typically a setting of interest, in which the literature

and Jin, 2010). Importantly, the information structure we study is closely related to those in MacLeod (2007) and Board and Meyer-ter-Vehn (2010). In the former, two types of situations are studied: the case of a ″normal good″, for which breakdowns are relatively rare, and ″innovative goods″, for which breakthroughs constitute rare events. Those two situations pertain to the actual result of effort, not to the information structure. In the latter paper, the setting is dynamic and news arrive according to a a Poisson process. Either there is no information revealed, or, with Poisson arrival, the information structure ″ticks″. Whether a tick is good or bad news depends on whether high or low quality ticks more often. Hence the signal combination can be either no news/good news or no news/bad news, and the interpretation of a tick depends on the parameters.

Our work is also related to the model of Kamenica and Gentzkow (2011). They study the design of informational environment, a term we borrow from their paper. However in their analysis the state of the world is fixed, and unknown to all agents, while in our model quality is endogenous (and the seller is obviously aware of the quality he has chosen). They are interested in how the informational environment influences an ex-post action, while our focus is on how (ex-ante) moral hazard is influenced by the informational environment.[5]

This allows to derive new results and to provide new insights on existing results–such as the interpretation of multiple equilibria. In the penultimate section of the paper, we relate our results to different branches of the literature.

is still at an early stage (Lyon, 2010; Lyon and Maxwell, 2011). In such context, models are still needed that account for both friendliness and hostility in information revelation. Typically, environments in which ″bad cops″ NGOs such as Greenpeace are a majority generate a hostile informational environments.

[5]There is also a connection with inspection games (see Avenhaus et al., 2002, for an overview). In an inspection game, the inspectee chooses to comply (at some cost) or not with a previously agreed course of action, and the inspector can conduct an inspection (a statistical test) to verify compliance. The inspector strategy is hence related to the informational environment we model. There are many difference with our setting, though, the most crucial difference being that inspection games are simultaneous, while our setting is sequential. Relatedly, the literature on monitoring in principal-agent models (e.g. Dye, 1986; Fagard and Sinclair-Desgagné, 2007) studies the design of ex-post audits (of effort) that are contingent on some realized outcomes. In our case, there is always a one-dimensional signal, and further conditional audit can not be conducted. In addition, our assumptions on commitment are different from the ones made in those two strands of the literature.

The structure of the paper is straightforward. Section 2 presents the model. In sections 3 and 4, we characterize equilibria and conduct comparative statics exercises. In Section 5, we explore how our results relate to the existing literature. In particular, we apply our analysis to three research streams: statistical discrimination, collective reputation and quality certification. Section 6 gathers final comments.

2 The model

We consider a game with a continuum of agents (firms, sellers) that each produces one unit of a good which quality is imperfectly observable by a representative buyer (consumer, user). The quality variable a is binary, with $a \in \{0, 1\}$, and set by each agent. Choosing $a = 0$ costs the agent nothing , whereas choosing $a = 1$ costs c, which is heterogenous across agents, and follows a cumulative distribution $F(c)$ and density $f(c)$ on $[\underline{c}; \bar{c}]$. F and f are common knowledge, but each agent privately observes the realization of his cost.[6]

The buyer's (marginal) willingness to pay for the high quality ($a = 1$) is Δ and it is 0 for the low quality ($a = 0$). For the sake of clarity, most of the analysis is carried under the normalization $\underline{c} = 0$ and $\bar{c} = \Delta = 1$, which is qualitatively unimportant.[7] Finally, we also assume that the agents can fully extract the buyer surplus, which first gives the most incentives to the agents and second makes welfare analysis transparent. Under these assumptions, if quality were perfectly observable, the social optimum would be attained since all agents would choose $a = 1$ and sell the good at price Δ, since $\Delta \geq c$ for all c.

The buyer receives an imperfect signal $s \in \{H, L, \varnothing\}$ on the quality of the good supplied by each agent. The monitoring technology is asymmetric: the signal is generated according to the probabilities $g = \Pr[s = H|a = 1]$ and $b = \Pr[s = L|a = 0]$ where g can be less or higher than b. With probability $1 - g$ when $a = 1$ and with probability $1 - b$

[6]Equivalently, we model an agent with unknown cost and unobservable effort.

[7]We show in the appendix that the results go smoothly through when some agents have high costs, and never choose $a = 1$. The insights also carry through the case where some agents have negative costs. These agents could be interpreted either as intrinsically motivated by high quality production, or equivalently, as "honest" agents always producing the reference high quality. Finally, an important assumption that we also discuss regards the willingness to pay for low quality being 0.

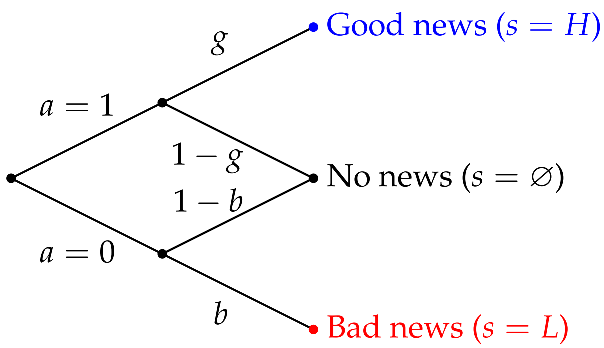

Figure 1: Information structure.

when $a = 0$, an empty signal $\varnothing$ is generated. We interpret this outcome as ”no news”, in the sense that it carries no evidence, but this signal may of course still lead to Bayesian interpretation. Overall, the information structure is thus one of partial hard information since signals H and L are evidence of the actual quality.[8]

The sequence of events, illustrated in figure 1, is as follows:

- **Stage 1:** The seller privately learns his cost c, then chooses a.
- **Stage 2:**
  - If the seller has chosen $a = 1$, the buyer learns this ”good news”, $s = H$, with probability g. With probability $(1 - g)$, no information is disclosed ($s = \varnothing$).
  - If the seller has chosen $a = 0$, the buyer learns this ”bad news”, $s = L$, with probability b. With probability $(1 - b)$, no information is disclosed ($s = \varnothing$).
- **Stage 3:** The buyer transfers Δ to the seller if $s = H$, but gives nothing for a bad news $s = L$. When no evidence is generated about the real action, the buyer forms a belief $\mu = \Pr[a = 1 | s = \varnothing]$. Given this belief, the seller only receives the conditional expected value $\mu\Delta$, which is the willingness to pay of the uninformed buyer.[9]

[8] We show in Appendix A how to extend our result to imperfect signals of quality, i.e. to a soft information structure.

[9] There are several extensive forms that corresponds to this payment scheme. A natural situation would be one in which the seller makes a take-it-or-leave-it price offer. The timing of this offer deserve some careful treatment. It could be made ex-ante, before the seller learns its costs and thus consists of a conditional price

Our main goal is to study the equilibrium distribution of qualities depending on the informational environment (g, b) and the cost distribution F.

The setting is sufficiently abstract to capture many real-world situations. The buyer can represent consumers in a final market with vertical differentiation in which endogenous quality is not perfectly observable (e.g., environmental attributes in a market with green consumers), a firm hiring employees whose intrinsic productivity resulting from past efforts is uncertain, lenders in a capital market, etc. The model can also apply to non-market situations such as school testing: A teacher needs to grade students whose level of performance is not always observed for monitoring resources are limited. In this context, Δ represents the grade for a student whose level of performance $a = 1$ is disclosed, 0 is the grade for a student whose observed performance is $a = 0$. Finally, $\mu\Delta$ is the grade of students for whom the teacher has no information.

3 Equilibrium analysis

3.1 Incentives and cutoff equilibria

We adopt the notion of perfect Bayesian equilibrium where each agent chooses its best reply to the market's belief and the buyer's belief is consistent with the quality set by the different types of agents.

Let $\Pi(a, c)$ be the expected payoff of an agent with cost c. For a given belief μ, the possible expected payoffs are: $\Pi(1, c) = [g\Delta + (1 - g)\mu]\Delta - c$ and $\Pi(0, c) = (1 - b)\mu\Delta$.

(on the signal to be realized), ex-post, once information is public, or at an interim stage, once the seller knows his type, but before the public signal is known. For the interim case, one could assume that the offer is made before or after quality is chosen. Ex-ante and ex-post offers clearly lead to the same outcome. In turn, interim offers are more subtle to analyze, as prices can be used as signalling devices (e.g. Milgrom and Roberts, 1986; Bagwell and Riordan, 1991). We take the view that the price offers should be renegotiation proof (or alternatively offers are made ex-post), having in mind that the seller could simply withdraw its product in case the offer is not profitable ex-post. Then equilibria are pooling and the extensive form does not matter. Another straightforward way to justify this payoff structure is that there are at least two identical buyers interested in the product.

The agent then chooses $a = 1$ whenever[10] $\Pi(1, c) \geq \Pi(0, c)$, which translates into:

$$c \leq g\Delta - (g - b)\mu\Delta \tag{1}$$

Note that the last RHS term $(g - b)\mu\Delta$ of this incentive constraint is negative if $g > b$ and positive if $g < b$. That is, if the environment is friendly, an increase in the belief μ reduces incentives to increase quality whereas it raises incentives in a hostile environment. We will show below that this asymmetry has fundamental consequences on equilibria.

An almost immediate consequence of the incentive constraint is that all the Bayesian equilibria of the game has a cutoff structure, namely they will all be characterized by a cost threshold below which agents choose quality 1 and above which the others choose quality 0. A first lemma states this formally.

Lemma 1 *All Bayesian equilibria of the game described above are cutoff equilibria. Namely, any equilibrium is characterized by some c^* such that all firms with $c \leq c^*$ choose $a = 1$ and all firms with $c > c^*$ choose $a = 0$. When $0 < c^* < 1$, this cutoff satisfies:*

$$c^* = g\Delta - (g - b)\mu^*\Delta. \tag{2}$$

The corresponding equilibrium belief μ^* is consistent with the cutoff c^* according to Bayesian updating:

$$\mu^* = \frac{(1 - g)F(c^*)}{(1 - g)F(c^*) + (1 - b)(1 - F(c^*))} \tag{3}$$

Proof. Consider an equilibrium of the game in which the beliefs of the stakeholder upon receiving no news is some μ^*, and suppose that there exists a $\hat{c}$ such that $\hat{c} \leq g\Delta - (b - g)\mu^*\Delta$. Then for all $c \leq \hat{c}$ we have $c \leq \hat{c} \leq g\Delta - (b - g)\mu^*\Delta$, hence the best reply to the equilibrium belief μ^* of all types below $\hat{c}$ is to choose $a = 1$. Similarly, if some type chooses $a = 0$, then all types above choose not to self-regulate. This establishes that an equilibrium is characterized by a cutoff c^*, possibly the extreme types 0 or 1. Bayesian revision obtains from the fact that the fraction of self-regulating firms is $F(c^*)$, the mass of types lower than the cutoff. ■

[10] Mixing by one type is here unimportant given it has zero weight, and we assume that the unique indifferent type chooses $a = 1$ over $a = 0$.

Given that an equilibrium is entirely characterized by a cost cutoff c^*, we will often refer directly to c^* as an equilibrium, the corresponding belief being unequivocally obtained using (3).

3.2 Existence and stability

To investigate existence and stability properties of the equilibria we first use a standard fixed point representation of the problem. We also impose $\Delta = 1$ to simplify notations. Combining (3) and (2) yields:

$$c^* = \Phi(c^*) \quad \text{with} \quad \Phi(c) \equiv \frac{g(1-b) - (g-b)F(c)}{(1-b) - (g-b)F(c)} \tag{4}$$

The function Φ is continuous with $\Phi(0) = g$ and $\Phi(1) = b$. Hence existence of equilibrium is ensured by the intermediate value theorem. It is also straightforward that the equilibrium is unique when $g \geq b$ as Φ is then a (global) contraction mapping of c:

$$\frac{d\Phi}{dc} < 1 \Leftrightarrow -(g-b)f(c)\left(\frac{(1-g)(1-b)}{[(1-b)-(g-b)F(c)]^2}\right) < 1$$

In contrast, there can exist several equilibria in hostile environments ($g < b$) as the graph of Φ possibly cross several times the 45 degree line. This ultimately depends on the properties of F.

To explore further the case with multiple equilibria, we introduce the following notion of stability:

Definition 1 *An equilibrium c^* is stable if $\frac{d\Phi}{dc}(c^*) < 1$. An equilibrium is unstable or is a tipping point if $\frac{d\Phi}{dc}(c^*) \geq 1$.*

This is a standard definition[11] which amounts to say that a stable c^* is an attractive fixed point. Underlying this notion is the idea of dynamic adjustment: In the case where small perturbations induce a little deviation of the equilibrium from c^* to $c^* + \varepsilon$, the cost cutoff will return to c^* if the equilibrium is stable. This occurs if the iterated function

[11] A closely-related version of stability is considered in Jackson and Yariv (2007), which does not require Φ to be continuously differentiable. In our framework, both turn out to be equivalent. See also Brock and Durlauf (2001).

sequence $c^* + \varepsilon$, $\Phi(c^* + \varepsilon)$, $\Phi(\Phi(c^* + \varepsilon))$,... converges to c^*. Alternatively, the equilibrium is unstable if small perturbations lead to a diverging dynamic adjustment towards a nearby stable equilibrium. c^* is then said to be a tipping point since a downward deviation $c^* - \varepsilon$ and an upward deviation $c^* + \varepsilon$ lead to two different stable equilibria. Note that, as $\Phi(0) = g < \Phi(1) = b$ in hostile environments, $\frac{d\Phi}{dc}(c^*) < 1$ for the lowest and the highest equilibrium.

3.3 Equilibrium characterization

For the remainder of the exposition, an alternative representation of equilibria will prove most useful. When $g \neq b$, we suggest rewriting Equation (4) as follows:

$$F(c^*) = M(c^*|g,b) \quad \text{with} \quad M(c|g,b) \equiv \frac{(g-c)(1-b)}{(g-b)(1-c)} \tag{5}$$

The function M will be referred to as the *monitoring line*, as it contains the data pertaining to the informational environment, and is independent of the distribution F. We derive some properties that will be used in the subsequent analysis:

Lemma 2 *Properties of the monitoring line:*

1. *M is is decreasing and concave in a friendly environment, and increasing and convex in a hostile environment. It is a vertical line (i.e. not a function) in a neutral environment (when $g = b$).*

2. *In the case where $g < b$, an equilibrium is stable (resp. a tipping point) if $f(c^*) < \frac{\partial M}{\partial c}$ (resp. $f(c^*) \geq \frac{\partial M}{\partial c}$).*

Proof. For the first part of the lemma, note that $\frac{\partial M}{\partial c} = -\frac{(1-b)(1-g)}{(g-b)(1-c)^2}$ and $\frac{\partial^2 M}{\partial c^2} = -\frac{2(1-b)(1-g)}{(g-b)(1-c)^3}$. Both derivatives are negative (resp. positive) when $g < b$ (resp. $g > b$). The second part follows directly from the stability condition $\frac{\partial \Phi}{\partial c} < 1$ and $\frac{\partial \Phi}{\partial c} = f(c^*)\left(\frac{\partial M}{\partial c}\right)^{-1}$.

■

Using this representation, an equilibrium cutoff is such that the graph of the cumulative distribution intersects the auditing line, as pictured on figure 1[12]. There might be

[12] In the symmetric case $g = b$, M actually consist of a vertical line at $c^* = g = b$.

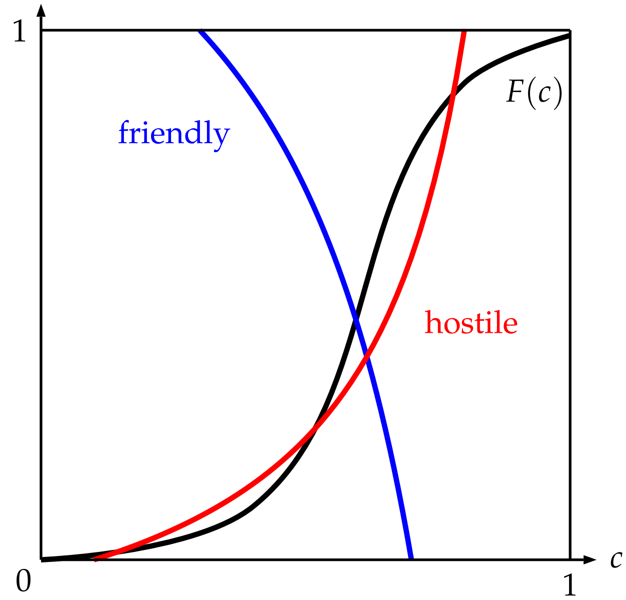

Figure 2: Equilibrium representation in the $(c, F(c))$-space.

multiple intersections between F and M when $g < b$. The lemma says that stable equilibria correspond to crossings of M by F from above.

We summarize the analysis of this section in a first proposition:

Proposition 1 *Characterization of equilibria.*

1. *In a neutral environment ($g = b$), there exists a unique equilibrium with $c^* = g = b$.*

2. *In a friendly environment ($g > b$), there exists a unique perfect Bayesian equilibrium defined by (4). Moreover the cutoff is such that $b \leq c^* \leq g$.*

3. *In a hostile environment ($g < b$), there might be multiple equilibria defined by (4), and there are all such that $g \leq c^* \leq b$. Moreover the lowest and the highest equilibrium are stable.*

The first part of the proposition corresponds to the traditional symmetric monitoring technology ($g = b$). In this case, the equilibrium cutoff $c^* = g = b$ does not depend on the distribution F. Still, the fraction of agents choosing high quality is of course $F(g)$. But there is no strategic effect between the different types of agents.

The major result lies in the existence of multiple equilibria in hostile environments (and only in hostile environments). To understand what drives this result, note first that the equilibrium belief (3) increases with the number of agents opting for high quality, $F(c^*)$. Hence, quality choice generates a positive externality as equilibrium payoffs of all agents increase with the belief, whether they choose $a = 1$ or not.

This externality has different impacts on quality choice in the two informational environments. To explain why, one needs to turn back to the RHS of the incentive constraint (1). In particular, the last term $(g - b)\mu\Delta$ is negative if $g > b$ and positive if $g < b$. This means that:

- if the environment is friendly, the externality passing through the belief reduces incentives to increase quality in a friendly environment (a free riding effect);
- if the environment is hostile, it provides more incentives (a bandwagon effect).

Loosely speaking, quality choices of different types of agents are strategic substitute in a friendly environment. They are strategic complement in a hostile environment. As a result, the beliefs tend to be self-fulfilling in the latter case. Consider for instance the extreme case where $g = 0$ and $b = 1$. If the belief is $\mu = 0$, the buyer basically thinks the agents will never choose high quality, and this belief can never be proven wrong because $g = 0$. Agents have thus no interest in raising quality since they will never get any premium for that. We get just the opposite result if $\mu = 1$: every agent has no choice but to provide high quality as low quality would always be revealed (since $b = 1$).

For the sake of applications, our equilibrium characterization has two corollaries pertaining to equilibrium multiplicity, the first one regarding the information structure and the second one regarding the cost distribution.

Corollary 1 *A necessary condition for multiple equilibria is that the informational environment is hostile.*

This is typically important for models of statistical discrimination, that are discussed in the application section. It is also linked to a discussion on multiple equilibria in compliance model. Compliance by definition creates a hostile environment. However, while hostility is necessary for multiplicity in our framework, it is not sufficient. In particular, we provide a condition for uniqueness on the distribution function:

Corollary 2 *A sufficient condition for the equilibrium to be unique irrespective of the information structure is that the cost distribution is concave and $g > 0$. When $g = 0$, $c^* = 0$ is also an equilibrium, but it is unstable.*

Hence in particular if the distribution of costs is uniform the equilibrium is unique.

4 Comparative Statics

We now investigate how changes in informational environment or in cost distribution influence equilibria.

4.1 Changes in informational environment

How do variations in g and b affect the equilibrium? It is easy to obtain the following comparative statics:

Proposition 2 *All else equal, the equilibrium cutoff (and hence the fraction of agents opting for high quality),*

1. *increases with the quality of information when the equilibrium is stable:* $\frac{\partial c^*}{\partial g}, \frac{\partial c^*}{\partial b} \geq 0$.
2. *decreases when the equilibrium is a tipping point:* $\frac{\partial c^*}{\partial g}, \frac{\partial c^*}{\partial b} < 0$.

Proof. Differentiating the equilibrium condition (5) with respect to g and b one obtains:

$$\frac{\partial c^*}{\partial g} = \frac{\frac{\partial M}{\partial g}}{f(c^*) - \frac{\partial M}{\partial c}} \text{ and } \frac{\partial c^*}{\partial b} = \frac{\frac{\partial M}{\partial b}}{f(c^*) - \frac{\partial M}{\partial c}}$$

In friendly environments, we know that $f(c^*) - \frac{\partial M}{\partial c} \geq 0$ as $\frac{\partial M}{\partial c} = -\frac{(1-b)(1-g)}{(g-b)(1-c)^2} < 0$. Then we differentiate M with respect to g and b : $\frac{\partial M}{\partial g} = \frac{(1-b)(c-b)}{(1-c)(g-b)^2}$ and $\frac{\partial M}{\partial b} = \frac{(1-g)(g-c)}{(1-c)(g-b)^2}$. $\frac{\partial M}{\partial g}, \frac{\partial M}{\partial b} \geq 0$ follows from $b \leq c^* \leq g$ (Proposition 1). In hostile environments, results from $\frac{\partial M}{\partial g}, \frac{\partial M}{\partial b} > 0$ (because $g < c^* \leq b$) and $f(c^*) < \frac{\partial M}{\partial c}$ (resp. $f(c^*) \geq \frac{\partial M}{\partial c}$) when the equilibrium is stable (resp. unstable). ■

The first part of the proposition is very intuitive: better information creates higher incentives to supply high quality. In contrast, the second could seem awkward at first glance. But but it is not if one adopts a dynamic view of the equilibrium. As explained above, a tipping point is a level of c at which any small upward deviation leads to a higher stable equilibrium, and any downward deviation leads to a lower equilibrium. Hence, the lower the tipping point, the higher the chance to shift to a higher equilibrium. The overall message remains thus the same: better information improves quality.

As the focus of the paper is the contrast between hostile and friendly informational environments, it sounds appropriate to see how the relative weight of friendly and hostile monitoring influences the equilibrium. A straightforward strategy for looking at this issue consists in solving the following optimisation program:

$$\max_{g,b} \quad c^*(g,b) \quad \text{subject to} \quad g+b=q \tag{6}$$

where q represents the overall quantity of information which is less than 1 because monitoring resources are limited for instance.[13] One example of interpretation would be the following: there is a limited number of independent imperfect tests (or questions), measured by q, that can be performed on quality. Each test has equal power, and can identify either low quality or high quality. The problem is then to choose the balance between the different type of tests to be run. Another example would be a committee with a given size q. Each committee member is imperfectly informed on some specific elementary aspect of quality, but some members have preferences in favor of the agent (they are friendly), other have preferences against the agent (they are hostile). Then the probabilities g and b depend linearly on the composition of the committee, known in advance by the agent. We establish a new proposition for such scenario:

Proposition 3 *Under the constraint $g+b=q$, c^* is maximized either with $g=0$ or $g=q$. More specifically,*

[13] Another modeling route would be to assume explicit monitoring costs. At any rate, it is not a priori clear what shape such a cost function should take.

1. *If the equilibrium is unique when $g = 0$, the maximum is attained for $g = 0$ if $F(q/(2-q)) > 1/2$, and $g = q$, otherwise.*

2. *If there exist multiple stable equilibria when $g = 0$, then:*

   (a) *if $F(q/(2-q)) > 1/2$, there exists at least one equilibrium with $g = 0$ which dominates the equilibrium with $g = q$.*

   (b) *if $F(q/(2-q)) \leq 1/2$, there exists at least one equilibrium with $g = 0$ which is dominated by the equilibrium with $g = q$.*

Proof. See the Appendix. ■

The overall message is clear: friendliness becomes preferable when the amount of information q decreases and/or when the weight of low-cost agents becomes less important (F is lower at $q/(2-q)$). The intuition is as follows: a low q entails that little information is disclosed about firms' activities. As a result, little reward is granted to high quality firms and incentives for quality are limited. Knowing that, the Bayesian buyer holds a pessimistic belief that one agent about whom he does not receive any feedback is producing high quality. Similarly he is pessimistic when F is low as it implies high quality is likely to be too costly for the agent.

In this context, Proposition 3 tells us it is more effective to go against the buyer's belief. That is, to motivate agents to provide high quality by increasing the probability to get the full reward when producing high quality. And conversely when the buyer is optimistic about the level of high quality supply he faces (because q is high and/or F is low).

4.2 Change in costs

In this section we study more generally the consequences of changes in costs. More specifically, we consider a decrease in the sense of First-Order Stochastic Dominance (FOSD): a distribution F first-order stochastically dominates a distribution G when $F(c) \leq G(c)$ for all c.

We establish a new proposition.

Proposition 4 *Consider a decrease in costs in the sense of FOSD:*

1. *The fraction of high quality supplied in equilibrium increases in any stable equilibria.*

However,

2. *The equilibrium cutoff decreases in a friendly environment.*

3. *The equilibrium cutoffs of stable equilibria increase in a hostile environment. Except for the higher equilibrium, some stable equilibria can even disappear after FOSD shifts of F.*

Hence a decrease in costs makes hostile environments relatively more efficient at inducing incentives than friendly environments.

Proof. Consider a family of distributions $F(c;\theta)$ with $\frac{\partial F}{\partial \theta} > 0$, so that, if $\theta' > \theta$, $F(.;\theta)$ first-order stochastically dominates $F(.,\theta')$. From (5) follows:

$$\frac{\partial c^*}{\partial \theta} = -\frac{\frac{\partial F}{\partial \theta}}{f(c^*) - \frac{\partial M}{\partial c}} \tag{7}$$

Hence $\frac{\partial c^*}{\partial \theta} > 0$ if $g > b$ as we know that $f(c^*) - \frac{\partial M}{\partial c} > 0$ in this case. $\frac{\partial c^*}{\partial \theta} < 0$ if $g < b$ follows from the fact that $f(c^*) - \frac{\partial M}{\partial c} < 0$ for stable equilibria in hostile environments.

Then we have

$$\frac{dF(c^*(\theta);\theta)}{d\theta} = f(c^*)\frac{\partial c^*}{\partial \theta} + \frac{\partial F}{\partial \theta}$$

Plugging (7) in this expression leads to

$$\frac{dF(c^*(\theta);\theta)}{d\theta} = -\frac{\frac{\partial M}{\partial c}\frac{\partial F}{\partial \theta}}{f(c^*) - \frac{\partial M}{\partial c}}$$

which is positive as $f(c^*) - \frac{\partial M}{\partial c}$ and $\frac{\partial M}{\partial c}$ have opposite signs in both environments for stable equilibria. ■

Figure 3 provides a representation of these results: a decrease in costs in the sense of first-order stochastic dominance is just a shift of the cost distribution from F to G. Figure 4 illustrates this point in a special case where the two environments initially performs equally well with F. When costs decrease, i.e. when the distribution becomes G, with $G > F$, the hostile environment induces more self-regulation.

The fact that decreasing the cost of quality raises high quality supply is not surprising. But why is the impact less important in friendly environments? Remember the contrast

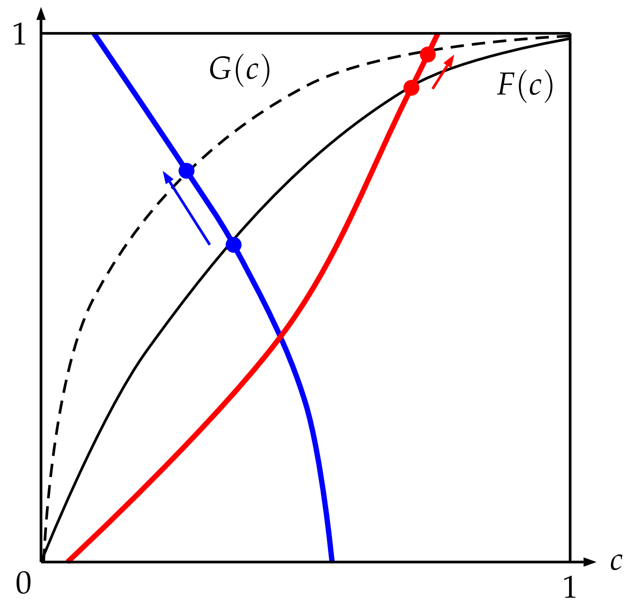

Figure 3: Decrease in costs (FOSD).

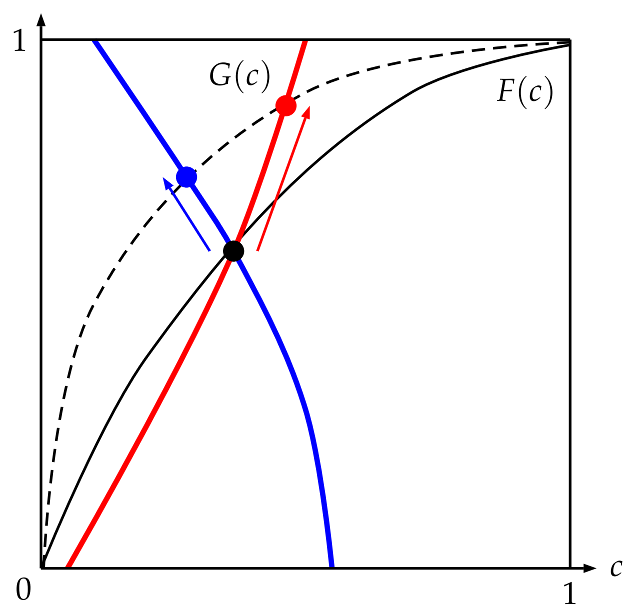

Figure 4: Comparative advantages.

we pointed out previously between the two environments. In friendly environments, rising the buyer's belief has a negative impact on one agent's incentive to choose high quality, resulting in a free riding effect. This attenuates the positive effect of increased supply of high quality on incentives. In hostile environments, a self reinforcing mechanism turns agent quality choices into strategic complements, which exacerbates the positive effect of increased supply of high quality.

5 Applications and relation to the literature

5.1 Statistical discrimination

Models of statistical discrimination are used to explain group inequality. In the literature originated from the seminal contribution of Arrow (1973), average group differences are endogenously derived in equilibrium without assuming any ex ante exogenous differences between groups.[14]. A prominent example is the model by Coate and Loury (1993) which describes the interaction between employers and groups of workers whose individual productivity is imperfectly observed. Discrimination then amounts to the existence of multiple equilibria, which implies that ex ante identical groups could exhibit different average level of qualifications ex post (and thus different wages). Our model sheds a new light on the multiple equilibria issue by giving sensible conditions for the equilibrium to be unique, and relates the multiplicity problem to the informational environment.

To explain this, we now put our analysis in a discrimination context. The agents are potential workers which belong to a given sexual or racial group. Individual agent's level of qualification can either be low ($a = 0$) or high ($a = 1$) and is endogenously chosen by making an investment c in human capital for $a = 0$ and 0 for $a = 0$. The buyer is an employer who intends to hire workers. If the worker is qualified, the employer gets a return Δ and 0 if it is not. The problem is that the employer does not observe a when making hiring decisions. But it can rely on a test which reveals the level of qualification

[14]In contrast to Phelps (1972), where ability is on average different in different groups.

with a probability g if the agent is qualified and b if it is not.[15] As we continue to assume that the agent extracts the total expected employer's return[16], the employer offers the wage Δ when the test reveals the agent is qualified, 0 when $a = 0$ is disclosed and $\mu\Delta$ when the test is not conclusive where μ is the probability assigned by the employer that the tested agent is actually qualified. Under these assumptions, Proposition 1 replicates the standard result of the statistical discrimination literature that there might multiple equilibria.

The literature on statistical discrimination does not convey any clear messages over the conditions under which discrimination occurs. Proposition 1 in Coate and Loury (1993, p. 1126) does give a necessary and sufficient condition for multiple equilibria, but the condition is not interpretable (and they do not actually try to interpret it). In contrast, our model invites to focus on the role of technology used to evaluate worker productivity in the labor market and we show that multiplicity can only occur in hostile environments. This may generate novel policy implications on the design of evaluation tests and procedures. They need to be more friendly in the sense we give to friendliness in this paper. More specifically, our results suggest that in hostile environments with discrimination:

- increasing g is more efficient than reducing b for it increases the average productivity (Proposition 2) and can eliminate discrimination.
- If resources is limited so that it is not possible to increase g without reducing b, reducing discrimination is in line with increasing average qualification when $g + b$ is low as suggested by Proposition 3. That is, when evaluation is difficult for intrinsic reasons or because resources are limited.

Our framework also gives insights on the effectiveness of other policy approaches. For instance, subsidizing education investments of discriminated groups (decreasing c in our model) is less effective when evaluation hostility is not too severe (Proposition 4). At this

[15] The literature uses a different information structures with soft information.

[16] Although not realistic in a labor market context, this assumption does not influence qualitatively the results.

stage, these points are not thoroughly established but they give interesting new directions of research.

5.2 Collective reputation

An interesting application of our model consists in reinterpreting it as a collective reputation model.[17] Firms with different types (costs) are pooled into a single group. While sometimes the results of quality investment by a given firm turns out to be individually observable (with probability g or b, depending on whether the investment is high or low), it can escape pooling with other firms for which no signal is available. The collective reputation effect hence arises when no information is revealed, in which case all firms obtaining the null signal are treated equally, i.e. under the same pricing umbrella, and receive in equilibrium a price of $\mu^*\Delta$.

The seminal contribution in this area has been made by Tirole (1996) who describes a principal who contracts with an agent only if is sufficiently confident that the agent will not engage in corrupt activities. This a model in which, at each date t, the principal is matched with a new agent. There is imperfect information about the agent's past behavior: with a probability b, the principal knows that the agent has been found corrupt at least once. In addition the probability b is higher when the agent has cheated more in the past. As a result, the agent trade-off the current benefit of corruption and the loss in reputation. The model then looks at the steady states of the model. Tirole (1996, proposition 1, p. 9) shows the existence of multiple steady states stemming from the dynamic complementarity between past and future reputations, implying that corruption may persist or, conversely, may be maintained at a low level. Our analysis shows that these results rests on the hostility of the informational disclosure mechanism: the principal can only learn something when the agent has been corrupt. Hostility implicitly originates here from the fact that the quality variable is conformity with a rule. Being found non compliant initiates a judicial process, which selectively discloses information to the general public on

[17] One could also think of it as a one player reputation problem. However some interpretations would require , especially as regards stability and externality across agents' type. See Bar-Isaac and Tadelis (2008) for a recent survey on reputation, framed precisely in a buyer/seller setup.

low quality.[18]

Note that, as Levin (2009) points out, the collective reputation model can be interpreted as a dynamic version of standard statistical discrimination models. Blume (2005, 2006) provides a bridge between the two approaches by considering the dynamic version of a canonical statistical discrimination model.

5.3 Quality disclosure and certification

The model can also provide novel insights on quality certification (for a survey of this literature, see Dranove and Jin, 2010). Consider for instance the case of a product label which selectively signals high quality. The certifier decides to grant the label on the basis of a private evaluation, obtained with a given inspection ability. More precisely, assume that the certifier owns an imperfect but symmetric monitoring technology: with probability $m < 1$, it is fully informative on a, while with probability $1 - m$ it generates an empty signal. The monitoring technology is thus a priori neutral: it either uncovers the true quality or discloses nothing. A first possible labeling policy is *strict*. Under the strict policy, the certifier grants the label only if $a = 1$. The other possible rule is a *lenient* policy, which consists in labeling products only if the certifier does not observe $a = 0$. Suppose finally that firms first choose the quality of their product, then apply for certification, and that consumers observe only whether a product is certified or not. That is, labeling is a public coarsening of the (private) information of the certifier, as shown on Figure 5 We also assume for simplicity that the certifier is truthful. Under the strict policy, consumers will be sure that the quality of a labeled product is high. Under the lenient policy, consumers remain uncertain about labeled products quality,[19]. In turn, a lenient policy informs fully on non-labeled products, which are necessarily of low quality.

Our model predicts very different equilibrium qualities under the two labels:

- When the certifier uses a strict rule, the agent's expected profit opting for high qual-

[18] Obviously, the process can also exonerate the defendant, but this outcome is arguably less frequent, and not accounted for in Tirole (1996).

[19] Harbaugh et al. (2011) deal with a related problem in which the it is the quality of monitoring inside the label that is uncertain. We consider here that the potential consumers know whether the label has a strict or a lenient policy.

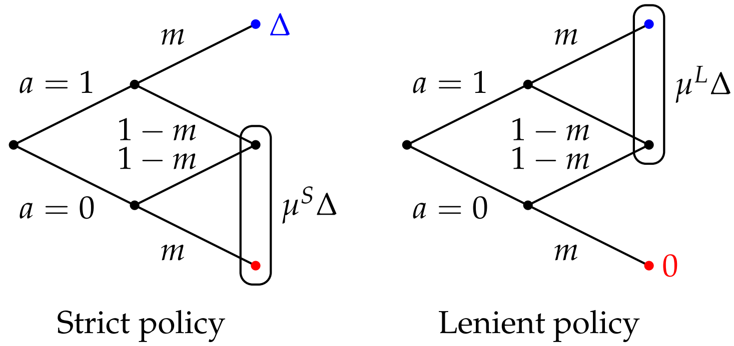

Figure 5: Labeling policies.

ity is $\Pi(1, c) = [m + (1-m)\mu^S]\Delta - c$ and $\Pi(0, c) = \mu^S\Delta$. Hence, the incentive constraint is $c \leq m(1-\mu^S)\Delta$. Hence this it corresponds to a purely friendly environment with $g = m$ and $b = 0$.

- When the certifier uses a lenient rule, we have $\Pi(1, c) = \Delta - c$ and $\Pi(0, c) = (1 - m)\mu^L\Delta$, implying the following incentive constraint $c \leq m\mu^L\Delta$. This corresponds to a purely hostile environment with $g = 0$ and $b = m$.

Strict labels signal only high quality, and are thus associated with higher product prices. In turn, non-labeled products when the label is strict are of mixed quality, and their equilibrium price is intermediate. Lenient labels, on the other hand, filter out bad products ones, hence they allow the consumers to identify perfectly poor quality–hence non-labeled product trade at the lowest price. Indeed, prices for both labeled *and* non-labeled products are higher when the label is strict. What matters for overall incentives is the gap between the price of a labeled product and the price of a non-labeled product, and which gap is bigger depends on the quality of information m and the cost distribution. A first immediate consequence of Proposition 1 is that with lenient labels, there may be multiple equilibria. The label can reach recognition and create strong incentives, or fall flat, and create only mild differences between labeled and non-labeled products.

Assume now that the cost distribution is concave, so that there is a equilibrium stable equilibrium under both policies (Corollary 2). Proposition 3 suggests the following policy implications: strict labels are socially preferable when evaluation is difficult (low m) and/or when the cost of high quality tends to be relatively high. Our model relates this

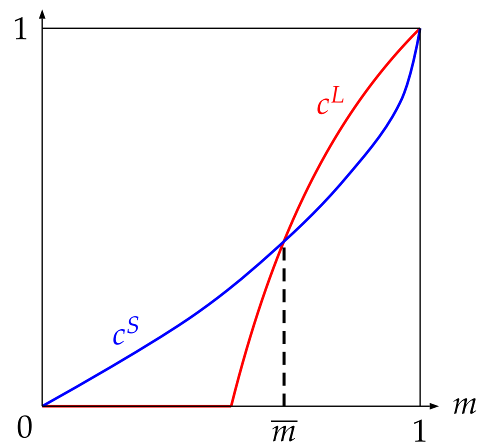

Figure 6: Equilibrium cutoffs under the two policies (uniform distribution).

to the friendliness or hostility that labeling does generate. Figure 6 illustrates this in the case of the uniform distribution.

6 Conclusion

The model developed in this paper tackles an issue which has been thoroughly explored in the literature: the impact on sellers' incentives of asymmetric information on the quality they supply. We provide an original dichotomy of informational environments, between friendly environments, in which high quality is more easily identified by buyers than low quality, and hostile environments, in which this is the opposite.

We show that whether the informational environment is friendly or hostile makes a considerable difference. In particular, hostility gives rise to a bandwagon effect across agents which can generate in certain circumstances multiple equilibria, which can not occur in friendly environments. The generic insight we uncover is that it is more effective to rely on the monitoring technology which can lead to the most substantial revisions of the buyers' equilibrium prior. For instance, hostility provides higher incentives for quality than friendliness when the cost of quality is low (which implies that buyers are optimistic over the average level of quality). Similarly, friendliness is more effective in increasing quality when information is poor, as this means little incentives to provide quality and thus pessimistic beliefs.

A Appendix: Soft Information

The basic model features "hard" information, or "evidence", when information is disclosed. Arguably, this could be seen as a limitation for two reasons. First, assuming that consumers are sometimes perfectly certain that quality is good (or bad) is quite strong: signals are often not perfect while still being informative. Second, one might precisely like to talk about signal precision in our context. We show here that the model easily generalizes to "soft" information and preserves our qualitative results, provided the signal structure is sufficiently precise (hence close–in some sense–to the hard information structure).

Consider an information structure with a ternary imperfect signal $s \in \{L, M, H\}$. Let $p(s|a)$ be the probability of receiving signal s when the true action is a. Soft information means here that both $a = 0$ and $a = 1$ can lead to any signal: $p(s|a) > 0, \quad \forall s$ (i.e. there is full support). To fix ideas, the hard information structure can be written as: $p(H|1) = g$, $p(M|1) = 1 - g$ and $p(L|1) = 0$, while $p(H|0) = 0$, $p(M|0) = 1 - b$ and $p(L|0) = b$. Then H is evidence that $a = 1$ and L evidence that $a = 0$. The intermediate signal M is consequently exactly our previous "no news" outcome, which could originate in either action.

The generic structure has four free parameters, and we will use a parameterization that allows a clear connection with the hard information structure. Let $p(H|1) = g\gamma$, $p(M|1) = 1 - g$ and $p(L|1) = g(1 - \gamma)$, while $p(H|0) = b(1 - \beta)$, $p(M|1) = 1 - b$ and $p(L|0) = b\beta$. Hence we keep the probability of receiving the signal interpreted as "no news", M, to be g or b as before. The additional parameters γ and β pertain to signal precision. In order to keep the interpretation of "good" and "bad" news, the following assumption is needed (this is without loss of generality, it is only a matter of labeling):

$$\frac{p(H|0)}{p(L|0)} \leq \frac{p(H|1)}{p(L|1)} \Leftrightarrow \gamma + \beta \geq 1$$

This corresponds to the correct notion of good and bad news in this context, as shown by Milgrom (1981). It just says that when the signal is H, the likelihood of $a = 1$ is higher than that of $a = 0$ (and conversely for signal L). Note that this condition is in particular automatically satisfied under the stronger requirement that γ and β are both higher than

1/2, saying that the signal H is more often received than the signal L when $a = 1$, and conversely when $a = 0$.

Finally, we need to introduce the notion that the signal M can be interpreted as the no news signal, to finish the parallel with the hard information model. This corresponds to the idea of a signal being "in between", less conclusive than h or L. This is again captured by the information order proposed by Milgrom (1981), in the following sense:

$$\frac{p(L|1)}{p(L|0)} \leq \frac{p(M|1)}{p(M|0)} \leq \frac{p(H|1)}{p(H|0)},$$

which we rewrite as the following condition:

$$\frac{1-\beta}{\beta} \leq \frac{(1-g)b}{g(1-b)} \leq \frac{\gamma}{1-\gamma} \tag{8}$$

Intuitively, this series of inequalities should be satisfied for information structures that are sufficiently close to hard information. This is the case in the following sense:

Lemma 3 *For any $(g, b) \in (0, 1)^2$, the inequalities (8) are satisfied provided the information structure is precise enough, in the sense of γ and β high enough.*

Proof. Consider first the right inequality in isolation. It is equivalent to $\gamma \geq \frac{1}{1+\frac{(1-g)b}{g(1-b)}}$, which is always strictly smaller than 1 for $b \in (0, 1)$. Hence a high precision conditional on $a = 1$ makes the inequality satisfied. Similarly, one obtains for the left inequality the condition $\gamma \geq \frac{1}{1+\frac{(1-b)g}{b(1-g)}}$. Note in particular that each precision has to be higher than 1/2. ∎

Now, assume that condition (8) holds, and denote by $\mu(s)$ the belief of the consumer upon receiving signal s, which determine the price to be paid. Assuming for simplicity $\Delta = 1$, recall that the incentive constraint is:

$$p(H|1)\mu(H) + p(M|1)\mu(M) + p(L|1)\mu(L) - c \geq p(H|0)\mu(H) + p(M|0)\mu(M) + p(L|0)\mu(L),$$

which can be rewritten as:

$$c \leq (g\gamma + b\beta - b)\mu(H) - (g-b)\mu(M) + (g - g\gamma - b\beta)\mu(L) \tag{9}$$

The crucial point is that the coefficient of the posterior $\mu(M)$, which is the equivalent of μ in the hard information model, is precisely the same, $-(g-b)$. Under the mild additional assumption that $g\gamma + b\beta \geq \max\{g, b\}$, which again is just a precision assumption, the incentive constraint is hence qualitatively the same as with hard information.

Concluding that the analysis would be exactly the same under soft and hard information is however erroneous. While under soft information there exist equilibria qualitatively similar to the ones we have studied, the set of equilibria is larger under soft information. In particular, the zero-quality equilibrium is always an equilibrium of the soft information game–precisely like a "babbling" equilibrium in a cheap talk framework. Indeed an equilibrium beliefs that no firms choose $a = 1$ can not be contradicted, and is self-fulfilling. No evidence can contradict this belief, and any message can then legitimately be interpreted as coming from an action $a = 0$–just as any message is interpreted as noise in a babbling cheap talk equilibrium.

B Appendix: motivated agents, inefficient agents and hazardous products

This appendix considers three extensions of the basic model: one in which some agents are motivated and always provide high quality (i.e. a fraction of agents has negative costs), another one in which some agents have costs higher than the willingness to pay, and finally on in which the bad quality product is dangerous, i.e. yields a negative utility to the buyer. In this last case, the participation constraint of the buyer should be taken into account: it can be better not buying rather than buying a too risky product.

C Appendix: Omitted Proofs

C.1 Proof of Proposition 3

We first compare the two corner cases: $g = 0$ and $g = q$. To simplify notations, we write $c_0^* \equiv c^*(0, q)$ and $c_q^* \equiv c^*(q, 0)$. Plugging in (5) yields:

$$F(c_0^*) = M_0(c_0^*) \quad \text{with} \quad M_0(c) \equiv \frac{c(1-q)}{q(1-c)}$$
$$F(c_q^*) = M_q(c_q^*) \quad \text{with} \quad M_q(c) \equiv \frac{(q-c)}{q(1-c)}$$

Then, we adopt a geometrical approach to compare c_0^* and c_q^*. Straightforward calculations yield properties which we use to draw the functions M_0 and M_q in Figure X:

- $M_0' > 0$, $M_0(0) = 0$, $M_0 \to +\infty$ if $c \to 1$ and $M_0(c) = 1$ if $c = q$.
- $M_q' < 0$, $M_q(0) = 1$, $M_q \to -\infty$ if $c \to 1$ and $M_q(c) = 0$ if $c = q$.
- $M_0 = M_q$ if $c = q/(2-q)$ and $M_0\left(q/(2-q)\right) = M_q\left(q/(2-q)\right) = 1/2$.

The graph shows that $F(q/(2-q)) > 1/2$ is sufficient for having at least one stable equilibrium c_0^* which dominates c_q^*. The reason is simple: If $F(q/(2-q)) > 1/2$, we necessarily have $c_q^* < q/(2-q)$ whereas F crosses M_q from above beyond the intersection between M_0 and M_q in $c = q/(2-q)$.

Things are less clear when $F(q/(2-q)) \leq 1/2$. To begin, we know that $c_q^* \geq q/(2-q)$. Then we know that, if there exist a unique equilibrium c_0^*, it is less than $q/(2-q)$[20]. Hence $c_q^* \geq c_0^*$. If there exist multiple equilibria c_0^*, it remains that there exists at least one equilibrium c_0^* such that $c_0^* \leq c_q^*$. But there also might be higher equilibria with certain distribution functions.[21]

Finally, we show that no interior solution dominates the corner solution. The program (6) is equivalent to $\max_g \; c^*(g, q-g)$. Let g_i denote a candidate interior solution. g_i should

[20]Note that this equilibrium is $c_0^* = 0$ if $f - \partial M_0/\partial c < 0$, which is equivalent to $f(0) < (1-q)/q$.

[21]This is however never the case when F is convex beyond $c = q/(2-q)$ because F never crosses M_0 above $q/(2-q)$ in this case. This happens for instance if F is globally convex or if F is unimodal with a mode less than $q/(2-q)$.

satisfy

$$\frac{dc^*}{dg}(g, q-g) \quad = \quad \frac{\partial c^*(g, q-g)}{\partial g} - \frac{\partial c^*(g, q-g)}{\partial b} = 0$$

$$\iff \quad \frac{1}{f - \frac{\partial M}{\partial c}} \frac{(c^* - q + g)(1 - q + g) - (g - c^*)(1 - g)}{(q - 2g)^2 (1 - c^*)} = 0$$

And thus

$$c^*(g_i, q - g_i) = H(g_i) \text{ with } H(g) = \frac{q(1-q) + 2g(q-g)}{2-q}$$

H is maximized in $g = q/2$ because $H'' = -\frac{4}{2-q} < 0$, which implies that H is maximized in $H' = \frac{q(1-q)+2g(q-g)}{2-q} = 0$. Solving for g yields $g = q/2$. As $q/2$ is the unconstrained maximum of H, we necessarily have $H(g_i) \leq H(q/2)$ as g_i is constrained by the equilibrium condition $F[c^*(g_i, q - g_i)] = M[c^*(g_i, q - g_i)]$. Hence $c^*(g_i, q - g_i) \leq q/2$ (as $H(q/2) = q/2$). Finally, we have seen that the corner maximum (c_0^* or c_q^* depending on q and F) is higher than $q/(2-q)$. This completes the proof as $q/2 < q/(2-q)$.

References

Akerlof, G., 1970. The Market for Lemon's: Quality Uncertainty and the Market Mechanism, *Quarterly Journal of Economics*, 84(3), 488-500.

Arrow, K., 1973. The Theory of Discrimination, In O. Ashenfelter and A. Rees, eds., *Discrimination in Labor Markets*. Princeton, N.J.: Princeton University Press, 3-33. 20

Avenhaus, R., von Stengel, B. and S. Zamir, 2002. Inspection games, Chapter 51, 1947-1987, in *Handbook of Game Theory*, eds. R. J. Aumann and S. Hart, North-Holland, Amsterdam. 6

Bagwell, K. and Riordan, M., 1991. High and Declining Prices Signal Product Quality, *American Economic Review*, Vol. 81(1), 224-239. 9

Bar-Isaac, H. and Tadelis, S., 2008. Seller Reputation, *Foundations and Trends in Microeconomics* 4(4), 273-351. 22

Blume, L., 2005. Learning and Statistical Discrimination, *American Economic Review (Papers and Proceedings)*, 95(2), 118-121. 23

Blume, L., 2005. The Dynamics of Statistical Discrimination, *Economic Journal*, 116(515), F480-F498. 23

Board, S. and Meyer-ter-Vehn, M., 2010. Reputation for Quality, *mimeo* UCLA. 6

Brock, W. and Durlauf, S., 2001. Discrete Choice with Social Interactions, *Review of Economic Studies*, Vol. 68(2), 235-260. 11

Coate, S., Loury, G.C., 1993. Will Affirmative-Action Policies Eliminate Negative Stereotypes? *American Economic Review*, 83(5), 1220-1240. 5, 20, 21

Dranove, D., and Jin, G.Z., 2010. Quality Disclosure and Certification Theory and Practice, *Journal of Economic Literature*, 48(4), 935-963. 5, 23

Dye, R., 1986. Optimal Monitoring Policies in Agency, *RAND Journal of Economics*, 17(3), 339-350. 6

Fagart, M.-C., and Sinclair-Desgagné, B., 2007. Ranking Contingent Monitoring Systems, *Management Science*, 53(9), 1501-1509. 6

Jackson, M. O. and Yariv, L., 2007. Diffusion of Behavior and Equilibrium Properties in Network Games, *American Economic Review (Papers and Proceedings)*, 97(2), 92-98. 11

Harbaugh, R., Maxwell, J. and Roussillon, B. 2011. Label Confusion: The Groucho Effect of Uncertain Standards. *Management Science*, forthcoming. 23

Kamenica, E. and M. Gentzkow, 2011. Bayesian Persuasion, *American Economic Review*, Vol. 101, 2590-2615. 6

Kessler, A., 2001. Revisiting the Lemons Market, *International Economic Review*, Vol. 42, 25-41. 5

Levin, J., 2001. Information and the market for lemons, *RAND Journal of Economics*, Vol. 32(4), 657-666. 5

Levin, J., 2009. The Dynamics of Collective Reputation, *The B.E. Journal of Theoretical Economics*, Contributions, vol. 9, art. 27, http://www.bepress.com/bejte/vol9/iss1/art27. 23

Lizzeri, A., 1999. Information Revelation and Certification Intermediaries, *RAND Journal of Economics*, Vol. 30(2), 214-231. 5

Lyon, T. (Ed.), 2010. *Good Cop/Bad Cop: Environmental NGOs and Their Strategies*, RFF Press, Washington, DC. 6

Lyon, T. and Maxwell, J., 2011. Greenwash: Corporate Environmental Disclosure under Threat of Audit, *Journal of Economics & Management Strategy*, Vol. 20(1), 3-41. 6

MacLeod, W. B., 2007. Reputations, Relationships, and Contract Enforcement, *Journal of Economic Literature* 45(3), 595-628. 6

Milgrom, P., 1981. Good News and Bad News: Representation Theorems and Applications, *Bell Journal of Economics*, Vol. 12(2), 380-391. 26, 27

Milgrom, P., and Roberts, J., 1986. Price and Advertising Signals of Product Quality, *Journal of Political Economy*, Vol. 94(4), 796-821. 9

Phelps, E., 1972. The Statistical Theory of Racism and Sexism, *American Economic Review*, 62, 659-61. 20

Sarath, B., 1996. Public Information Quality with Monopolistic Sellers, *Games and Economic Behavior*, Vol. 16, 261-279. 5

Tirole, J., 1996. A Theory of Collective Reputation (with applications to the persistence of corruption and to firm quality). *Review of Economic Studies*, 63, 1-22. 5, 22, 23