Incentives for Quality in Friendly and Informational Environments*

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

December 2015

Abstract

We develop a model of costly quality provision under biased disclosure. We define as friendly an environment in which the disclosure probability increases with quality, and as hostile an environment in which the opposite holds. Hostile environments produce a positive externality among sellers and potentially multiple equilibria. In contrast, friendly environments always yield a unique equilibrium. We establish that the environment that maximizes quality generates signals contradicting buyers' expectations. Hence hostility produces greater incentives for quality than friendliness when costs are low and monitoring resources high.

Keywords: Moral Hazard, Asymmetric Information, Biased Disclosure, Multiple Equilibria, Lemons.

JEL classification: D82.

*For both friendly comments and useful criticism, we would like to thank Heski Bar-Isaac, Simon Board, Alessandro Lizzeri, Bentley MacLeod, David Martimort, Moritz Meyer-Ter-Vehn, Nicolas Roux, Bernard Salanié, Andy Skrzypacz, Joel Sobel, Gabor Virag and seminar and conference participants at Bielefeld, Cergy, Columbia, CREST, EM Lyon, Guelph, IUPUI, Kellogg (Northwestern), Lille, Mines, MPI Bonn, Paris School of Economics, Polytechnique, Stony Brook, UCLA, EAERE 2011, GAMES 2012 and EARIE 2012.

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

[‡]MINES ParisTech, PSL - Research University, Centre for Industrial Economics, 60 Boulevard St Michel, 75006 Paris. Email: glachant@mines-paristech.fr.

[§]MINES ParisTech, PSL - Research University, Centre for Industrial Economics, 60 Boulevard St Michel, 75006 Paris. Email: gabrielle.moineville@mines-paristech.fr.

1 Introduction

We here consider how different informational environments can help shape quality provision. The distinctive feature of our analysis is the contrast between friendly and hostile environments. We define as friendly an environment in which favorable evidence on quality is more likely to be disclosed, and as hostile an environment in which the disclosure probability of unfavorable evidence is higher.

Consider, for concreteness, the example of a politician under media scrutiny seeking re-election. While in office, he can expend more or less effort on delivering a policy that will increase voter welfare. Voters cannot directly observe this effort, but receive news from a media that collects evidence about the politician's behavior and reports it, potentially in a selective way. The voters use this news to decide whether to re-elect the politician. Clearly, the nature and quality of media reporting is an essential determinant of the politician's behavior.

The media may be neutral. In this case, it will disclose any available piece of information on the politician's behavior. Alternatively, it can be hostile to the politician, in which case it reveals information only on bad behavior, as such news reduces the likelihood that the politician is re-elected. Or the media can be friendly, and thus selectively publicize only good behavior.

In this context, we try to answer two types of question. First, how do these differences in media reporting shape the politician's incentives? We thus start by analyzing the qualitative difference between equilibria in friendly and hostile environments. Second, is the politician more likely to behave better (i.e. supply quality) when the media is hostile, neutral or friendly? To answer this question, we consider the optimal design of the informational environment, and apply the results to evaluate different informational environments.

These questions are of substantial importance in understanding situations where the news available to a voter, or a buyer, varies by the type of information provided. Many monitors selectively report information. In the online-rating systems which play an important role in mitigating information asymmetry in electronic transactions, it has been found that traders on eBay are more likely to report positive feedback (Dellarocas and Wood, 2008). In a totally different context, environmental NGOs, which are an important source of information regarding corporate environmental performance, tend to specialize: using the terminology of Lyon (2010), some are "bad cops", like Greenpeace, and mostly report objectionable behavior to the general public. Others are "good cops", like the World Wildlife Fund, which develops partnerships with corporations in order to signal good performance. In a Court of Law, the Defense Lawyer collects evidence which is favorable to the defendant and is not meant to reveal unfavorable information; the Prosecutor does the exact opposite (Dewatripont and Tirole, 1999). Reporting bias may also be embedded in the underlying technology: doping tests, for instance, can only provide unfavorable evidence.

The evaluation of credence goods by multiple organizations with heterogeneous disclosure behavior is in fact widespread: media, auditors, rating agencies, non-governmental organizations, labeling institutions, peers, etc. Whether a particular informational environment is then neutral, friendly or hostile ultimately depends on the number, resources and preferences of teh monitors which are active in this environment. In our last section concerning applications, we will limit ourselves to discussing labeling policies, media bias and statistical discrimination in labor markets (Coate and Loury, 1993), but the insights apply to many other situations.

We present a simple model which covers all of the different environments mentioned above in a market setting. The model describes a continuum of agents who produce and sell a good whose quality is non-contractable. Quality can either be low, at no cost, or high, at a cost which varies across agents. Each agent first chooses the quality level, but this choice is not directly observable by the potential buyer. After the agent's choice of quality, there is monitoring. This discloses quality with a probability g if quality is high ("good news") and probability b if quality is low ("bad news"). After receiving this news (or not), the buyer updates their beliefs about quality and decides whether to purchase the good. We say that the informational environment is friendly if $g > b$ and hostile if $g < b$.

We can easily anticipate that differences between g and b will be anything but neutral. In particular, when the buyer doesn't receive any news after monitoring, their belief updating will be totally different: when the informational environment is friendly, monitoring mainly filters out high quality. As such, when no news is received about a product, quality beliefs become more pessimistic. In a hostile informational environment, no news will conversely increase beliefs. The model will describe how agents' quality investments interact with the buyer's belief updating.

We find sharp differences between the equilibria in friendly and hostile environments. An important result is the potential multiplicity of equilibria in hostile environments, whereas the equilibrium is unique in friendly environments. This difference is driven by the type of reputational externalities across (different types of) agents. When no news is revealed, buyer expectations naturally increase with the number of agents opting for high quality in the two environments, but this has the opposite effect on quality choice. If the environment is friendly, increased expectations reduce the individual incentive to increase quality. This produces free-riding amongst agents, as choosing high quality reduces other agents' incentives to do so: quality choices are strategic substitutes. If the environment is hostile, an increase in beliefs provides greater incentives (a bandwagon effect): quality choices are strategic complements. As a result, beliefs tend to be self-fulfilling in hostile environments, hence the possibility of multiple equilibria.

We also examine design issues. In particular, we characterize the information structure that maximizes incentives for quality.[1] In our framework, quality is obviously maximized when information is perfect. Hence the design question becomes interesting only if monitoring resources are limited. We introduce a cost for the information structure $K(g, b)$, which rises in g and b and which is symmetric: $K(g, b) = K(b, g)$ for all g and b. This symmetry assumption enables us to carry out an unbiased comparison between friendly and hostile environments. We show that a neutral environment with $g = b$ is generically inefficient and that, for a given information cost, a marginal tilt towards hostility increases quality when the cost of quality is low while a tilt towards friendliness is preferable when the cost of quality is high.

[1] Under our assumptions, this corresponds both to the socially-optimal environment and to the environment that a seller would choose before learning his cost.

With this result, we uncover a broad logic of dis-confirmation: a high cost of quality comes with little incentive to invest and, correspondingly, buyers' pessimistic beliefs about quality, on which they do not receive any feedback. In this context, our analysis says that acting against the buyer's beliefs via the emission of good news is the most effective way of increasing quality. We then fully characterize the optimal structure in the case where $K(g, b) = K(b + g)$. In such a case, the results are extreme: the optimal environment features either $g = 0$ or $b = 0$. We also identify the environment that maximizes the amount of news in equilibrium and derive a number of comparative-statics results.

Related Literature. To organize the discussion of the literature, it is useful to start with the classic distinction in information theory between binary symmetric communication channels (BSC) and binary erasure channels (BEC).[2] Both channels can transmit one of two symbols – say, low quality and high quality –, but this transmission is characterized by different types of imperfection. With a BSC, the receiver obtains the wrong bit with some positive probability; with a BEC sometimes the bit is "erased" so that the receiver has no idea what the bit was. A BEC thus generates three possible signals: the null signal and two perfect signals (high quality, low quality) in contrast with a BSC which conveys two possible noisy signals.

As noted above, our contribution is twofold: to highlight the key differences in incentive properties between friendly and hostile environments, and to design the informational structure that maximizes quality. We model the difference between friendliness and hostility by introducing heterogeneity in the erasure probabilities of a BEC. In this respect, our work is related to Board and Meyer-ter-Vehn (2013). The information structure analyzed there is similar to ours except that they focus on two special cases: a structure which they refer to as good news learning, where only good news is revealed ($b = 0$) and the opposite structure with $g = 0$. Their focus is on reputational dynamics, which leads them to adopt the setting pioneered by Abreu et al. (1991) where news is emitted over time according to a Poisson process. Their central result concerns the multiplicity of equilibria in what we term hostile

[2] See for instance MacKay (2003, pp. 147-148).

environments.[3] We here go further by considering the optimal design of informational environments, which requires a more general three-signal structure in order to compare biased disclosure with a standard neutral BEC.[4] In particular, the central results of our paper are Propositions 7 and 8 that show that neutral environments never maximize the incentives for quality and provide the conditions under which friendliness or hostility is preferable (i.e., the logic of dis-confirmation).

Our work is indirectly related to a paper by MacLeod (2007), who also explicitly distinguishes friendly and hostile environments under a BEC-type information structure.[5] MacLeod's work falls in the tradition of contract theory. He compares various contracts and enforcement mechanisms designed to counter the problem of informational asymmetry. By way of contrast, quality is not contractable in our paper, and is traded on a spot market at a price reflecting the information available ex-post. There is neither the potential for screening via a price menu, which is realistic for many applications, nor for signaling, as low-quality agents can duplicate at zero cost any pricing decision by high-quality agents.[6]

Gill and Sgroi (2012) consider a different BSC-type information structure with two noisy signals into which they introduce heterogeneous probabilities. In contrast to Board and Meyer-ter-Vehn, their objective is to study the choice of information structures. In their model, the quality of a new product is tested before launch. They deal with the "toughness" of different tests which can yield two results, a pass or a fail. A test is tough when the probability that a low-quality product passes is low, but at the price of a high-quality product potentially failing. A tough test is thus related to a friendly environment, as it identifies high-quality products better. In contrast to our contribution, quality is exogenous, implying that the incentive dimension at the heart of our analysis is absent from theirs.

[3] They also consider a variant with noisy signals, but do not provide a full characterization of the equilibria in that case, which is substantially more complicated.

[4] Board and Meyer-ter-Vehn's setting does not allow for neutral environments except in the limit cases with perfect information and no information.

[5] MacLeod's terminology is "normal good", for goods traded under hostile informational environments, and "innovative goods", for those traded in friendly environments. Information arrives only through good news, or only through bad news.

[6] Designing incentive solutions under asymmetric information is the preserve of contract theory. In the standard setting, the information structure is given, with only a few notable exceptions such as Kim (1995).

Apart from the work discussed above, the literature has in general overlooked the distinction between hostile and friendly environments. Nevertheless, our work is also related to a number of other contributions. On the design of information structures, an important contribution is the work of Kamenica and Gentzkow (2011) who provide a characterization of the optimal information structure with commitment in a sender-receiver framework. Their state, message, and action spaces in general contrast with ours, which features two qualities and three signals. However, as in Gill and Sgroi (2012), the state of the world is fixed,[7] whereas quality is endogenous in our paper. The nature of the design problem is thus fundamentally different: in our model, we are primarily interested in the *ex ante* incentives provided to the seller by informational feedback; Kamenica and Gentzkow (2011) and Gill and Sgroi (2012) focus on the *ex post* influence of the signals on the buyer.

A related question in information economics is the choice of information structure in price-discrimination problems (see Lewis and Sappington, 1994; Ottaviani and Prat, 2001; Johnson and Myatt, 2006, for important contributions). In that literature, a seller discloses information to exploit buyers' heterogeneity. Hence, the focus is again on the consequences of information disclosure on buyers, not on the incentives of sellers.

Formally, the statistical-discrimination model of Coate and Loury (1993) and project-evaluation model of Taylor and Yildirim (2011) are also closely related to ours, as they both consider the incentive effect of various institutional arrangements in a static setting. However, the distinction between friendliness and hostility remains implicit. The same remark applies to work on collective reputation (Tirole, 1996; Levin, 2009). Our model allows us to revisit some of the insights from this research, pointing out the necessity of hostility for equilibrium multiplicity. In other contributions, 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 the incentives of certifiers.[8] In this area of research, a voluminous literature in industrial organization has looked at the behavior of information intermediaries (e.g. Lizzeri, 1999; Strausz, 2005).

[7] Note also that Kamenica and Gentzkow (2011) analyze design at the ex ante stage, while Gill and Sgroi examine interim signalling.

[8] See in particular Dranove and Jin (2010) for a survey of these issues.

In the field of Law and Economics, the research by Dewatripont and Tirole (1999) on advocacy explores how competition between organizations that are biased towards special interests may jointly produce better information for decision making than a non-partisan investigator (see also Shin, 1998). They do not however address the incentive and deterrence properties that these systems generate in the first place. Our contribution is complementary, as they focus on the incentives of monitors and their strategic interaction while we ignore these aspects.

The same holds in the burgeoning literature in media economics which investigates how media may distort news reporting (see Gentzkow and Shapiro, 2008, for a recent review), but without looking at the consequences for incentives, with the notable exception of the empirical analysis of Snyder and Strömberg (2010). We discuss applications and point out the connections with this literature in our last section.

Finally, it is worth mentioning the recent literature on online trading platforms, with empirical analysis that has sought to measure reporting bias (Dellarocas and Wood, 2008) and the impact on seller performance (Cabral and Hortaçsu, 2010; Cai et al., 2014).

The structure of the paper is straightforward. Section 2 presents the model. In Sections 3 and 4, we characterize the equilibria and carry out comparative-statics exercises. In Section 5, we characterize the informational environment that maximizes quality incentives, and Section 6 then discusses applications. The last section concludes.

2 The model

We consider a game with a continuum of agents (sellers) who each produce one unit of a good whose quality is imperfectly observable by a representative buyer. We refer to sellers and a buyer for ease of presentation, although the setting is sufficiently abstract to cover many real-world situations. The buyer can represent consumers in a final market with vertical differentiation in which endogenous quality is not perfectly observable, a firm hiring employees whose intrinsic productivity resulting from past education investments is uncertain, lenders in capital markets, etc. The model can also be applied to non-market situations, such as the example of the politician set out in the introduction. Another application is school testing: a teacher has to grade students whose performance is not always observed, since monitoring intensity is limited by the teacher's time and the number of students being evaluated.

The quality variable q is binary, with $q \in \{0, 1\}$. Providing quality q costs the agent $c(q, \theta) = \theta q$, where θ varies across agents with a cumulative distribution $F(\theta)$ and a positive and continuous density $f(\theta)$ over $\left[\underline{\theta}, \overline{\theta}\right]$. The distribution F is common knowledge, but each agent privately observes their own actual cost.[9] For most of the presentation, we will assume that $\left[\underline{\theta}, \overline{\theta}\right] = [0, 1]$.

The buyer's willingness to pay for quality is exactly q. We assume that agents can fully extract the buyer's surplus. This provides the most incentives to the agents, and makes the welfare analysis transparent. Under these assumptions, were quality to be perfectly observable, the social optimum would result, as all agents would choose $q = 1$ and sell the good at price 1, since $\theta \leq 1$ for all types. In the Appendix, we relax these payoff assumptions and show that the results continue to hold when some agents have high costs, and hence never choose high quality, or when some agents have negative costs, and hence always choose high quality. We also discuss the case with a negative willingness to pay for low quality.

The buyer receives a signal $s \in \{q, \varnothing\}$ on the quality of the good supplied by each agent. Hence we assume that either quality is perfectly-revealed or no evidence is received.[10] Figure 1 illustrates the information structure. The signal is generated with the probabilities

$$g = \Pr[s = 1|q = 1] \text{ and } b = \Pr[s = 0|q = 0].$$

A signal $\varnothing$ is generated with probability $1 - g$ when $q = 1$ and $1 - b$ when $q = 0$. We label this signal as "no news", in the sense that it reveals no evidence, but is of course subject to Bayesian interpretation by the buyer.[11] The central feature of the model is that we allow

[9] Equivalently, this model represents a single agent with unknown cost and unobservable effort.

[10] We discuss in the Appendix the case where all signals are imperfect.

[11] The representation of a friendly (resp. hostile) environment by $g > b$ with binary quality carries through to a general model with continuous quality, where friendly (hostile) environments correspond to a revelation probability that rises (falls) in quality. The posterior belief upon receiving no news then dominates the prior in the sense of a monotone likelihood ratio if and only if the environment is hostile.

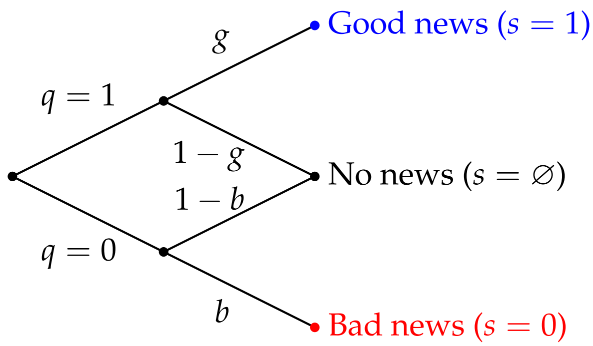

Figure 1: Information structure.

the disclosure technology to be asymmetric: the probabilities g and b may differ.

The full sequence of events is as follows:

- **Stage 0:** All players are informed of the nature of the informational environment (g, b).[12]

- **Stage 1:** The seller privately learns his cost θ and then chooses quality q.

- **Stage 2:** If the seller has chosen $q = 1$, the buyer receives "good news", $s = 1$, with probability g, and no news ($s = \varnothing$) with probability $(1 - g)$. If the seller has chosen $q = 0$, the buyer receives "bad news", $s = 0$, with probability b, and no news ($s = \varnothing$) with probability $(1 - b)$.

- **Stage 3:** The buyer buys at price 1 if $s = 1$, and at price zero if $s = 0$. When no news is received about the real quality, the buyer forms a belief $\mu = \Pr[q = 1|s = \varnothing]$. Given this belief, the seller can only selle the good at its conditional expected value μ, which

[12] This is by no means a restrictive assumption, as g and b can represent the probabilities inferred by players from the number of auditors and the knowledge of their preferences. In the politician and media example, this amounts to saying that voters are aware of the political bias in the media. More generally, any game that the auditors may play after receiving signals will ultimately generate an informational environment described by these two probabilities.

is the willingness to pay of the uninformed buyer: $\mu = \mu \cdot 1 + (1 - \mu) \cdot 0$.[13,14]

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

3 Equilibrium analysis

3.1 Incentives and cutoff equilibria

We adopt the notion of a perfect Bayesian equilibrium, where each type of agent chooses their best reply to the market's belief, and the buyer's belief is consistent with the distribution of quality offered.

Let $\Pi(q, \theta)$ be the expected payoff of an agent with cost θ. For a given belief μ, the possible expected payoffs are: $\Pi(1, \theta) = g + (1 - g)\mu - \theta$ and $\Pi(0, \theta) = (1 - b)\mu$. The agent then chooses $q = 1$ whenever[15] $\Pi(1, \theta) \geq \Pi(0, \theta)$, which translates into:

$$\theta \leq g - (g - b)\mu. \tag{1}$$

An almost immediate consequence of the incentive constraint is that all of the Bayesian equilibria of this game have a cutoff structure: they are all characterized by a cost threshold θ^* below which agents choose high quality and above which they choose low quality. Our first lemma states this formally.

Lemma 1. *All Bayesian equilibria of the game described above are cutoff equilibria. Any equilibrium is characterized by a threshold θ^* such that all sellers with $\theta \leq \theta^*$ choose $q = 1$*

[13] We may imagine that high-quality agents will try to signal their quality in the case of no news. But since quality is not contractable, a price offer can only be made after the emission of the signals. In the case that no evidence is disclosed, all equilibria are pooling since a low-quality agent can imitate the price offer of high-quality agents at no cost. There is thus no room for signaling, contrary to Milgrom and Roberts (1986) and Bagwell and Riordan (1991). The fact that all agents then offer μ can be justified on the grounds that this is the best offer for any agent under the pooling constraint. Alternatively, the same payoffs obtain when at least two identical buyers compete in price on the basis of the public signal.

[14] As in Gill and Sgroi (2012), this price plays a role similar to the acceptance/hiring standards in Coate and Loury (1993) and Taylor and Yildirim (2011).

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

and all sellers with $\theta > \theta^*$ choose $q = 0$. When $0 < \theta^* < 1$, this cutoff is given by:

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

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

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

Proof. Consider an equilibrium in the game in which the buyer's beliefs upon receiving no news is some μ^*, and suppose that there exists a $\hat{\theta}$ such that $\hat{\theta} \leq g - (b - g)\mu^*$. Then for all $\theta \leq \hat{\theta}$ we have $\theta \leq \hat{\theta} \leq g - (b - g)\mu^*$, so that the best reply to the equilibrium belief μ^* of all types below $\hat{\theta}$ is to choose $q = 1$. Similarly, if some type chooses $q = 0$, then all types above also choose $q = 0$. This establishes that an equilibrium is characterized by a cutoff θ^*, possibly at the boundaries of 0 or 1. Equation (3) results from Bayesian revision as the fraction of high-quality sellers is $F(\theta^*)$, the mass of types below the cutoff. □

3.2 Equilibrium characterization and multiplicity

It follows directly from the equilibrium condition (2) that a neutral environment with $g = b$ induces a unique equilibrium cutoff $\theta^* = g = b$. In this case, the cutoff does not depend on the cost distribution F. There is no strategic effect between the different types of agents. Note that this is the case usually considered in the literature. To explore the case in which the environment is not neutral ($g \neq b$), we combine (2) and (3) to obtain the following relationship:

$$F(\theta^*) = M(\theta^*|g, b), \quad \text{where} \quad M(\theta|g, b) \equiv \frac{(g-\theta)(1-b)}{(g-b)(1-\theta)}. \tag{4}$$

In this representation of the equilibrium, the left-hand side is the average product quality in equilibrium, which determines total surplus in different informational environments, and the right-hand side is a function M which contains all the data pertaining to the informational environment, independent of the distribution F. In the following, we will refer to M as the *disclosure curve*[16]. We now derive some properties of M that will be used in the subsequent analysis:

Lemma 2. *1. $M(.|g,b)$ is continuous over the interval $[0,1]$.*

2. $M(0|g,b) \geq 0$ and $\lim_{\theta \to 1} M(\theta|g,b) = -\infty$ in a friendly environment ($g > b$), whereas $M(0|g,b) \leq 0$ and $\lim_{\theta \to 1} M(\theta|g,b) = +\infty$ in a hostile environment.

3. The disclosure curve is decreasing and concave in a friendly environment; it is increasing and convex in a hostile environment.

Proof. See the Appendix. □

The first two properties and the intermediate-value theorem ensure the existence of equilibrium. An equilibrium cutoff is such that the graph of the cumulative distribution of costs intersects the disclosure curve. As shown in Figure 2, the equilibrium is necessarily unique in a friendly environment, as the disclosure curve is decreasing while F is increasing. But in a hostile environment M is increasing, so that there may be multiple intersections between F and M, as shown in Figure 3. We bring these results together in our first proposition.

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

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

3. In a hostile environment ($g < b$), there may exist multiple equilibria defined by (4). These are all such that $g \leq \theta^* \leq b$.

The main result is the potential multiplicity of equilibria in hostile environments. This is driven by the existence of reputational externalities: the equilibrium belief (3), which is

[16] There are other ways to combine (2) and (3). For instance, we can use a fixed-point representation in μ, as below.

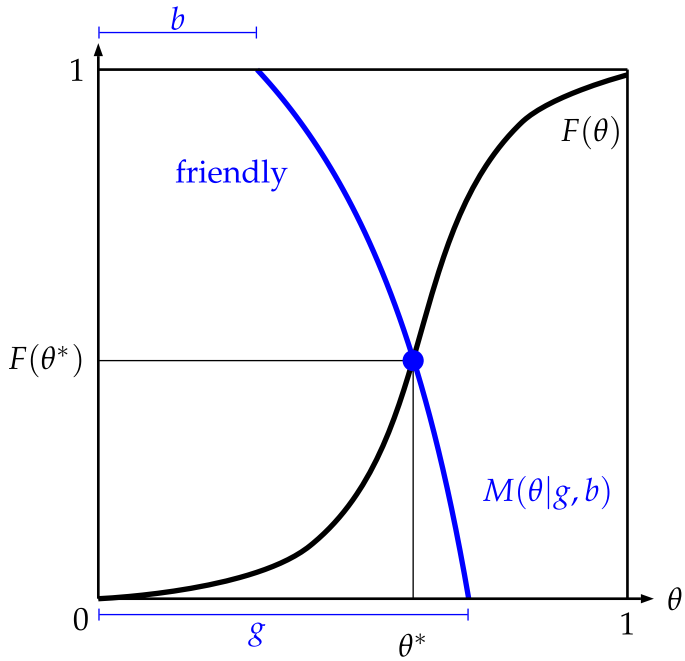

Figure 2: The representation of equilibrium in $(\theta, F(\theta))$-space.

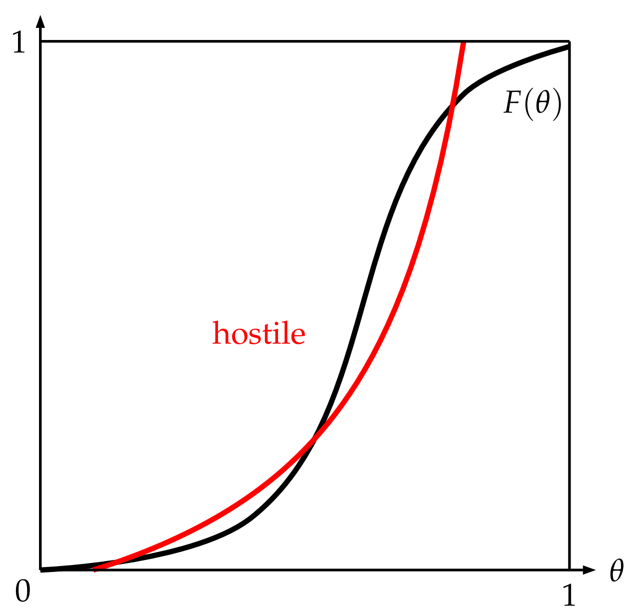

Figure 3: A case of a hostile environment with multiple equilibria in $(\theta, F(\theta))$-space.

also the price of unidentified products, increases with the mass of agents opting for high quality, $F(\theta^*)$. This creates externalities between agents which have opposite effects on individual quality choice in the two informational environments. To see this, we reconsider the incentive constraint (1): the last term on the right-hand side, $-(g-b)\mu$, is negative if $g > b$ and positive if $g < b$. This means that:

- If the environment is friendly, an increase in the belief μ reduces the incentive to increase quality. This produces free-riding amongst agents, as choosing high quality reduces other agents' incentives to do so: quality choices are strategic substitutes.
- If the environment is hostile, an increase in the belief provides greater incentives (a bandwagon effect): quality choices are strategic complements.

As a result, beliefs tend to be self-fulfilling in hostile environments. To illustrate this point, consider the simple example of a purely hostile environment with $g = 0$, $b = 2/3$, and where θ has a uniform distribution. Applying lemma 1, straightforward calculation shows that there are two equilibria: $\theta^* = 0$ and $\theta^* = 1/2$. A natural interpretation is as follows: if buyers are very pessimistic, with beliefs $\mu = 0$, this belief can never be proven wrong because the disclosure technology never produces good news. There is thus no point in sellers raising quality, as they will never be rewarded for doing so, and $\theta^* = 0$. In turn, in an equilibrium with optimistic beliefs, the price of unidentified products is positive, creating incentives for low-cost sellers to raise quality as avoiding bad news becomes worthwhile.

When there are multiple equilibria in hostile environments, we will now show that some equilibria are not stable in the sense that price adjustments following a small perturbation do not lead back to the initial equilibrium. To define stability in our context, we use a fixed-point representation of the equilibrium in the price of unidentified quality, μ. Substituting (2) in (3) yields:

$$\mu^* = \varphi(\mu^*), \quad \text{where} \quad \varphi(\mu) = \left(1 + \frac{1-b}{(1-g)Q(\mu)}\right)^{-1} \tag{5}$$

In this expression $Q(\mu) = \frac{F(g-(g-b)\mu)}{1-F(g-(g-b)\mu)}$ is the ratio of high-quality to low-quality goods contingent on μ. Equation (5) thus gives the price of unidentified products as a function of average market quality.

In hostile environments, we have seen that an increase in μ increases the incentives to provide high quality, which formally implies that $\varphi' > 0$. Consider now a small positive perturbation of the belief/price μ^* so that the buyer assigns a probability $\mu^*+\varepsilon$ that unidentified products are of high quality. Sellers then react by increasing quality (and thus φ), which in turn leads to an increase in the belief. If supply responds strongly (because there are many sellers with costs slightly higher than the cutoff θ^*), this initiates a dynamic movement which leads to divergence from the equilibrium price. More precisely, this occurs if the sequence $\mu^* + \varepsilon$, $\varphi(\mu^* + \varepsilon)$, $\varphi(\varphi(\mu^* + \varepsilon))$,... does not converge to μ^*.

This actually corresponds to a standard notion of stability, whereby the equilibrium is stable if μ^* is an attractive fixed point:[17]

Definition 1. *An equilibrium is stable if* $\left|\frac{d\varphi}{d\mu}(\mu^*)\right| < 1$. *An equilibrium is unstable if* $\left|\frac{d\varphi}{d\mu}(\mu^*)\right| \geq 1$.

Coming back to the geometric representation where the equilibrium is the intersection of M and F, simple calculations show that, in hostile environments, this definition of stability is equivalent to:[18]

$$f(\theta^*) < \frac{\partial M}{\partial \theta}.$$

Hence an equilibrium is stable if and only if the disclosure curve crosses the distribution from above. Consider the example in Figure 3 which features three equilibria. The lowest and highest equilibria are stable, as F crosses M from above. On the contrary, the intermediate equilibrium is not, meaning that an upward perturbation leads to the highest equilibrium

[17] 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. Hence the situation we analyze shares the stability properties of their adoption game.

[18] Differentiating φ yields $\varphi'(\mu) = (b-g)\left(\frac{\mu}{F(g-(g-b)\mu)}\right)^2\left(\frac{1-b}{1-g}\right) f(g-(g-b)\mu)$. The stability condition $\left|\frac{d\varphi}{d\mu}(\mu^*)\right| < 1$ is thus equivalent to $f(\theta^*) < \left(\frac{F(\theta^*)}{\mu^*}\right)^2 \frac{1-g}{(1-b)(b-g)}$ when $b > g$. We then substitute $\mu^* = \frac{g-\theta^*}{g-b}$ and $F(\theta^*) = M(\theta^*|g,b) = \frac{(\theta^*-g)(1-b)}{(b-g)(1-\theta^*)}$ in this expression. This leads to $f(\theta^*) < \frac{(1-b)(1-g)}{(1-\theta)^2(b-g)}$. It is then immediate that the LHS is equal to $\frac{\partial M}{\partial \theta}$.

and a downward perturbation leads to the lowest equilibrium. By analogy with adoption games, unstable equilibria represent tipping points between two stable configurations in hostile environments: they correspond to a critical mass of agents producing quality below which a perturbation would lead to a low-quality equilibrium and above which a high-quality equilibrium would prevail.

It is important to note that there are always unstable equilibria in the case of multiplicity. Stability analysis can thus yield the conditions for the existence of multiple equilibria.

Proposition 2. *The necessary and sufficient conditions for multiple equilibria are that (i) the informational environment is hostile ($g < b$) and (ii) there exists an equilibrium such that*

$$f(\theta^*) > \frac{(1-b)(1-g)}{(1-\theta^*)^2(b-g)}. \tag{6}$$

This result is typically important for statistical-discrimination models, as multiplicity is at the heart of the issue there: this is discussed in the application section. Condition (6) shows that multiple equilibria are more likely to occur when the gap between b and g is large. Conversely, for a given distribution, there always exists a unique equilibrium provided that $(b-g)$ is sufficiently small compared to the maximal value of f: supply is then not responsive enough that small changes in the price μ^* destabilize the equilibrium.

4 Comparative Statics

We now investigate how changes in the informational environment and cost distribution affect the equilibria.

4.1 Changes in the informational environment

4.1.1 Quality in equilibrium

We start by investigating how changes in g and b affect the cutoff θ^*, and thus average quality in equilibrium. It is easy to obtain the following comparative statics:

Lemma 3. *All else equal, the equilibrium cutoff rises with g and b when the equilibrium is unique. When there exist multiple equilibria, the cutoff rises with g and b if and only if the equilibrium is stable, and falls otherwise.*

Proof. Differentiating the equilibrium condition (4) with respect to g and b, we have:

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

We then differentiate M with respect to g and b:

$$\frac{\partial M}{\partial g} = \frac{(1-b)(\theta - b)}{(1-\theta)(g-b)^2} \text{ and } \frac{\partial M}{\partial b} = \frac{(1-g)(g-\theta)}{(1-\theta)(g-b)^2}.$$

If $g > b$, the equilibrium is unique, and we know that $f(\theta^*) - \frac{\partial M}{\partial \theta} \geq 0$ as $\frac{\partial M}{\partial \theta} < 0$. $\frac{\partial M}{\partial g}, \frac{\partial M}{\partial b} \geq 0$ follows from $b \leq \theta^* \leq g$ (Proposition 1). If $g < b$, we have $\frac{\partial M}{\partial g}, \frac{\partial M}{\partial b} < 0$ (as $g < \theta^* \leq b$). Hence the sign of $\frac{\partial \theta^*}{\partial g}$ and $\frac{\partial \theta^*}{\partial g}$ is the same as that of $f(\theta^*) - \frac{\partial M}{\partial \theta}$, which is ambiguous as $\frac{\partial M}{\partial \theta} > 0$, and corresponds exactly to the stability condition above. □

This lemma states that better information creates greater incentives to supply high quality.[19] This is immediately clear when the equilibrium is unique, as a higher cutoff means higher average quality in equilibrium. This is less direct when there are multiple equilibria, as the cutoff falls with g and b when the equilibrium is unstable. However, remember that unstable equilibria can be interpreted as tipping point, hence the lemma says that the associated critical mass falls in (g, b). Hence, we can write:

Proposition 3. *Average quality in equilibrium increases with g and b.*

4.1.2 News in equilibrium

We now consider how news arrival depends on the information structure. The quantity and nature of news in equilibrium is of primary interest for actors who are specialized in

[19] That more information is always better for trade, and hence for incentives in the first place, may not be true in the original lemons market of Akerlof (1970) with respect to seller's information (see Kessler, 2001; Levin, 2001). Here such considerations do not arise since the seller is perfectly informed.

information production, such as press agencies, certifiers, or the media. The question is also relevant in empirical work as, in many cases, the stream of news is more easily observed by the econometrician than the underlying disclosure parameters (g,b). This is, for instance, the case in Dellarocas and Wood (2008), who try to infer the propensity of trading partners to provide positive, negative, or no feedback on eBay from feedback observations.

From an empirical point of view, the variables of interest are $G(g, b) \equiv gF(\theta^*(g, b))$ and $B(g, b) = b[1 - (F(\theta^*(g, b))]$, which are the equilibrium quantities of good and bad news, respectively. Although well-defined in theory, the observation of a lack of news reflected in $1 - G - B$ poses empirical difficulties. We thus ignore it in the following.

In order to make the analysis transparent, we assume that the equilibrium is unique. We obtain the following proposition.

Proposition 4. *Suppose there is a unique equilibrium. Then:*

1. *The equilibrium quantity of good news rises in* (g, b)*.*

2. *The equilibrium quantity of bad news falls in* g*. The profile of bad news is non-monotonic in* b *since* $B(g, 0) = B(g, 1) = 0$*, and the effect of* b *is ambiguous in the interior.*

Proof. The result for good news is obvious as $G(g, b) = gF(\theta^*(g, b))$ rises in (g, b). Bad news falls in g since there is less low quality in equilibrium, and low quality is detected with a constant probability of b. In turn, we have $\frac{\partial B}{\partial b} = 1 - F(\theta^*(g, b)) - bf(\theta^*(g, b))\frac{\partial \theta^*}{\partial b}$, which in general is of ambiguous sign, except for extreme b's. □

Good news trivially increases in both g and b as more information provides greater incentives, as seen in Proposition 3, and so does $gF(\theta^*)$. But the main point in the proposition is that the effect of greater hostile disclosure on the equilibrium amount of bad news is in general ambiguous. The intuition is straightforward: although greater hostile disclosure directly increases the quantity of bad news, it is also associated with more agents choosing high quality, which then reduces the likelihood of bad news.

4.2 Changes in costs

In this subsection we analyze the consequences of a change in costs. In particular, we consider a fall in the sense of First-Order Stochastic Dominance (FOSD). We then establish the following proposition.

Proposition 5. *Consider a fall in costs in the sense of FOSD and suppose the equilibrium is unique. The average quality supplied in equilibrium then rises in all environments. However,*

1. *The equilibrium cutoff falls in friendly environments.*

2. *The equilibrium cutoff rises in hostile environments.*

Lower costs hence make hostile environments relatively better in the provision of incentives compared to friendly environments.

Proof. Consider a family of distributions $F(\theta;\lambda)$ with $\frac{\partial F}{\partial \lambda} > 0$, so that, if $\lambda' > \lambda$, $F(.;\lambda)$ first-order stochastically dominates $F(.;\lambda')$. From (4) we have:

$$\frac{d\theta^*}{d\lambda} = -\frac{\frac{\partial F}{\partial \lambda}}{f(\theta^*;\lambda) - \frac{\partial M}{\partial \theta}}. \tag{7}$$

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

For the equilibrium fraction of good quality, we then have:

$$\frac{d}{d\lambda}\left(F(\theta^*(\lambda);\lambda)\right) = f(\theta^*;\lambda)\frac{d\theta^*}{d\lambda} + \frac{\partial F}{\partial \lambda}.$$

Plugging (7) into this expression yields:

$$\frac{d}{d\lambda}\left(F(\theta^*(\lambda);\lambda)\right) = -\frac{\frac{\partial F}{\partial \lambda}\frac{\partial M}{\partial \theta}}{f(\theta^*;\lambda) - \frac{\partial M}{\partial \theta}},$$

which is positive, as $f(\theta^*;\lambda) - \frac{\partial M}{\partial \theta}$ and $\frac{\partial M}{\partial \theta}$ have opposite signs in both environments with stable equilibria. □

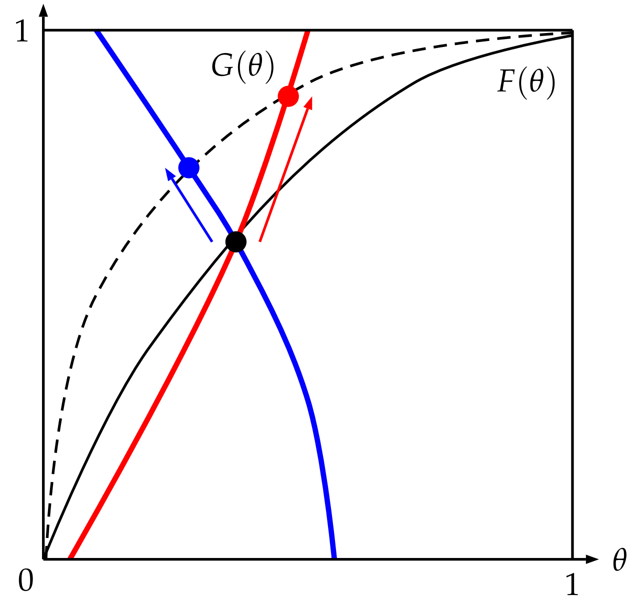

Figure 4: The relative advantage of hostile environments as costs fall.

Figure 4 illustrates these results for the special case where the two environments initially perform equally well with a cost distribution F. Lower costs in the sense of first-order stochastic dominance consist of a shift in the cost distribution from F to G. As costs fall from F to G, the hostile environment produces higher average quality.

The fact that reducing the cost of quality increases the supply of high quality is not surprising. But why is this effect smaller in friendly environments? It is useful to recall the contrast noted above between the two environments. In friendly environments, greater buyer belief μ reduces the incentive to supply high quality, leading to an incentive to free ride. This free-riding attenuates the positive effect of a greater supply of high quality from lower costs. On the contrary, in hostile environments the self-reinforcing mechanism renders agents' quality choices strategic complements, which amplifies the increased supply of high quality.

In order to streamline the presentation, we will not formulate a specific proposition for the case of a hostile environment with multiple equilibria. The situation is rather more complicated in this case, as some equilibria may disappear after FOSD shifts in F, except for the higher equilibrium. But bearing in mind that unstable equilibria are tipping points, the overall message is entirely in line with that of in Proposition 5, since lower unstable equilibria imply that there is a greater probability of shifting to a higher equilibrium.

Turning now to the impact of changing costs on the quantity and nature of news in equilibrium, the analysis is straightforward, as a change in costs essentially increases the share of high-quality agents. We summarize this in the next proposition; the proof is left to the reader.

Proposition 6. *Suppose the equilibrium is unique. All else equal, if costs fall in the sense of First-Order Stochastic Dominance:*

1. *The quantity of equilibrium good news increases and that of equilibrium bad news falls.*

2. *The quantity of news generated in equilibrium,* $G(g, b) + B(g, b)$*, rises in friendly environments and falls in hostile environments.*

This proposition might help us to identify the nature of the informational environment if we observe an exogenous shock to costs. This in particular holds for the second part of the proposition, as it is sometimes hard to disentangle empirically whether news is good or bad. This may allow us to draw inferences on the nature of the informational environment uniquely on the basis of whether the topic becomes more or less popular in a news stream after an exogenous cost shock.

5 The design of the informational environment

Up to this point, g and b have been considered as exogenous. In this section, we endogenize the informational environment. More precisely, we here try to identify the informational structure (g, b) that maximizes social welfare, which amounts to maximizing expected quality in equilibrium. Under full information, the equilibrium is first-best efficient. A non-trivial analysis thus requires the introduction of some frictions which prevent the choice of $g = 1$ or $b = 1$.[20] One obvious way of doing so consists in introducing a cost of the information

[20] In the influential work of Kamenica and Gentzkow (2011) on the design of the information structure, the designer – their terminology is "sender" – can commit ex ante to a structure which transmits all of the

structure,[21] $K(g,b)$. To remain as general as possible, we only assume that K rises in g and b and is symmetric: $K(g,b) = K(b,g)$ for any b. The symmetry assumption allows for an unbiased comparison between friendly and hostile environments.

We start by showing that a neutral environment ($g = b$) is generically inefficient. We previously saw that the equilibrium incentive constraint is $\theta \leq g - (g-b)\mu$, and that the equilibrium is unique in a neutral environment. The incentive to choose high quality is thus captured by the term $g - (g-b)\mu^*$. A marginal rise in g then increases incentives by $1 - \mu^* - (g-b)\frac{d\mu^*}{dg}$, which simplifies to $1 - \mu^*$ in a neutral environment. In the same way, a marginal increase in b increases incentives by $\mu^* - (g-b)\frac{d\mu^*}{db}$, which simplifies to μ^*. Note moreover that $\mu^* = F(\theta^*)$ when $g = b$.[22] As a result, average quality rises more with b than with g when $F(\theta^*) > \frac{1}{2}$, and conversely if $F(\theta^*) < \frac{1}{2}$. In other words, a marginal rise in g is more efficient than a marginal increase in b if and only if less than half of quality is good. Only when $F(\theta^*) = \frac{1}{2}$ are the marginal increases in g and b equivalent. It is thus always possible to strictly increase quality in a neutral environment by tilting the environment marginally, i.e. by trading off g for b in one direction or the other, without altering total costs (as K is symmetric).

We have thus proved the next proposition.

Proposition 7. *When the cost of the information structure $K(g,b)$ is symmetric in its two arguments, a neutral environment with $g = b$ is generically not efficient. To increase quality, keeping the cost constant, a marginal change towards friendliness is more effective when $F(\theta^*) < \frac{1}{2}$ and towards hostility when $F(\theta^*) > \frac{1}{2}$.*

This proposition highlights the positive effect of disconfirming buyer's beliefs on incentives. When the Bayesian buyer is pessimistic, and rightly so as only a minority of agents

information. She/he will not choose this structure as some news may reduce her/his payoffs. Our game is very different. Quality is endogenous and the designer is thus primarily concerned by how the structure affects the level of incentives. As average quality rises with the quantity of information (see Proposition 3), we need a restricted set of signals to yield non-trivial insights.

[21] Recently, Martin (2012) and Gentzkow and Kamenica (2014) have used a cost function for information structure based on entropy, inspired by Sims (2003). In this approach, the cost depends on entropy reduction with respect to a given prior. We cannot apply this approach here for the fundamental reason that the prior over quality before trade is endogenous.

[22] This results directly from plugging $g = b$ into Eq.(3).

supply high quality ($F(\theta^*) < 1/2$), a marginal deviation towards friendliness is preferable. Alternatively, if beliefs are optimistic ($F(\theta^*) > 1/2$) it is then more efficient to motivate agents by contradicting the buyer's beliefs with bad news.

The limit of Proposition 7 is that it only applies to marginal changes the informational structure around the neutral environment. In order to identify the optimal information structure, we will further specify the cost function. We now introduce the assumption that $g + b \leq r$, where $r < 1$. This amounts to saying that K depends on the sum $g + b$ and that the budget is limited, so that, abusing notation: $K(g + b) \leq K(r)$.[23] This assumption can also be directly interpreted, with r being the maximum amount of information that can be disclosed, for instance due to limited airtime in the case of media, or as monitoring resources are limited. One example of such a constraint is a committee of experts in charge of quality evaluation with a given size. Each committee member is imperfectly informed over some particular aspect of quality, but some members have preferences in favor of the agent (they are friendly), and other members have preferences against the agent (they are hostile). Experts may selectively withhold any evidence they obtain, if doing so favors their interests. The probabilities g and b then depend linearly on the composition of the committee. For clarity, we also restrict the analysis to concave distributions, which implies the uniqueness of the equilibrium in all environments. [24]

Under these assumptions, the socially-optimal informational environment (g^*, b^*) is the solution of the following optimization program:

$$\max_{g,b} \quad \theta^*(g, b) \quad \text{subject to} \quad g + b \leq r \tag{8}$$

We then show:

Proposition 8. *Suppose F is concave and monitoring resources are limited: $g + b \leq r$. Then:*

[23] Gill and Sgroi (2012) make a comparable linearity assumption on available information structures. In their model, the linear relation comes from a bound on total type I and type II errors in an imperfect test.

[24] The proof appears in the Appendix. Alternatively, the next proposition would read identically were we to restrict our attention to the highest equilibrium, which is always stable.

1. *The socially-optimal informational environment is extreme:* $g^* = 0$ *or* $g^* = r$*.*

2. *The socially-optimal informational environment is purely friendly if* $F(\frac{r}{2-r}) \leq \frac{1}{2}$ *and hostile otherwise.*

Proof. See the Appendix. □

The condition on $F(\frac{r}{2-r})$ confirms the logic of disconfirmation which we highlighted in Proposition 7. When $F(\frac{r}{2-r}) < \frac{1}{2}$ friendliness is preferable when there is only limited information (low r) and the cost of quality is high (in the sense of FOSD). Low r means that little information is disclosed about firm activities. As a result, there is little gain to being a high-quality firm and incentives are limited. With this in mind, the Bayesian buyer holds pessimistic beliefs in the absence of news. The buyer will be similarly pessimistic when F is low, as this implies that high quality is likely too costly for the agent.

The proposition's result that the optimal structure is extreme (either purely hostile or purely friendly) holds more generally under substitutability between g and b in information costs. If there is on the contrary strong complementarity between investing in g and b (say with a negative cross-partial derivative: $\partial^2 K/\partial g \partial b << 0$), then this extreme-information result will in turn no longer hold. But for lower degrees of complementarity, the extreme result will still hold, as can be seen from the proof in the Appendix. Clearly, at one extreme, if a rise in either g or b is mirrored in the other parameter – say, the information cost $K(g, b)$ depends only on $\max(g, b)$– then the optimal environment is obviously neutral. At the other extreme if, say for credibility reasons, increasing g is not feasible when $b \neq 0$, or vice versa, then the optimal information structure will mechanically be extreme. Which extreme structure should then be chosen is indicated by the proposition.

6 Applications and discussion

Our analysis yields two broad results. The first concerns the conditions for multiple equilibria and the second the socially-optimal environment. In this section, we consider different branches of the literature in order to illustrate how these findings can generate new insights in specific applications.

6.1 Quality disclosure and certification

The model can provide novel insights into quality certification (for a survey of this literature, see Dranove and Jin, 2010).

We will study the standard case of a binary product label which selectively signals high quality. The certifier decides to grant the label on the basis of a private evaluation. The certification test is imperfect: with probability $m < 1$, the test is fully informative about q, while it generates a null signal with probability $1 - m$. The testing technology is thus a priori neutral: it either uncovers the true quality or discloses nothing.

Consider now two alternative labeling policies in our framework. Under the *strict* policy, the certifier grants the label only if $q = 1$. Under the *lenient* policy, the label is granted only if the certifier does not observe $q = 0$. The consumer observes only whether a product is certified or not. We also assume for simplicity that the certifier is truthful. Under the strict policy, the consumer will thus be sure that the quality of a labeled product is high, but she is uncertain about the quality of non-labeled products. Under the lenient policy, she remains uncertain about the quality of labeled products, but is perfectly informed about the (low) quality of non-labeled products. Finally, we adopt the timing of the model where firms first choose quality, then apply for certification, and the representative consumer then purchases the good at a price equal to her willingness to pay. [25]

The corresponding information structure is shown on Figure 5. In the case where the label is strict, the price is thus 1 for labeled products (as the consumer knows that $q = 1$) and μ for non-labeled products. Under a lenient label, the price is μ for labeled products and 0 for the others. Strict labels, which only signal high quality, are thus associated with higher product prices.

[25] We abstract here from a number of complexities: the allocation of the gains from trade between the buyer and the sellers; the participation constraint of the consumer – she always purchases the good even if quality is bad. Relaxing these assumptions would however not alter the analysis qualitatively. The latter conditions is actually relaxed in Appendix A.1.

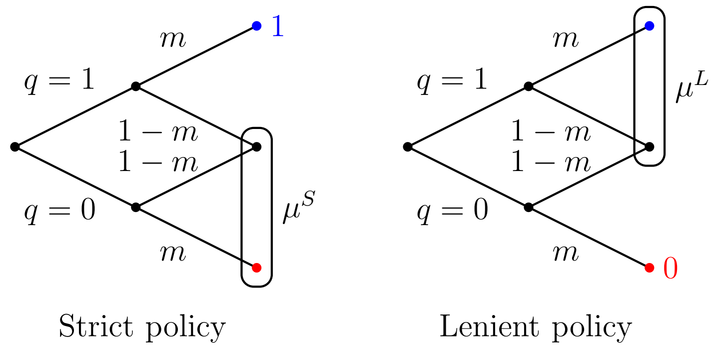

Figure 5: Labeling policies.

It is then easy to map this structure onto our friendly versus hostile distinction:

- When the certifier uses a strict rule, the firm's expected profit from high quality is $\Pi(1,\theta) = [m + (1-m)\mu^S] - \theta$ and $\Pi(0,\theta) = \mu^S$, where μ_S is the price of non-labeled products. Hence, the incentive constraint is $\theta \leq m(1-\mu^S)$. This corresponds to a purely friendly environment with $g = m$ and $b = 0$.

- When the certifier uses a lenient rule, we have $\Pi(1,\theta) = \mu^L - \theta$ and $\Pi(0,\theta) = (1-m)\mu^L$, where μ_L is the price of labeled products, implying the incentive constraint $\theta \leq m\mu^L$. This corresponds to a purely hostile environment with $g = 0$ and $b = m$.

We can thus assimilate a strict (lenient) label to a friendly (hostile) informational structure. A first immediate application of Proposition 1 is then that there may be multiple equilibria with lenient labels. The label can be recognized and create strong incentives, or fall flat, and lead to mild or even non-existent differences between labeled and non-labeled products–which equilibrium will emerge ultimately depends on producer coordination and buyer expectations. By way of contrast, a strict label has more predictable market consequences, and always creates incentives, even though it allows non-labeled products to partially free-ride on average quality (as $\mu^S > 0$).

Assume now that the cost distribution is concave, so that there is a stable equilibrium under both policies (see the proof in the Appendix). Proposition 8 indicates that strict labels are socially preferable when evaluation is difficult (low m) and/or when the cost of

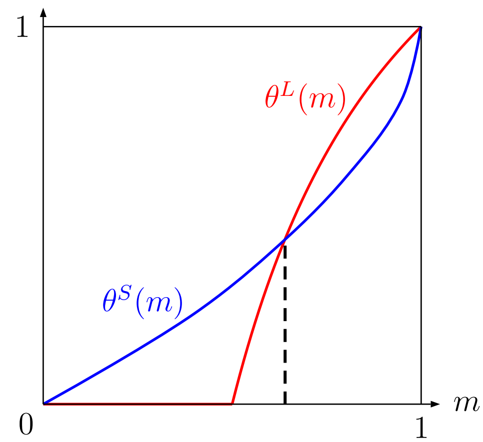

Figure 6: Strict vs Lenient: average quality as a function of m (uniform distribution).

high quality is relatively high. The case of a uniform distribution appears in Figure 6, showing that average quality is higher under a lenient label if and only if the monitoring technology is good enough (with m being above a threshold).

Noisy certification tests and uncertain labels have been analyzed by Harbaugh et al. (2011) and Mason (2011), who both propose a model where labels provide consumers with uncertain information and investigate the firms' incentives to adopt these labels. In our analysis, however, quality is exogenous, which rules out the key mechanism generating multiplicity and the assessment of quality incentives.

6.2 Statistical discrimination

Equilibrium multiplicity is the central notion in models of statistical discrimination, which are used to explain group inequality. In the literature starting from the seminal contribution of Arrow (1973), average group differences endogenously result in equilibrium without the assumption of any ex ante exogenous differences between groups.[26] One prominent contribution is 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, implying that ex ante identical groups

[26] In contrast to Phelps (1972), where ability is on average different in different groups and/or information on the various groups is more or less precise.

can have different average levels of qualifications ex post (and thus different wages). The literature on statistical discrimination does not provide any clear statement on the conditions under which this discrimination occurs. Proposition 1 in Coate and Loury (1993, p. 1126) does provide a necessary and sufficient condition for multiple equilibria, but the condition is not easy to interpret (and they indeed do not actually try to interpret it).

Our model sheds new light on this issue by relating the potential for multiple equilibria to hostile informational environments. It thus invites us to focus on the role of the technology used to evaluate students, and worker productivity in the labor market, which may generate novel policy implications for the design of evaluation tests: these should be friendlier, in the sense of friendliness in this paper.

However, the results also suggest a possible contradiction with the objective of increasing qualifications: in a hostile environment with multiple equilibria, making testing technologies friendlier reduces average qualifications in equilibrium when evaluation is relatively easy for intrinsic reasons, or monitoring resources are not that scarce ($g + b$ can be high), or when increasing qualifications is cheap, as suggested by Proposition 8. But even in the case where friendliness induces the highest level of average qualification (when $g + b$ is low and qualifications are costly), making testing friendlier has opposing impacts on the equilibria. In particular, it increases the lowest equilibrium, where those discriminated against tend to be found, but reduces the top equilibrium, which is likely to apply to those who are not discriminated against.

Our framework also provides some insights into the effectiveness of other policy approaches. For example, subsidizing the educational investments of discriminated groups (reducing θ in our model) is all the more effective when evaluation is hostile (Proposition 5). At this stage, these points have not been thoroughly established, but they do suggest new research directions.

6.3 Media bias

Following on from empirical observations (Groseclose and Milyo, 2005), a number of papers have proposed theoretical explanations of bias in news reporting. Focusing on the demand side, Mullainathan and Shleifer (2005) derive media bias from the assumption that consumers hold beliefs that they like to see confirmed – a bias for confirmation that has been extensively confirmed in experimental psychology. Alternatively, Baron (2006) and Gentzkow and Shapiro (2006) focus on the supply side and analyze the incentives that reporters and editors have to manipulate news. Besley and Prat (2006) consider the possibility of media capture by the government.

The representation of media bias varies across papers. In Mullainathan and Shleifer (2005), the media selectively filters facts. Other authors assume false reporting (Baron, 2006). In line with the idea that the function of media is to summarize complex issues by aggregating across a number of dimensions, Duggan and Martinelli (2011) formalize the concept of media slant as a relative emphasis on the different dimensions. None of these theoretical papers look at the impact of media bias on the behavior of the agents (politicians, corporations) who are under scrutiny, and the possible feedback on media behavior. Empirically, Snyder and Strömberg (2010) consider the feedback loop in a political context, from media reporting to politicians' and voters' behavior, and provide evidence of incentive effects on politicians.

Our information structure can be interpreted in terms of the selective filtering of facts, as in Mullainathan and Shleifer (2005): the media receives signals about agents' behavior (e.g., politicians, civil servants, firms). Then, as airtime or the number of pages is limited, the media has to select the news to be reported. In this context, reporting is biased if g differs from b, and $g + b \leq r$ represents the capacity constraint, where r is the total number of pages or airtime available to communicate news. Note that since the news receiver is Bayesian, this produces no gap between equilibrium beliefs and the truth, on average, even though reporting is biased.

In the media economics literature, media bias is viewed negatively. In contrast, by looking at incentives, we come to the opposite conclusion that neutrality can reduce social welfare as it does not maximize agents' incentives to adopt prosocial behavior. Proposition 8 sets out the types of bias which improve or reduce welfare: for example, the demand-driven confirmation bias highlighted by Mullainathan and Shleifer (2005) is socially detrimental compared to neutrality.

Our framework can also generate supply-driven biases which reduce welfare. Consider for example the case of firms that are primarily interested in producing evidence, such as news agencies. Their objective is thus to maximize the quantity of news in equilibrium. Let us generically call them monitors Assume that these monitors are atomistic, to capture the idea that they do not have individual market power. As such, they will not take the incentive feedback of their informational strategy into account. Which monitoring technology will these firms choose?

Let $x \in [0, 1]$ denote the index of these monitors, which each choose $g(x)$ and $b(x)$. Monitor x seeks to maximize the quantity of news they report with limited resources, so that $g(x) + b(x) \leq r(x) < 1$. Let $g = \int_0^1 g(x)dx$ and $b = \int_0^1 b(x)dx$; we hence assume each monitor is specialized so that they all report different pieces of information,[27] and let the aggregate potential information flow be $r = \int_0^1 r(x)dx$. Given the atomistic assumption, each news provider takes $\theta^*(g, b)$ as given since their own impact is negligible. Formally, monitor x thus solves:

$$\max_{g(x),b(x)} \quad g(x)F(\theta^*) + b(x)[1 - F(\theta^*)] \quad \text{subject to} \quad g(x) + b(x) \leq r(x) \tag{9}$$

An equilibrium is then completed by the condition that $F(\theta^*) = M(\theta^*, g, b)$. It is then almost immediate to obtain the following proposition:

Proposition 9. *Atomistic monitors maximizing news are prone to a confirmation bias. They either choose to be purely friendly ($g(x) = r(x)$ and $b(x) = 0$) if $F(\theta^*) > \frac{1}{2}$, or they choose to be purely hostile if $F(\theta^*) < \frac{1}{2}$. They only choose different monitoring technologies*

[27] This is just for simplicity: any (smooth) aggregation of the $g(x)$'s such that g is not affected by a point-wise change yields the same qualitative insights, including the case with multiple reports of the same news item.

in the non-generic case where $F(\theta^*) = \frac{1}{2}$. The equilibrium never maximizes incentives for quality.

Proof. Substituting $b(x) = r(x) - g(x)$ into $g(x)F(\theta^*) + b(x)(1 - F(\theta^*))$, the objective function is linear in $g(x)$. The monitor hence chooses $g(x) = r(x)$ if $F(\theta^*) \geq \frac{1}{2}$ and $b(x) = r(x)$, otherwise. This implies that there are only extreme environments in equilibrium, and the best-reply of monitors involves only the two situations mentioned in the proposition. Note that multiple equilibria are possible. □

The confirmation logic of these monitors runs counter to that in Proposition 7. The monitor focuses on the likeliest quality, as this increases the probability of generating news. This proposition suggests that media firms primarily interested in producing evidence (e.g., news or rating agencies, but also small online information providers) will tend to prefer monitoring technologies which do not maximize quality incentives.

7 Conclusion

The very simple model developed here allows us to tackle an issue that has been extensively explored in the literature: the impact of asymmetric information on the quality supplied by sellers. We provide an original split of informational environments into friendly environments, in which high quality is more often identified by buyers than low quality, and hostile environments, in which the opposite holds.

We show that friendly or hostile informational environments differ remarkably. In particular, hostility produces a bandwagon effect across agents, which can generate multiple equilibria; this does not occur in friendly environments.

We also characterize the informational environment which maximizes quality incentives. Our general insight is that it is more effective to go against buyers' equilibrium beliefs. For instance, hostility produces greater incentives for quality than does friendliness when the cost of quality is low (which implies that buyers are optimistic regarding average quality). Similarly, friendliness is more effective in increasing quality when information is poor, as this means there are fewer incentives to provide quality and so beliefs are pessimistic. In a nutshell, disconfirmation is efficient.

These results shed new light on the literature on statistical discrimination and media bias, although many other applications are possible. One is the design of online feedback systems. Facebook, for example, forces the feedback environment to be friendly, by only allowing the emission of good news if the Like button is clicked, or no news if it is not. Including a third option is currently being debated. YouTube or Reddit, for example, do provide three options (good news, bad news, no news). It would be of interest to use our model to consider how the feedback system influences published content, possibly by adding heterogeneous tastes (which are absent in our model). Law and Economics is another area where informational control is complex and possibly asymmetric – evidence is disclosed by bodies which may be hostile (e.g., prosecutors), neutral (judges), or friendly (defense lawyers) – and where the understanding of how institutional design influences behavior is crucial.

Finally, in our setting, spot prices reflect all the information available ex-post, but no commitment is made ex-ante on this price. One area for future research is the analysis of solutions including contracts. The shape of optimal contracts in this setting will depend in an interesting way on the informational environment. Commitment also matters, with respect to how much information is disclosed. This is especially important in cases where the nature of the environment is rooted in the monitors' preferences: the ex-ante willingness to maximize incentives clashes with the ex-post incentives of friends, who prefer to withhold bad news ex-post. Both these concerns and the contractual solutions are left for future research.

A Appendix

A.1 Alternative payoff assumptions

For ease of presentation, we assumed here that the willingness to pay for quality is equal to the quality q, and that the support of costs is $[0, 1]$. Most of our results hold under more general assumptions. To check, assume that the willingness to pay is w_0 for $q = 0$, w_1 for $q = 1$, and the support of costs is $\left[\underline{\theta}, \overline{\theta}\right]$. It is easy to show that the results are unchanged if the willingness to pay for low quality is positive ($w_0 \geq 0$), so that the buyer always purchases the product. The incentive constraint simply becomes:

$$\theta \leq (g - (g - b)\mu)(w_1 - w_0).$$

where the difference between the willingness to pay for quality enters multiplicatively in the constraint. As we always have $g - (g - b)\mu < 1$, all sellers supply high quality and the nature of the informational environment is irrelevant if $\underline{\theta} < w_1 - w_0$. We obtain the symmetric result that all sellers supply high quality if the cost to all sellers is negative (that is, if $\overline{\theta} < 0$). Beyond these extreme cases, some agents may have costs so high that they never offer high quality ($\overline{\theta} > w_1 - w_0$). As such, we add a fixed mass of bad quality to the market, which systematically reduces μ in equilibrium, all else equal. Following the now familiar logic, this produces more incentives in a friendly environment (free-riding is reduced), and dampens the bandwagon effect in hostile environments. This effect reinforces the message of proposition 8. The opposite insight applies straightforwardly if some sellers always offer high quality because $\underline{\theta} < 0$. This again confirms the broad insights of proposition 8 that the more favorable the circumstances, the more efficient is hostility relative to friendliness.

Things become slightly more complicated when we introduce buyer-participation constraints and $w_0 < 0$. The buyer incurs a loss when consuming the low-quality product, and hence may not want to buy an unidentified product. Formally, participation is ensured if $\mu + (1 - \mu)w_0 \geq 0$: consumers will only buy if their perception μ that the product is good is high enough. The main difference with the scenario in the body of the paper is that lower

w_0 boosts incentives in friendly environment, since free-riding on average quality becomes less valuable, while it reduces incentives in a hostile environment via a reduced bandwagon effect. This can be seen from the modified incentive constraint:

$$\theta \leq g - (g - b) \cdot max(\mu + (1 - \mu)w_0, 0) \tag{10}$$

Note that if μ or w_0 is low enough, this constraint boils down to $\theta \leq g$. The following proposition characterizes these equilibria and can be compared to proposition 1.

Proposition 10. *For any (g, b), there exists $\underline{w}(g, b)$ such that*

1. *if $w_0 < \overline{w}(g, b)$, the equilibrium is unique, with $\theta^* = g$.*
2. *if $w_0 \geq \overline{w}(g, b)$:*
    - *In a friendly environment, the equilibrium is qualitatively the same as in proposition 1. The equilibrium threshold falls in w_0.*
    - *In a hostile environment, the equilibria with $F(\theta^*) \geq \frac{(1-b)w_0}{(1-b)w_0-(1-g)}$ are qualitatively the same as in proposition 1; otherwise they collapse to $\theta^* = g$. The equilibrium threshold of stable equilibria rises in w_0.*

Proof. Consider the case where the buyer's participation constraint does not bind. Then the analysis is almost identical to that in the body of the paper, with the function M being replaced by the more general:

$$\hat{M}(\theta|g, b) = \frac{(1 - b)(g - \theta - (g - b)w_0)}{(g - b)((1 - \theta) - (1 - b)w_0)} \tag{11}$$

This modified function has the same qualitative properties as the function M in Lemma 2, except that $\hat{M}(g-(g-b)w_0|g, b) = 0$ (i.e. the intersection of M with the horizontal axis now shifts by $(g - b)w_0$). Hence the highest equilibrium is such that $\theta^* \leq max(b, g - (g - b)w_0)$. Note that, from (3), the Bayesian belief μ consistent with some threshold θ always rises in θ. Hence, in hostile environments, the upper bound for an equilibrium belief is reached at $\theta = b$, associated with belief $\hat{\mu}$ strictly lower than 1; there hence exists a threshold $\overline{w}(g, b)$

such that $\hat{\mu} - (1-\hat{\mu})\overline{w}(g,b) = 0$. In a friendly environment, the unique intersection of F and $\hat{M}$ is interior, since $\hat{M}(1|g,b) < 1$. The corresponding belief is strictly lower than 1, hence the existence of a threshold $\overline{w}(g,b)$. In both cases, for any $w_0 \leq \overline{w}(g,b)$ the participation constraint of the buyer is not satisfied for an unidentified product (in any equilibrium); there hence exists a unique equilibrium, with $\theta^* = g$.

Conversely, in a friendly environment where the participation constraint does not bind, the unique equilibrium is qualitatively similar to that in the baseline scenario. In hostile environments, an equilibrium of the same form as that in Proposition 1 survives only if the corresponding belief is high enough so that the unidentified product is still bought, which applies under the given criterion.

Finally, the comparative statics on w_0 are immediate from the inspection of $\frac{\partial \hat{M}}{\partial w_0}$. □

The first point corresponds to an extreme case where an unidentified product is never bought. This completely shuts down the Bayesian externality across the population of agents. The second bullet point underlines that the other case is similar to the main case analyzed in the paper, but that a different equilibrium may result which does not feature a Bayesian externality. Shutting down the Bayesian externality is desirable in friendly environments, but not in hostile environments. Finally, note that under a constraint $g + b \leq r$, a friendly environment is better than a hostile environment when w_0 is sufficiently low (i.e. lower than $\overline{w}(g,b)$ for all $b < r$), in that hostile disclosure can even serve no purpose in this case, as the equilibrium depends only on g. Given the comparative statics on w_0 in the second point, it is also the case for non-constrained equilibria that reducing w_0 makes friendliness relatively more efficient than hostility. This reinforces the message of proposition 8 on environment design: the less favorable are the fundamentals (here the lower is w_0), the more friendliness is socially desirable.

A.2 Imperfect signals

The base model features two perfect signals: good and bad news. This arguably can be seen as a limitation for two reasons. First, quality signals received by buyers are often imperfect in reality, while still being informative. Second, this prevents discussion of the impact of signal precision. We here introduce imperfect signals into the model and underline the difficulties arising from this generalization.

Consider an information structure with the same ternary signal support $s \in \{q, \varnothing\}$. We introduce noise as indicated in the following table, where each entry represents the probability $p(s|q)$ of receiving signal s when the true quality is q:

	$q=1$	$q=0$
$s=1$	$g(1-\varepsilon_g)$	$b\varepsilon_b$
$s=\varnothing$	$1-g$	$1-b$
$s=0$	$g\varepsilon_g$	$b(1-\varepsilon_b)$

Imperfect signals here imply that both $q=0$ and $q=1$ can lead to all types of news. This generalizes the structure that appeared in the main body of the paper in which $(\varepsilon_g, \varepsilon_b) = (0,0)$. Note that the intermediate signal $\varnothing$ is unchanged.

We need additional assumptions to ensure that the two extreme signals capture the correct notion of good and bad news in this new context. Following Milgrom (1981), we impose that:

$$\frac{p(0|1)}{p(0|0)} \leq \frac{p(\varnothing|1)}{p(\varnothing|0)} \leq \frac{p(1|1)}{p(1|0)},$$

which we can rewrite as the following condition:

$$\frac{\varepsilon_g}{1-\varepsilon_b} \leq \frac{(1-g)b}{g(1-b)} \leq \frac{1-\varepsilon_g}{\varepsilon_b} \tag{12}$$

Clearly, these conditions are satisfied if the extreme signals are sufficiently close to perfect: When the errors are small enough, the LHS tends to zero and the RHS to infinity.

We now denote by $\mu(s)$ the consumer's belief on receiving signal s. The incentive constraint is $\sum_s [p(s|1) - p(s|0)]\mu(s) \geq \theta$, which indicates that the equilibria here are also cutoff equilibria. Expected quality conditional on a signal s with a cutoff θ is then $\mu(s) = \frac{p(s|1)F(\theta)}{p(s|1)F(\theta)+p(s|0)(1-F(\theta))}$. θ is $\mu(s) = \frac{p(s|1)F(\theta)}{p(s|1)F(\theta)+p(s|0)(1-F(\theta))}$. The incentive constraint can now be written as:

$$\theta \leq g\mu(1) - b\mu(0) - (g-b)\mu(\varnothing) - (g\varepsilon_g + b\varepsilon_b)(\mu(1) - \mu(0)) \tag{13}$$

This condition is worthy of a number of remarks. The first is that zero quality ($\theta^* = 0$) is always an equilibrium as long as $\varepsilon_b > 0$. The reason is that, for all s, $\mu(s) = 0$ when $\theta = 0$, in contrast to the perfect-signal case in which $\mu(1) = 1$ and $\mu(0) = 0$ for any value of θ. This equilibrium is conceptually very similar to the babbling equilibrium in a cheap-talk game. Any message can be interpreted as indicating $q = 0$, just as any message is interpreted as noise in a babbling cheap-talk equilibrium. There is thus no incentive to invest in any effort, leading to a self-fulfilling prophecy.[28] This has important implications for the design of information structure as this inefficient equilibrium disappears as soon as $\varepsilon_b = 0$.

A second important point is that introducing imperfect signals creates a last term in the incentive constraint (13) which reduce incentives.[29] This leads to the general, albeit straightforward, result that a structure with three imperfect signals is always less efficient than the structure of the base model. The first three terms in (13) are familiar from the perfect-signal case. In particular the coefficient on the posterior $\mu(\varnothing)$, which is the equivalent of μ in the base model, is exactly the same, $-(g-b)$.

The last term features, in general, a coefficient of the form $F(\theta)(1 - F(\theta))$, which introduces a non-monotonicity that significantly complicates the analysis.[30] In order to illustrate the problem, consider the simple example of a neutral structure with $m = g = b$, but asymmetric errors.[31] An equilibrium cutoff is a solution to:

$$\theta = \frac{m(1-\varepsilon_g-\varepsilon_b)^2 F(\theta)(1-F(\theta))}{((1-\varepsilon_g)F(\theta) + \varepsilon_b(1-F(\theta)))(\varepsilon_g F(\theta) + (1-\varepsilon_b)(1-F(\theta)))}.$$

[28] A zero-quality equilibrium also always exists in a purely hostile environment with perfect signals as no news can also be interpreted as $q = 0$.

[29] $\mu(1) - \mu(0)$ is always positive given the assumption on signal ordering.

[30] This is similar to the case in Board and Meyer-ter-Vehn (2013) when dealing with imperfect signals.

[31] This information structure with errors generalizes that in Gill and Sgroi (2012), by adding an inconclusive (no news) result to their binary test. Their condition (1) amounts in our setting to exactly imposing $g = b = m$, as can be easily seen by substituting the relevant $p(s|q)$. The toughness of the test τ in their definition (2) amounts in our setting to $\tau = m.\varepsilon_g$. Combining (1) and (2) leads to $m.(\varepsilon_b + \varepsilon_g) = 1 - \kappa$, where κ is what they call the expertise of the test.

It is easy to see that the equilibrium might not be unique even in this environment. One source of multiplicity comes from the good signal being polluted by bad quality (while the opposite error does not play an equivalent role). In particular, uniqueness prevails in the neutral environment as long as $\varepsilon_b = 0$. In this case the equilibrium condition is $\theta = \frac{m(1-\varepsilon_g)(1-F(\theta))}{1-F(\theta)+\varepsilon_g F(\theta)}$, the RHS of which is increasing in θ. Conversely, if $\varepsilon_g = 0$, the equilibrium satisfies $\theta = \frac{m(1-\varepsilon_b)F(\theta)}{(1-\varepsilon_b)F(\theta)+\varepsilon_b}$, and multiple equilibria are possible because the RHS now increases in θ. Bearing in mind that reducing the error ε_g actually makes the *bad* signal more precise, this analysis mirrors the friendly vs. hostile contrast highlighted in the base model where multiplicity arises when more evidence is available on bad rather than good quality.

A.3 Proof of Lemma 2

Note that $\frac{\partial M}{\partial \theta} = -\frac{(1-b)(1-g)}{(g-b)(1-\theta)^2}$ and $\frac{\partial^2 M}{\partial \theta^2} = -\frac{2(1-b)(1-g)}{(g-b)(1-\theta)^3}$. Both derivatives are negative (positive) when $g < b$ ($g > b$). Furthermore, $M(0|g,b) = \frac{g(1-b)}{(g-b)}$, which is higher than $F(0) = 0$ if $g > b$ and lower otherwise. $M(.|g,b) \to \infty$ (resp. $-\infty$) when $c \to 1$ and $g > b$ ($g < b$). When $g = b$, $c = b = g$ means that the disclosure curve becomes a vertical line.

A.4 Equilibrium uniqueness with a concave distribution

We here establish that a sufficient condition for the equilibrium to be unique for any information structure with $g > 0$ is that the cost distribution be concave. According to proposition 1, the equilibrium is unique in a friendly environment. In hostile environments, the disclosure curve is convex in θ and $M(0|g,b) > 0$ when $g > 0$, so that when F is concave the intersection of M and F is unique. Note that when $g = 0$, $\theta^* = 0$ is always an equilibrium; there is another equilibrium when $f(0) > \frac{\partial M(0|0,b)}{\partial \theta}$, since then the distribution is above the disclosure curve around $\theta = 0$, but below it for θ close to 1. The non-generic case $g = 0$ is specific to the support of costs starting exactly at 0, and the corresponding equilibrium with $\theta^* = 0$ disappears for any arbitrarily small g.

A.5 Proof of Proposition 8

We prove the points of the proposition in turn. Note that since F is concave the equilibrium is unique, up to the unstable $\theta^* = 0$ equilibrium in the limiting case $g = 0$. Since for an arbitrarily small g this equilibrium disappears, we will ignore it, keeping in mind that a strict $g = 0$ could be understood as arbitrarily small. Figure 7 is meant to illustrate the argument of the proof (note that it represents a case where a fully hostile environment is optimal).

We first show that for any given $y \in [0, 1]$, the highest θ such that $y = M(\theta, g, b)$ and $g+b = r$ is found with a disclosure curve with either $g = 0$ or $g = r$. Consider the inverse (in θ) of the disclosure curve: $M^{-1}(y|g, b) = \frac{(g-b)y-g(1-b)}{(g-b)y-(1-b)}$. Along the budget constraint, we aim to maximize $M^{-1}(y, g, r-g)$. But $\frac{d^2}{dg^2} M^{-1}(y|g, r-g) = \frac{(2(2-r)^2 y(1-y))}{(1+g-r+y(r-2g))^3)}$, which is always positive since $0 \leq y \leq 1$ and $0 \leq g \leq r < 1$. Hence M^{-1} is convex, and its maximum is attained for an extreme g for any given y. The maximal equilibrium $\theta^*(g, r-g)$ has to be on a disclosure curve, by the definition of M. Suppose that this is not on an extreme disclosure curve. There then exists an extreme disclosure curve which, as we have seen, is to the right of $\theta^*(g, r-g)$. Then, since F is increasing, the intersection of F and this extreme disclosure curve has to be to the right of $\theta^*(g, r-g)$, which implies that the corresponding equilibrium has higher average quality. This proves the first point: any optimal equilibrium has to be on one of the two extreme curves.

Second, consider the point of intersection of these two extreme disclosure curves. In terms of θ, this is the solution to $M(\theta|r, 0) = M(\theta|0, r)$. Since the first disclosure curve is decreasing and the second is increasing, the intersection is unique, i.e. there exists a unique $\hat{\theta}$ solving the equation, which is $\hat{\theta} = \frac{r}{2-r}$. This implies that $M^{-1}(y|r, 0)$ is higher than $M^{-1}(y|0, r)$ if and only if $y \leq F(\hat{\theta})$. As such, the intersection of F with an extreme disclosure curve occurs either with $M(\theta|r, 0)$ if $F(\hat{\theta})$ or with $M(\theta|0, r)$ otherwise, which yields the criterion in the second point.

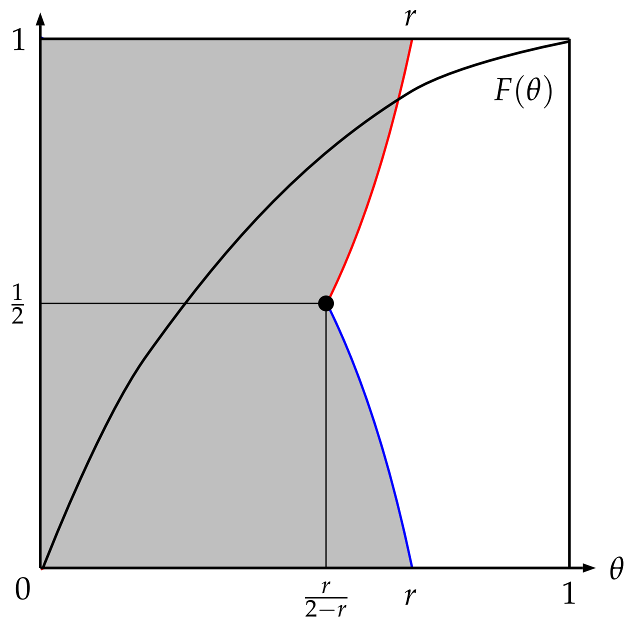

Figure 7: Illustration of the proof of proposition 8.

References

Abreu, D., Milgrom, P. and Pearce, D., 1991. "Information and Timing in Repeated Partnerships." *Econometrica* 59(6):1713-1733.

Akerlof, G., 1970. "The Market for Lemons: 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.

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

Baron, D. P., 2006. "Persistent media bias." *Journal of Public Economics* 90: 1-36.

Besley, T. and Prat, A., 2006. "Handcuffs for the Grabbing Hand? Media Capture and Government Accountability." *American Economic Review* 96(3):720-736.

Board, S. and Meyer-ter-Vehn, M., 2013. "Reputation for Quality." *Econometrica* 81(6):2381-2462.

Cabral, L. and Hortaçsu, A., 2010. "The Dynamics of Seller Reputation: Evidence from eBay." *Journal of Industrial Economics* 58:54-78.

Cai, H., Jin, G., Liu, C. and Zhou, L. 2014. "Seller Reputation: From Word-of-mouth to Centralized Feedback." *International Journal of Industrial Organization* 34:51-65.

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

Dellarocas, C. and Wood, C., 2008. "The Sound of Silence in Online Feedback: Estimating Trading Risks in the Presence of Reporting Bias." *Management Science* 54(3):460-476.

Dewatripont, M. and Tirole., 1999. "Advocates." *Journal of Political Economy* 107(1):1-39.

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

Duggan, J. and Martinelli, C., 2011. "A Spatial Theory of Media Slant and Voter Choice." *Review of Economic Studies* 78:640-666.

Gentzkow, M. and Kamenica, E., 2014. "Costly Persuasion." *American Economic Review (Papers and Proceeding)* 104(5):457-462.

Gentzkow, M. and Shapiro, J., 2006. "Reputation and Media Bias." *Journal of Political Economy* 114(2):280-316.

Gentzkow, M. and Shapiro, J., 2008. "Competition and Truth in the Market for News." *Journal of Economic Perspectives* 22(2):133-154.

Gill, D. and Sgroi, D., 2012. "The Optimal Choice of Pre-Launch Reviewer." *Journal of Economic Theory* 147:1247-1260.

Groseclose, T. and Milyo, J., 2005. "A Measure of Media Bias." *Quarterly Journal of Economics* 120(4):1191-1237.

Harbaugh, R., Maxwell, J.W. and Roussillon, B., 2011. "Label Confusion: The Groucho Effect of Uncertain Standards." *Management Science* 57(9):1512-1527.

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.

Johnson, J. P. and Myatt, D.P., 2006. "On the Simple Economics of Advertising, Marketing, and Product Design." *American Economic Review* 96(3):756-784.

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

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

Kim, S. K., 1995. "Efficiency of an Information System in an Agency Model." *Econometrica* 63(1):89-102.

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

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.

Lewis, T.R.M. and Sappington, D.E.M.:, 1994. "Supplying information to facilitate price discrimination." *International Economic Review*, 35:309-327.

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

Lyon, T. (Ed.), 2010. Good Cop/Bad Cop: Environmental NGOs and Their Strategies towards Business, Resources for the Future Press, Washington, DC.

MacKay, D. J. C., 2003. Information Theory, Inference, and Learning Algorithms, Cambridge: Cambridge University Press.

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

Martin, D., 2012. "Strategic Pricing with Rational Inattention to Quality." mimeo.

Mason, C. F., 2011. "Eco-Labeling and Market Equilibria with Noisy Certification Tests." *Environmental and Resource Economics* 48:537-560

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

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

Mullainathan, S. and Shleifer, A., 2005. "The Market for News." *American Economic Review* 95(4):1031-1053.

Ottaviani, M. and Prat., A., 2001. "The value of public information in monopoly." *Econometrica* 69(6):1673-1683.

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

Shin, H.S., 1998. "Adversarial and Inquisitorial Procedures in Arbitration." *RAND Journal of Economics* 29:378-405.

Sims, C., 2003. "Implications of Rational Inattention." *Journal of Monetary Economics*, 50(3), 665-690.

Snyder, J.M. and Strömberg, D., 2010. "Press Coverage and Political Accountability." *Journal of Political Economy* 118(2):355-408.

Strausz, R., 2005. "Honest Certification and the Threat of Capture." *International Journal of Industrial Organization* 23:45-62.

Taylor, C. and Yildirim, H., 2011. "Subjective Performance and the Value of Blind Evaluation." *Review of Economic Studies* 78: 762-794.

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.