BDPEMS
Berlin Doctoral Program in Economics and Management Science

Working Paper Series
2013-01

Informational opacity and honest certification

April 2013

Martin Pollrich, Humboldt-Universität zu Berlin

Lilo Wagner, DIW Berlin

Sponsored by

EINSTEIN Foundation.de
Einstein Stiftung Berlin
Einstein Foundation Berlin

BDPEMS
Spandauer Straße 1
10099 Berlin
Tel.: +49 (0)30 2093 5780
mail: bdpems@hu-berlin.de
http://www.bdpems.de

Informational opacity and honest certification

Martin Pollrich* and Lilo Wagner†

April, 2013

Abstract

This paper studies the interaction of information disclosure and reputational concerns in certification markets. We argue that by revealing less precise information a certifier reduces the threat of capture. Opaque disclosure rules may reduce profits but also constrain feasible bribes. For large discount factors a certifier is unconstrained in the choice of a disclosure rule and full disclosure maximizes profits. For intermediate discount factors, only less precise, such as noisy, disclosure rules are implementable. Our results suggest that contrary to the common view, coarse disclosure may be socially desirable. A ban may provoke market failure especially in industries where certifier reputational rents are low.

Keywords: Certification; Bribery; Reputation

JEL Classification Numbers: L15; D82; L14; L11

*Humboldt-Universität zu Berlin, *martin.pollrich@hu-berlin.de*

†DIW Berlin, *lwagner@diw.de*

We would like to thank Pio Baake and Roland Strausz for insightful comments and discussions. We thank Helmut Bester, Renato Gomez, Alessandro Lizzeri, Patrick Rey and seminar participants in Berlin, Mannheim and Toulouse for helpful comments and advice. Martin Pollrich gratefully acknowledges financial support by ANR and DFG through project "Market Power in Vertically Related Markets" and by DFG through SFB/TR 15.

1 Introduction

In markets that exhibit informational asymmetries, product quality is typically reduced. This in turn may provoke a breakdown of trade. The lack of credible communication between informed and uninformed parties may result in the emergence of certification intermediaries. Certifiers inspect products whose characteristics are private information to agents, and publicly reveal this information. Examples abound: rating agencies, eco-labels, wine certificates or technical inspections. Often however, results are revealed on a coarse scale, although the information at hand would allow for a more precise disclosure. We provide a new explanation for such opaqueness. We show that partially revealing rules can serve as a safeguard against fraud: certifiers may be tempted to accept bribes for releasing favorable certificates. This behavior, which we call capture, enables the certifier to extract payments other than the certification fee. If consumers are aware of this threat of capture, then the certifier must find ways to credibly commit to honesty. We show that one way to do so is to employ an opaque disclosure rule. Opaqueness reduces the producer's willingness to pay for bribery, because a more opaque disclosure rule lowers differences in the values of certificates. Hence, opaqueness can be welfare enhancing since it may prevent market failure. This result is surprising because it contradicts the commonly held view that reducing informational asymmetries is socially desirable per se. We show our result in a model with moral hazard where, in each period, short-lived producers first have to make an investment choice, which in turn determines the probability distribution of their product's quality. Thus, the payoffs assigned to each quality outcome determine the incentives to invest. The long-lived certifier has two instruments at his disposal: a flat certification fee and the disclosure rule. Consumers experience the true quality of a product only after consumption. If it does not match the awarded certificate, capture is detected. This makes the certifier face a classical reputation dilemma because he trades off short-run gains from capture against regular future profits.

We characterize feasible disclosure rules in this setting. Our major finding is that for sufficiently low discount factors, honest certification requires partial disclosure of quality information, which in our model implies noisy disclosure. In the short run, a certifier may gain from making a capture offer that is acceptable for at least some producers. The maximum producer willingness to pay for bribes is directly affected by the publicly announced disclosure rule. It is greatest for full disclosure and can be substantially reduced by revealing less precise information. But if consumers detect a bribe and therefore lose trust, a certifier gives up his future profits. Static certifier profits are maximal for full disclosure and any deviation will typically reduce the long-run loss from losing credibility. As will be shown, the first effect exceeds the latter.

We moreover obtain the counterintuitive result that a threat of capture increases social welfare.[1] Whenever information is fully revealed, sharing profits necessarily reduces producer investments as compared to the first-best level, obtained under complete information. We show that whenever capture offers are made before a certifier observes the true quality level, these are such that they are accepted by either all producers or only by low quality producers. If the highest threat of capture stems from offers that are accepted by all producers and the disclosure rule is noisy, credibility can be maintained by making honest certification more attractive to high quality producers. This in turn increases equilibrium investment levels as compared to full information disclosure.

Related literature. A stream of literature seeks to explain why certifiers often choose to only partially reveal quality information. Lizzeri (1999) finds that it is optimal for a monopolistic certifier in a static adverse selection environment to reveal almost no information. In this setting, this result is robust to introducing capture because a no revelation policy simply annihilates producer incentives to bribe. In the presence of moral hazard however, information revelation is necessary to create incentives for the provision of quality. Albano and Lizzeri (2001) study optimal disclosure rules in a static model of both moral hazard and adverse selection. In their setting, a certifier chooses to employ noisy disclosure if his set of actions is restricted to flat fees. According to Farhi et al. (2012), opaqueness in certification markets is caused by information averse sellers. In Dubey and Geanakoplos (2010), it is shown that coarse grading schemes can help to

[1] We analyze a belief system that substantially restricts the set of feasible disclosure rules. For different belief systems and sufficiently low discount factors, other (opaque) rules may be chosen by the certifier. The effect on social welfare is therefore not a general result.

induce all students to employ effort if they are disparate and care about their status in class. Kartasheva and Yilmaz (2012) explain imprecise ratings in a model with partially informed investors and heterogeneous liquidity needs of issuers. A static adverse selection model where quality is not fully observable by the seller is analyzed by Faure-Grimaud et al. (2009). They identify conditions under which the ownership of certification results is left to firms and under which firms reveal their ratings.

The threat of capture in certification markets has been analyzed by Strausz (2005). In a pure adverse selection setting with full disclosure, he analyzes the effects of a threat of capture on certification prices. He finds that in order to maintain credibility, for low discount factors, a certifier raises fees above the static monopoly price. This result is consistent with our finding in that as less information is disclosed, the certification fee generates a cut-off value that specifies a minimal certified producer quality. A larger fee increases this cut-off but this implies that less information is revealed in equilibrium. Although this effect is also present in Strausz (2005), he however does not explicitely point it out. As in the present paper, credibility is maintained by reducing the maximal willingness to bribe. In Strausz (2005), this is affected by the value of not being certified, which, in turn, is an increasing function of the certification fee.

There is a rich literature on reputation building in markets with informational asymmetries. For example, Shapiro (1983) analyzes the forces at work when sellers build reputation. Biglaiser (1993) investigates the role of market intermediaries when sellers are unable to build their own reputation. Examples of works that treat reputational concerns of rating agencies are Mathis et al. (2009) and Bolton et al. (2012). In contrast to the present paper, these works follow the asymmetric information approach to reputations, where certifiers are assumed to always be committed (i.e. honest) with positive probability.[2] This, however, restricts the set of allowed certification fees and disclosure rules for non-committed certifiers. The reason is that a departure from the equilibrium strategy immediately reveals the certifier type. Instead of assuming that testing by the certifier is imperfect as is done in those works, we show how imperfect testing may endogenously

[2] See Mailath and Samuelson (2012, Chapter IV) for a discussion of this approach.

arise in equilibrium.

Levin (2003) extends the standard moral hazard setting to situations where contractual agreements are enforceable only to a certain degree and where reciprocal relations are long-term. The optimal contract derived by Levin has a coarse structure, which parallels our finding of coarse disclosure being optimal.

The remainder of the paper is organized as follows. Section 2 presents the model. Section 3 analyzes the static game in the absence of bribery. In section 4, we treat the general case of certification under the threat of capture. Section 5 concludes. All proofs are presented in the appendix.

2 The setup

We consider a dynamic framework in discrete time. In each period $t = 1, 2, \ldots, \infty$, a short-lived monopolistic producer is born. He produces a single unit of quality $q_t \in \{q^l, q^h\}$, where $0 \leq q^l < q^h$. In the following, we refer to a *high type* producer if his product quality is q^h and to a *low type* producer otherwise. Prior to production, a producer chooses some investment level $e_t \in [0, 1]$. Quality is stochastic and the probability of the produced good being of high quality q^h is given by $Prob(q_t = q^h | e_t) = e_t$. This probability function is independent of t, i.e. quality levels are independent across time. Investment costs are given by the function $k(\cdot)$. We assume $k(\cdot)$ to be increasing and strictly convex. For technical reasons we assume a non-negative third derivative, so that the certifier's profit function is concave and to guarantee interior solutions we additionally assume $k'(0) \leq q^l$ and $k'(1) \geq q^h$.

Consumers' reservation prices equal (expected) qualities. Both investment and quality level are private information to the producer. Consumers observe the product quality only after consumption. All other components of the model are common knowledge. The equilibrium concept we use throughout the paper is that of perfect Bayesian equilibrium. Each producer is short-lived and leaves the market at the end of a period. Goods are

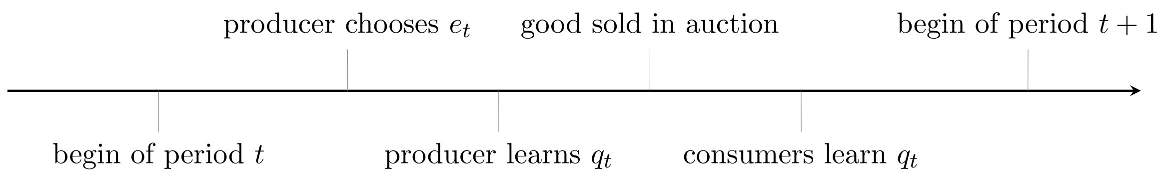

Figure 1: Timing in one period without certification

sold in a second-price auction.[3] Figure 1 summarizes the timing in period t.

To simplify notation, we set $q^l \equiv 0$ and define $v := q^h - q^l$. In the benchmark of *complete information* high quality goods are sold in the second-price auction at price v and low quality goods are sold at price 0. The producer then chooses e to maximize expected profits $ev - k(e)$. The first-best investment level e^* is thus given by $k'(e^*) = v$, which lies in the interval $[0, 1]$ due to our assumption $k'(1) \geq v$. In particular we have $e^* > 0$.

Under *asymmetric* information and in the absence of any further economic institution, a producer cannot persuade consumers that he offers a high quality good and the market price can therefore not be made contingent on a good's quality. It is standard to show that the Perfect Bayesian market outcome involves a market breakdown. In such an outcome, consumers form a belief q_t^e about the offered quality, which reflects their willingness to pay. In equilibrium, this belief has to be consistent with the actual expected quality, $E(q_t|e_t)$. Given any belief, the producer's optimal choice of investment will be $e_t = 0$, as he maximizes $q_t^e - k(e_t)$. But since $E(q_t|0) = 0$, the unique equilibrium has producers choosing $e_t = 0$ in every period and the quality of the good is zero in each period. The result is a market failure: high quality is never offered in equilibrium. We summarize this finding in the following lemma.

Lemma 1. *Without certification, producers choose $e_t = 0$ in each period. In equilibrium, quality is given by $q_t = 0$ and the price is 0 in each period.*

This inefficiency calls for the emergence of alternative market institutions to facilitate supply of high quality. The focus of this paper lies on certification as one such institution. Assume that an infinitely long-lived certifier enters the market. She offers to disclose the

[3]The second price auction results in a standard monopoly price that equals consumers' valuations. It circumvents signaling issues, e.g. letting the informed party take a publicly observed action that might be interpreted as a signal.

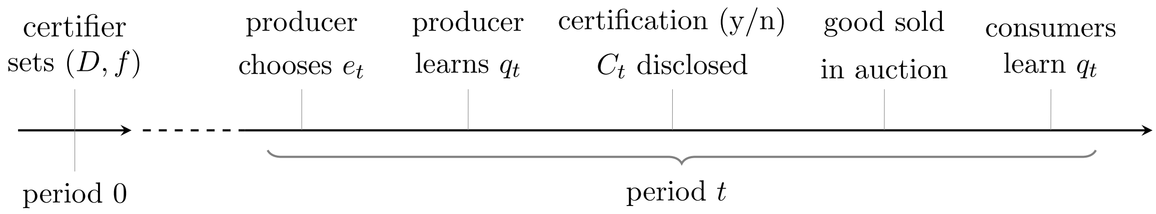

Figure 2: Timing with certification

result of some potentially imperfect test of the good's quality, prior to it being sold. More precisely, at the beginning of the game, in period $t = 0$, the certifier announces a fee $f \geq 0$ and a disclosure rule $D = (\mathcal{C}, \alpha^l, \alpha^h)$. The fee is to be paid by any producer who wishes to have his product tested. The disclosure rule consists of a set $\mathcal{C} = \{C^1, \ldots, C^m\}$ of potential certificates and probability vectors α^l and α^h, where the k-th entry of vector α^i reflects the probability that a product of quality q^i is awarded certificate C^k whenever tested. We do not assume that those probabilities add up to one, i.e. we allow for $\sum_{k=1}^{m} \alpha_k^i < 1$. Hence, a product may remain uncertified with the conforming probability and will be sold as such. We assume that consumers cannot observe whether a product was tested, unless it is offered with a certificate.[4] Possible disclosure rules encompass for example *full disclosure*, where $\mathcal{C} = \{q^l, q^h\}$ and $\alpha^h = (0, 1)$ as well as $\alpha^l = (1, 0)$, or *no disclosure*, where $\mathcal{C} = \{C\}$ and $\alpha^i = (1)$. Finally we assume that the certifier's inspection costs are zero[5] and that she discounts future profits at rate $\delta \in (0, 1)$. Figure 2 illustrates the timing of the game with certification.

An interpretation of the disclosure rule, which we shall use throughout the paper, is the following: the certifier can create any test that leads to a grading scheme with grades from the set $\mathcal{C}$ and results in the respective grades with conforming probabilities. This may be done with a computer program or a statistical test. In particular, after the test result is obtained, the certifier and the consumers share the same beliefs on product quality.

[4] Hence products which "failed" the test are sold under the same label as products that didn't even take the test. This assumption is not crucial, since the certifier can replicate any outcome of a game where consumers are able to observe whether a product applied for certification.

[5] This assumption simplifies the analysis without substantially affecting the results, which continue to hold as presented here for small but strictly positive inspection costs. Large inspection costs leave most of our results still valid, but create cumbersome case distinctions.

3 Optimal honest certification

In this section, we analyze certifier equilibrium strategies when the certifier is honest. By the stationary structure of the model, we can restrict our analysis to the certifier decision (D, f) plus a single period of production. Let $\pi^D(f)$ denote the equilibrium profit of the certifier, when adopting disclosure rule D with certification fee f.

We first study the case of full disclosure in some detail, as it will turn out that this disclosure rule can be used to achieve maximal profits. Under full disclosure, any product certified as high quality is sold at a price v, whereas a certificate of low quality makes the product worthless to consumers. Assume that consumers believe that uncertified products are of low quality with certainty. Therefore any high quality producer is willing to get his product certified, whereas low quality producers prefer to sell their product as uncertified.[6] A producer chooses his investment according to

$$e = arg \max_{\widetilde{e}} \ \widetilde{e} \cdot (v - f) - k(\widetilde{e}). \tag{3.1}$$

This implies $k'(e) = v - f$ in equilibrium and certifier expected equilibrium profits can be expressed as

$$\widehat{\pi}^{FD}(e) = e \cdot (v - k'(e)). \tag{3.2}$$

Denote e^{FD} the equilibrium effort level under a full disclosure rule and f^{FD} the respective fee that maximizes certifier profits under full disclosure. The following lemma proves that these values do exist and are unique.

Lemma 2. *Under full disclosure, there exists a unique fee f^{FD} that maximizes certifier profits. The uniquely defined equilibrium investment level e^{FD} is implicitly given by*

$$k''(e^{FD}) \cdot e^{FD} = v - k'(e^{FD}). \tag{3.3}$$

[6] Given consumers' beliefs, a high quality producer has a strict preference for certification whenever $f < v$. Low quality producers strictly prefer remaining uncertified for any $f > 0$. It is easy to see that certification fees $f \geq v$ and $f = 0$ yield zero profits to the certifier and we consequently restrict our analysis to $f \in (0, v)$.

The fee is $f^{FD} = v - k'(e^{FD})$ *and the (subgame-) equilibrium profit is* $\pi^{FD} = e^{FD} \cdot f^{FD}$.

We continue analyzing general disclosure rules. The entire set of disclosure rules is complex and difficult to handle analytically. A closer look at equation (3.2), which allows us to express the certifier profit as function of the implemented investment level e, points to the advantages of using an indirect approach. We can express the attained profit of any certifier policy (D, f) solely in terms of the induced investment level e. This allows for a straightforward comparison of attained profits and leads us to the following proposition.

Proposition 1. *For any disclosure rule* $D = \left(\mathcal{C}, \alpha^1, \alpha^0\right)$ *and any fee* $f \geq 0$*, it holds that* $\pi^D(f) \leq \pi^{FD}$ *in equilibrium.*

Proposition 1 states that the certifier will always find it optimal to employ a full disclosure rule. The reason is that, investment incentives depend on the difference between payoffs from selling high and low quality products. Given full disclosure, the certification fee is sufficient to fully control this difference.

We conclude this section with pointing out that full disclosure is not the unique disclosure rule that yields the maximal certifier profit π^{FD}. One example of such a disclosure rule is the following: The certifier issues two different certificates C^1 and C^2. Low quality products are only eligible for certificate C^2, hence $\alpha^l = (0, 1)$. High quality products receive certificate C^1 with probability $\alpha \in (0, 1)$ and C^2 otherwise, therefore $\alpha^h = (\alpha, 1 - \alpha)$. With this structure, it is possible to sustain an equilibrium in which all producer types demand certification. The optimal certifier profit π^{FD} is then obtained by choosing f and α appropriately.[7]

Disclosure rules of this kind play a crucial rule for the remainder of this paper. We henceforth refer to them as *partial disclosure* rules.

4 The capture problem

So far we assume that the certifier sticks to the announced disclosure rule, in particular that she conducts the lottery honestly and grants the respective certificate. However,

[7] We formally show this in the proof of Proposition 6.

there is pressure from producers who wish to be awarded better certificates. For instance, if disclosure is meant to be noisy, a certifier might be willing to guarantee a producer a high value certificate in exchange for a bribe. In this section we address this issue by formally introducing the possibility of capture.

We follow Strausz (2005) in modeling the possibility of capture, using the framework of enforceable capture as initiated by Tirole (1986). Hence we assume that the certifier and the producer can write an enforceable side-contract with transfers. Consumers are fully aware of the possibility of these side-contracts, but cannot observe them. Specifically, we model capture as follows: after a producer has learned his type q_t, the certifier, without observing q_t, may make an offer (C, b) to the producer. The offer consists of a certificate C, issued in case of acceptance, and a financial transfer b to be paid by the producer. The certifier thus offers to "sell" the sure certificate C at the price b, circumventing the customary certification procedure given by the disclosure rule. Hence, (C, b) are the terms at which she is willing to become captured. A producer however can reject this offer and, if willing to do so, insist on honest certification by paying the fee f. This last assumption is motivated following Kofman and Lawarrée (1993) in assuming that the certifier cannot forge certification without the help of the producer.

Note that the choice of the disclosure rule puts some limits on the set of *feasible* capture offers. For a general disclosure rule $D = \{\mathcal{C}, \alpha\}$ only offers of the form (C, b) with $C \in \mathcal{C}$ are feasible.[8] With full disclosure this means offers of the type (q, b) with $q \in \{q^l, q^h\}$.

Within the framework presented here, capture may subvert honest certification for two reasons. First, producers with low quality products are willing to side-contract with the certifier in order to obtain better certification. Second, high types may want to avoid uncertainty if disclosure is noisy.

In this section we are interested in the existence and characterization of equilibria where the certifier resists the temptation of making any capture offer of the above described kind. Throughout, we will work with different specifications of trigger beliefs. This

[8] This will be made more precise when formally introducing consumer beliefs. Granting a certificate which is not contained in D is certainly perceived as cheating by consumers. Consequently consumers believe to be faced with a worthless product and they will lose trust in the certifier's credibility.

becomes necessary as the ability of consumers to detect capture varies across disclosure rules. We assume consumers are able to perfectly observe quality after consumption. Therefore, if D is full disclosure or if certain certificates are awarded exclusively to high types, capture detection is also perfect.

Our particular idea behind the consumers' beliefs is the following: They stop trusting the certifier immediately if a false testimony about a product's quality is detected. Then, producers are not willing to pay for certification anymore. Consequently the certifier will lose future demand and makes zero profits henceforth. This prevents the certifier from becoming captured in the first place. We shall make this more precise in the following subsections.

4.1 Capture under full disclosure

Because, by Proposition 1, a certifier would want to employ full disclosure whenever possible, we start by investigating capture under a full disclosure rule. We assume that consumers trust certificates as long as they have not detected a deviation. A certifier who anticipates this behavior may be prevented from succumbing to the temptation of becoming captured by the fact that losing credibility will leave her without demand in future periods.

Denote $h_t = (n_t, q_t^c, q_t)$ the certification outcome in period t, where $n_t \in \{0, 1\}$ indicates whether certification in period t took place, q_t^c is the testified quality level in period t and q_t is the true quality observed after consumption.[9] If certification in period t did not take place, then $n_t = 0$ and $q_t^c = \emptyset$. Now let $H_t = (h_1, \ldots, h_{t-1})$ summarize the history of certification at the beginning of period t. Finally, we denote $q_t^e(n_t, q_t^c, H_t)$ a consumer's belief in period t when faced with a product certified as being of quality q_t^c and when having observed history H_t. The following assumption on consumers' beliefs comprehends the described behavior.[10]

[9] This specification of a history and Assumption 1 are particular for the case of full disclosure. With noisy disclosure one has to replace q_t^c by the disclosed certificate and the trigger beliefs have to be adjusted, since there will be cases where consumers are uncertain whether capture occurred. We make this precise in section 4.2.

[10] Note that consumers do not lose trust in the certifier when a product is awarded a low certificate, although this should not happen in equilibrium. It is not necessary to include this case into consumers'

Assumption 1. *The consumers' beliefs* $q_t^e(n_t, q_t^c, H_t)$ *satisfy* $q_t^e(1, q_t^c, H_t) = q_t^c$ *whenever* $\{\tau < t | n_\tau = 1 \ \wedge \ q_t^c \neq q_t\} = \emptyset$*. Moreover* $q_t^e(1, q_t^c, H_t) = 0$ *whenever* $\{\tau < t | n_\tau = 1 \ \wedge \ q_t^c \neq q_t\} \neq \emptyset$ *and* $q_t^e(0, \emptyset, H_t) = 0$*.*

The assumption states that consumers trust the certifier if capture was not observed in the past. They however lose trust forever, once they detected cheating. Losing trust implies that consumers believe for any certifier's claim that the offered quality is zero. With full disclosure, there are (at most) two types of bribing offers that can be made: (q^1, b) and (q^0, b). Obviously, an offer (q^0, b) is turned down by all types of producers, as it is worth nothing. Hence, in the following we focus on offers (q^1, b) and talk of a bribe b rather than (q^1, b). An offer b is accepted by high producer types whenever $b < f$. Low quality producers accept any bribe $b < v$ because acceptance will yield positive profits compared to zero profits for rejection. In equilibrium, the certifier assigns probability $e(f)$ to the event that a producer is of high type, where $e(f)$ is the producer's optimal investment under full disclosure, derived from (3.1). We are interested in equilibria where capture does not occur. In all such equilibria, a producer chooses his optimal investment level knowing that he will not receive an acceptable capture offer. The acceptance probability $p(b|f)$ of bribing offer b given the certification fee f is given by

$$p(b|f) = \begin{cases} 1, & \text{if } b < f, \\ 1 - e(f), & \text{if } f \leq b < v, \\ 0, & \text{if } b \geq v. \end{cases} \tag{4.1}$$

We denote by $\Pi^D(f) = \sum_{t=1}^{\infty} \delta^{t-1} \pi^D(f) = \pi^D(f)/(1-\delta)$ the certifier's expected profit from honest certification under disclosure rule D and fee f. The certifier's expected profit from offering bribe b is denoted by $\widehat{\Pi}^D(b|f)$ and depends on whether the consumer detected capture as follows: whenever $b < f$, all producer types will accept the bribe, but only for low quality producers this is detected. Hence, $\widehat{\Pi}^{FD}(b|f) = b + e(f)\delta\Pi^{FD}(f)$. For $f \leq b < v$, only low quality producers accept the bribe and $\widehat{\Pi}^{FD}(b|f) = (1-e(f))b+$

beliefs, because any such event can only follow a non-profitable deviation by the certifier.

$e(f)(f + \delta\Pi^{FD}(f))$. Whenever $b \geq v$, all producers reject the bribe and the certifier obtains $\widehat{\Pi}^{FD}(b|f) = \Pi^{FD}(f)$.

If $\widehat{\Pi}^{FD}(b|f)$ exceeds $\Pi^{FD}(f)$, the certifier is actually better off becoming captured with the associated probability $p(b|f)$. We say that certification at a fee f is *capture proof* if and only if

$$\Pi^{FD}(f) \geq \widehat{\Pi}^{FD}(b|f) \tag{4.2}$$

for all b. Note that $\widehat{\Pi}^{FD}(b|f)$ is increasing in b, both on $[0, f)$ and $[f, v)$ and it is constant for $b \geq v$. Furthermore $\widehat{\Pi}^{FD}(\cdot|f)$ is continuous at $b = f$.[11] Therefore, certifier profits from bribery are largest when b approaches v. Evaluating this yields the following proposition:

Proposition 2. *If D is full disclosure, an equilibrium satisfying Assumption 1 is capture proof. It exists if and only if*

$$\delta \geq \delta^{FD}(f) \equiv \frac{v}{v + \pi^{FD}(f)} \tag{4.3}$$

The proposition highlights the crucial role the discount factor plays for the existence of honest, i.e. capture proof, equilibria: the critical discount factor determines the relative weights of the short run gain - the bribe b - and the long run loss of capture - foregone future profits from certification. To see this, note that all bribes $b < v$ are accepted with some positive probability and therefore, the largest possible short-run gain equals v. In the long run, a certifier risks her per-period profits $\pi^{FD}(f)$. As the certification fee only enters via the per-period profit, $\delta^{FD}(f)$ depends on f only through $\pi^{FD}(f)$, which is concave in f. Therefore $\delta^{FD}(f)$ must be convex in f and minimized at the profit maximizing fee f^{FD}. The following corollary summarizes.

Corollary 1. *For any discount factor $\delta \geq \delta^{FD}$ there exists an interval of fees $[f_l(\delta), f_h(\delta)]$, which sustains capture-proof certification under full disclosure, where*

$$\delta^{FD} \equiv \frac{v}{v + \pi^{FD}}. \tag{4.4}$$

[11] To see this compare the left and right limit: $\lim_{b \nearrow f} \widehat{\Pi}^{FD}(b|f) = f + e(f)\delta\Pi^{FD}(f) = (1 - e(f))f + e(f)\big(f + \delta\Pi^{FD}(f)\big) = \lim_{b \searrow f} \widehat{\Pi}^{FD}(b|f)$.

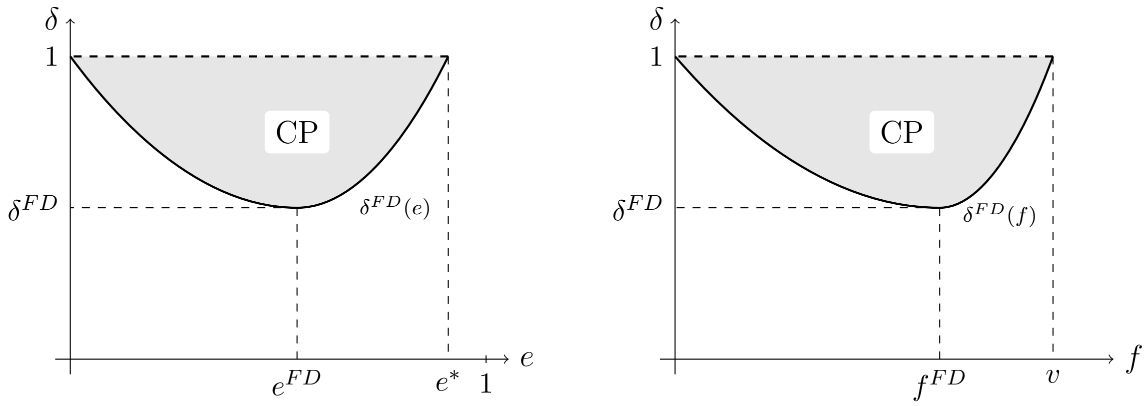

Figure 3: Capture proof combinations of (e, δ) resp. (f, δ) under full disclosure.

In the right part of Figure 3 the set of feasible (δ, f)-combinations for full disclosure is depicted.

An immediate consequence from this is that the static monopoly fee f^{FD} can sustain honest certification for all discount factors $\delta \geq \delta^{FD}$. Alternatively one might ask the question, what level of producer investment can be implemented via capture-proof certification with a full disclosure rule? The analysis follows the same arguments as above, only that certifier profits in the inequality of Proposition 2 are expressed in terms of e.

Proposition 3. *For any $\delta \geq \delta^{FD}$ there exists an interval of investment levels $[e_l^{FD}(\delta), e_h^{FD}(\delta)]$ that can be implemented in a capture-proof equilibrium. A particular investment level $e \in [0, e^*]$ can be implemented in a capture-proof equilibrium with full disclosure if and only if*

$$\delta \geq \delta^{FD}(e) \equiv \frac{v}{v + e \cdot (v - k'(e))} \tag{4.5}$$

.

The set of feasible (e, δ)-combinations is depicted in the left part of Figure 3.

Note that the first-best investment level e^* can only be (virtually) implemented for $\delta = 1$. Whenever $\delta < 1$, fees must be strictly positive in order to induce the certifier to remain honest. But then, the producer does not obtain the entire return on his investment. Hence, it must be that $e < e^*$.

4.2 Capture under partial disclosure

We next argue that alternative noisy disclosure rules can improve certifier credibility in the sense that they increase the range of discount factors that allow for capture-proof equilibria. To gain an intuition for this consider condition (4.3). This condition summarizes the trade-off between short-run gains and long-run losses. A larger profit $\pi^D(f)$ reduces the critical discount factor and full disclosure guarantees maximal per-period profits. On the other hand, $\delta^{FD}(f)$ is decreasing in v, which represents the the maximal bribe still accepted by low-type producers and therefore the largest possible short-run gain from capture. Using noisy disclosure the certifier can affect the maximal short-run gain in various dimensions. First of all, lowering the value of the best certificate or increasing the value of the worst certificate (resp. the value of uncertified products) decreases the gap between particular certification outcomes. This effect can be used to reduce the maximal bribe which producers are willing to pay. Second, with noisy disclosure the certifier can sustain an outcome where both producer types demand certification. Upon colluding with a producer type the certifier foregoes the regular certification fee, which reduces the effective gain from becoming captured.

Before analyzing noisy disclosure rules, we have to reconsider the detection possibilities by consumers. An implication of noisy rules is that consumers may hold probabilistic beliefs about a product's quality. In order to simplify matters and because it suffices to make our point clear, we focus on partial disclosure rules as as introduced in section 3. Other noisy disclosure rules are discussed in section 5 and in the appendix. Under partial disclosure, the high value certificate C^1 is awarded exclusively to high quality products which makes effective trigger punishment possible. In particular, it then suffices that the certifier is punished only if probability zero events (a low quality product was awarded certificate C^1) are observed. The fact that capture detection is not possible if bribes are being paid in exchange for the low value certificate C^2, which is assigned to both high and low types, turns out not to be crucial. This results from the fact that we can exclusively focus on equilibria where both producer types demand certification. Any partial disclosure equilibrium that has only high quality producers demand certification is outcome equivalent to a full disclosure rule. With the respective trigger beliefs it can be sustained as a capture-proof equilibrium under the same conditions as stated in the the last section. But whenever all producer types demand certification, receiving certificate C^2 is the worst possible outcome. Certificate C^2 can therefore not be part of a profitable bribing offer, as we will argue later.

To specify consumer beliefs, let $h_t = (n_t, C_t, q_t)$ denote the certification outcome in period t. Compared to the last section, we now explicitly state the disclosed certificate C_t. As before, $H_t = (h_1, \ldots, h_{t-1})$ describes the history of certification before period t. Based on these histories, we formulate trigger beliefs for consumers. Denote $\widehat{V}^D_{C_t}$ the (static) equilibrium value of a product endowed with certificate C_t under disclosure rule D.

Assumption 2. *The consumers' beliefs* $q^e_t(n_t, C_t, H_t)$ *satisfy* $q^e_t(1, C_t, H_t) = \widehat{V}^D_{C_t}$ *whenever* $C_\tau \in \mathcal{C}$ *for all* $\tau \leq t$ *and* $\{\tau < t | n_\tau = 1 \wedge Prob(C = C_t | q = q_t) = 0\} = \emptyset$. *Moreover* $q^e_t(1, C_t, H_t) = 0$ *when either* $C_\tau \notin \mathcal{C}$ *for some* $\tau \leq t$ *or* $\{\tau < t | n_\tau = 1 \wedge Prob(C = C_t | q = q_t) = 0\} \neq \emptyset$.

Note that in contrast to Assumption 1, we now also specify beliefs for cases where the testimony is inconsistent with the disclosure rule, i.e. when $C_t \notin \mathcal{C}$.[12]

Bribing offers can now be of two kinds: (C^1, b) and (C^2, b). Offer (C^2, b) is never beneficial. It would only be accepted for $b < f$, since any producer receives at least the certificate C^2 when applying for (honest) certification and the certifier gets f from any producer who is honestly tested. Thus, we can focus on bribing offers of the form (C^1, b), which we will simply refer to as b. Recall that certificate C^1 can only be awarded to high quality products. Hence, V_1, the value of a C^1-certified product, equals v. Denote V_2 the value of a C^2-certified product. Furthermore, recall that α is the probability with which a high type is awarded C^1.

A bribe b is accepted by low types whenever $V_2 - f < v - b$. High quality producers accept b if $\alpha v + (1 - \alpha) V_2 - f < v - b$. Denote $e(\alpha)$ the equilibrium investment.[13] Then

[12]In Assumption 1 this is done implicitly. Only quality levels are disclosed there and whenever $q^c_t \neq q_t$ beliefs react accordingly. In particular the trigger is pulled whenever $q \notin \{q^l, q^h\}$ is certified.

[13]The investment decision does not depend on the fee because in equilibrium, all types apply for certification and therefore pay f anyway. The expected producer profit is $e(\alpha V_1 + (1 - \alpha) V_2) + (1 - e) V_2 - f - k(e)$ and consequently the optimal investment level depends on α but not on f.

bribery acceptance probabilities are

$$p(b|\alpha,f)=\begin{cases}1, & \text{if } b<f+(1-\alpha)(v-V_2),\\ 1-e(\alpha), & \text{if } f+(1-\alpha)(v-V_2)\le b<f+(v-V_2),\\ 0, & \text{if } b\ge f+(v-V_2).\end{cases}$$

Let $\Pi^{PD}(\alpha,f)$ denote the expected profit from applying a partial disclosure rule and honestly disclosing information in each period. The corresponding expected certifier profits from bribing offer b are

$$\widehat{\Pi}(b|\alpha,f)=\begin{cases}b+e(\alpha)\delta\Pi^{PD}(\alpha,f), & \text{if } b<f+(1-\alpha)(v-V_2),\\ (1-e(\alpha))b+e(\alpha)\left(f+\delta\Pi^{PD}(\alpha,f)\right), & \text{if } f+(1-\alpha)(v-V_2)\le b<f+(v-V_2),\\ \Pi^{PD}(\alpha,f), & \text{if } b\ge f+(v-V_2).\end{cases}$$

Note that whenever high types accept the bribery offer, this is not perceived as cheating because the certificate then matches the observed quality level. Again, $\widehat{\Pi}(b|\alpha,f)$ is increasing in the respective subintervals. But the function now exhibits a downward-jump at $b=f+(1-\alpha)(v-V_2)$. The reason is that high types are willing to accept bribes strictly larger than the certification fee f to avoid the lottery between the good and the bad certificate. Therefore, at least locally, the certifier is better off bribing all producers instead of only the low types as it was the case with full disclosure. Furthermore, the maximal bribe that is accepted by at least some types is now $f+v-V_2$, which is weakly lower than under full disclosure, where the maximal bribe is v.[14] The analysis of condition (4.2) yields the following proposition.

Proposition 4. *With partial disclosure, an equilibrium satisfying Assumption 2 is capture-proof. It exists if and only if*

$$\delta\ge\delta^{PD}(\alpha,f)\equiv\max\left\{\delta^{l}(\alpha,f),\delta^{l,h}(\alpha,f)\right\}, \tag{4.6}$$

[14]In order to have all producer types demand certification it has to hold that $f\le V_2$. Consequently $f+v-V_2\le v$.

where $\delta^l(\alpha, f) = \frac{v-V_2}{v-V_2+f}$ *and* $\delta^{l,h}(\alpha, f) = \frac{(1-\alpha)(v-V_2)}{(1-\alpha)(v-V_2)+(1-e(\alpha))f}$.

The result gives a lower bound on the discount factor δ to guarantee existence of a capture-proof equilibrium with partial disclosure. The critical discount factor discount factor $\delta^{PD}(\alpha, f)$ depends on the parameters in the way how they affect short-run gain and long-run loss from capture and on which producer types accept the bribing offer that yields largest deviation profits. The term $\delta^l(\alpha, f)$ refers to the case where the largest threat stems from bribes accepted only by low types. The numerator $v-V_2$ is the effective bribe, defined as the bribery payment minus foregone payments. In the denominator we find again the effective bribe and the per-period profit f, reflecting the long-run loss from capture. The term $\delta^{l,h}(\alpha, f)$ refers to the case where the largest threat stems from bribes accepted by all types. Here the effective bribe is $(1-\alpha)(v-V_2)$. Since the long-run profit is only at stake if quality is low, long-run profits are lost with probability $(1-e(\alpha))$. Although the classical trade-off between short-run gain and long-run loss, that we already identified for full disclosure, prevails, the derivation of the maximal short-run gain is more involved for partial disclosure.

From Proposition 4 we identify a third notable difference between capture under full and noisy disclosure. Short-run gains from capture can be reduced due to the different equilibrium structure: all producers certify in equilibrium which implies that the certifier always loses fee payments if he is captured. Therefore, a larger fee f not only increases the long-run losses but at the same time reduces the short-run gains from capture.

It is now straightforward to see that $\delta^{PD}(\alpha, f)$ is decreasing in the certification fee f. This implies that for any partial disclosure rule (i.e. any α) the threat of capture is lowest when f is maximal. To keep all producers applying for certification, f cannot exceed V_2. It is therefore optimal to set $f = V_2$, which leaves low quality producers with an expected profit of zero. The following corollary summarizes.

Corollary 2. *With partial disclosure a capture-proof equilibrium satisfying Assumption 2 exists if and only if*

$$\delta \geq \delta^{PD}(\alpha) \equiv \max\left\{\delta^l(\alpha), \delta^{l,h}(\alpha)\right\}, \tag{4.7}$$

where $\delta^l(\alpha) = \frac{v-e(\alpha)(v-k'(e(\alpha)))}{v}$ *and* $\delta^{l,h}(\alpha) = \frac{1}{1+e(\alpha)}$.

Corollary 2 allows us to reduce the problem of finding the critical discount factor for partial disclosure to the one-dimensional problem of finding the optimal level of α, the probability that high quality is revealed. In fact, $\delta^{PD}(\alpha)$ depends on α only through the equilibrium value for producer investment $e(\alpha)$. The set of investment levels that can be implemented by partial disclosure is $(0, e^*)$, the same set as for full disclosure. Defining $\delta^{PD} \equiv \min_\alpha \delta^{PD}(\alpha)$ allows us to formulate the analog of Proposition 3 for partial disclosure.

Proposition 5. *For any* $\delta \geq \delta^{PD}$ *there exists an interval of investment levels* $[e_l^{PD}(\delta), e_h^{PD}(\delta)]$ *that can be implemented in a capture-proof equilibrium. A particular investment level* $e \in [0, e^*]$ *can be implemented in a capture-proof equilibrium with noisy disclosure if and only if*

$$\delta \geq \delta^{PD}(e) = \max\left\{\delta^{PD,l}(e), \delta^{PD,l,h}(e)\right\} \tag{4.8}$$

where $\delta^{PD,l}(e) = \frac{v-e(v-k'(e))}{v}$ *and* $\delta^{PD,l,h}(e) = \frac{1}{1+e}$.

Proposition 5 makes implementation of capture-proof equilibrium under full and partial disclosure directly comparable. Before investigating this in the next section we want to highlight some properties of the function $\delta^{PD}(e)$. Writing $e(v - k'(e)) = \pi^{PD}(e) = f$ the term $\delta^{PD,l}(e)$ can be expressed as $(v - f)/(v - f + \pi^{PD}(e))$. This resembles the trade-off between short-run gain and long-run loss, already identified above. Only the maximal short-run gain with partial disclosure is the maximal bribe minus foregone regular payments. The same trade-off leads to $\delta^{PD,l,h}(e)$, which is however independent of the producer's cost function $k(e)$. The maximal bribe that is accepted from both producer types in particular must be accepted from high quality producers. For them, the difference between the sure certificate C^1 and the lottery faced when certifying honestly matters. This difference is closely related to a producers' investment incentives, in fact one can show that the maximal bribe equals $v - k'(e)$. Now both short-run gain and long-run loss depend in a similar way on the investment incentives[15] and consequently

[15] As discussed, the short-run gain equals $v - k'(e)$. The long-run loss is the per-period profit, which was already shown to be $e(v - k'(e))$.

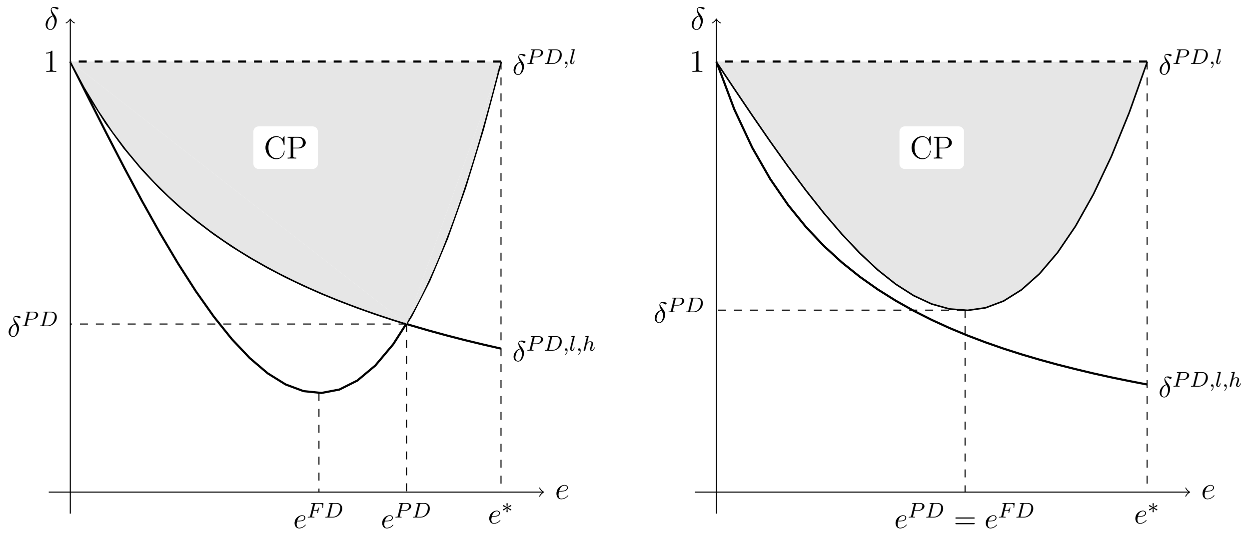

Figure 4: Capture-proof (e, δ)-combinations for low (left) and high (right) marginal costs k' at $e = e^{FD}$.

the fraction $\delta^{PD,l,h}(e)$ does not depend on the producers cost function anymore.

Which of the two terms, $\delta^{PD,l}(e)$ and $\delta^{PD,l,h}(e)$, is now larger? $\delta^{PD,l,h}(e)$ is decreasing in e, starting at 1 for $e = 0$ towards 1/2 for $e = 1$. On the other hand $\delta^{PD,l}(e)$ is convex in e with a unique minimum at $e = e^{FD}$. Furthermore $\delta^{PD,l}(0) = \delta^{PD,l}(1) = 1$. Therefore, δ^{PD} is either $\delta^{PD,l}(e^{FD})$, that is the minimum of $\delta^{PD,l}$, or it is the intersection of both fractions lying to the right of $e = e^{FD}$. Figure 4 illustrates the two cases, the latter in its left part.

4.3 Sub-optimality of full disclosure

In the previous sections, we identified the conditions under which capture-proof equilibria exist for full disclosure and a special class of noisy disclosure rules. These conditions are expressed in terms of the critical discount factors δ^{FD} and δ^{PD}. It is the aim of this study to show that opaque disclosure rules can be used by the certifier to improve his credibility. Comparing the critical discount factors δ^{FD} and δ^{PD} is short-hand for comparing the entire sets of (e, δ)-combinations, for which a capture-proof equilibrium exists with the respective disclosure rule. We are going to prove in this section that the two sets are different and, more importantly, that the respective set for full disclosure is contained in the respective set for partial disclosure. Consequently there exists an intermediate range of discount factors for which there does not exist a capture-proof equilibrium with full disclosure, but it is still possible to sustain capture-proof equilibria with partial disclosure.

As we have discussed several times throughout this paper, the key trade-off for implementing a capture-proof equilibrium is that of short-run gain versus long-run loss. Either disclosure rule leads to a per-period profit of $\pi(e) = e(v - k'(e))$ when implementing effort level e, the potential long-run loss is therefore the same. However, with partial disclosure the short-run gain from becoming captured by only low quality producers is $v - f$, compared to v for full disclosure. The resulting trade-off is resolved in favor of partial disclosure. So far this assumes that the largest threat of capture indeed stems from low quality producers. Although this is in general true for full disclosure, it ceases to hold for partial disclosure. When the maximal threat stems from a bribe accepted by all producer types, the long-run loss is reduced. Only when the producer is of low quality this is perceived as cheating by consumers and punished accordingly. So per-period profits are only lost with probability $1 - e$. On the other hand such a bribe must be smaller in order to be acceptable for high quality producers, which reduces the short-run gain. The following proposition proves that the latter effect outweighs the former.

Proposition 6. *It holds that $\delta^{PD} < \delta^{FD}$. For any $\delta \in [\delta^{PD}, \delta^{FD}]$, a capture-proof equilibrium can only be sustained applying a noisy disclosure rule. Furthermore, for any $\delta \geq \delta^{FD}$, we have that $[e_l^{FD}(\delta), e_h^{FD}(\delta)] \subsetneq [e_l^{PD}(\delta), e_h^{PD}(\delta)]$.*

Proposition 6 shows our main result that opaqueness can be used as a tool to improve certifier credibility. For any level of producer investment e, the range of discount factors that allow for capture-proof implementation of e is strictly larger for partial disclosure compared to full disclosure. Similarly, for any discount factor δ, the set of investment levels that are implementable in a capture-proof equilibrium with partial disclosure is strictly larger then the corresponding set for full disclosure. The superiority of partial disclosure therefore goes along two dimensions. Figure 5 displays these differences. The dark-grey area corresponds to (e, δ)-combinations that can be implemented as a capture-

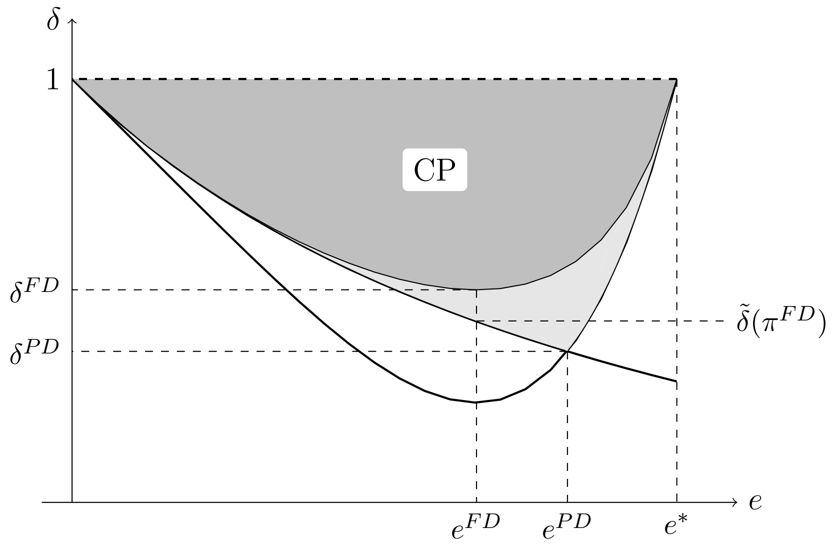

Figure 5: Dark-grey: capture-proof certification with full disclosure. Light-grey: (additional) capture-proof certification with noisy disclosure.

proof equilibrium under full disclosure. The light-grey area shows the *additional* (e, δ)-pairs that allow for implementation in capture-proof equilibrium under partial disclosure. In Section 3, we show that a certifier would always want to implement e^{FD} as this maximizes her per-period profits. With full disclosure, this is only possible when $\delta \geq \delta^{FD}$. Partial disclosure allows for capture-proof equilibria also for lower discount factors. It is remarkable that, at least for a range of discount factors, this can be achieved without waiving any profits. To see this, denote $\tilde{\delta}(\pi^{FD})$ the smallest discount factor, such that a capture-proof equilibrium is sustained and achieves per-period profits of π^{FD}. The following corollary is an immediate consequence of Proposition 6.

Corollary 3. *It holds that*

$$\tilde{\delta}(\pi^{FD}) = \max\left\{\frac{v - \pi^{FD}}{v}, \frac{1}{1 + e^{FD}}\right\} < \delta^{FD}.$$

4.4 Welfare properties of partial disclosure

In this subsection, we study welfare properties of capture-proof equilibria with partial disclosure. When $\tilde{\delta}(\pi^{FD}) = \delta^{PD}$ we also have $\delta^{PD} = (v - \pi^{FD})/v$. In this case, the largest threat of capture stems from low quality producers, i.e. the largest deviation profit for the certifier is achieved for $b = v$. Then the certifier can still achieve the maximal per-period profits π^{FD} in a capture-proof equilibrium for any $\delta \geq \delta^{PD}$, which implies implementing $e = e^{FD}$.

This is however not true when $\tilde{\delta}(\pi^{FD}) > \delta^{PD}$. As can be seen from Figure 5, for discount factors below $\tilde{\delta}(\pi^{FD})$ the profit maximizing level of investment e^{FD} is no longer capture-proof implementable. Instead only larger values of producer investment can be implemented when $\delta \in [\delta^{PD}, \tilde{\delta}(\pi^{FD}))$. To provide an intuition for this, note the following: Bribing offers b that are accepted by all producer types pose the largest threat. Now, implementing a larger e leads to a reduction in V_2, as otherwise profits would increase beyond π^{FD}. To incentivize producers to make larger investments, the certifier therefore has to increase α. As now shown, for high quality producers the difference in expected profits between the lottery of the certification process and the sure certificate v is reduced.[16] This in turn lowers the maximum bribe they are willing to pay for capture and therefore reduces the short-run gain for the certifier from any such offer. From a welfare perspective this increase in investment is beneficial. Social welfare is given by $e \cdot v - k(e)$ in each period. The first-best investment level e^* was shown to be strictly larger than e^{FD} and welfare is strictly increasing on $[0, e^*]$. Implementing certification with partial disclosure for discount factors $\delta \in [\delta^{PD}, \tilde{\delta}(\pi^{FD})]$ therefore increases social welfare compared to doing so for larger levels of the discount factor. Put differently, a severe threat of capture increases welfare. We summarize this in the following proposition.

Proposition 7. *Assume $\tilde{\delta}(\pi^{FD}) > \delta^{PD}$. For intermediate levels of the discount factor, i.e. $\delta \in [\delta^{PD}, \tilde{\delta}(\pi^{FD}))$, only investment levels that are strictly larger than e^{FD} can be capture-proof implemented with partial disclosure. This leads to increased social welfare.*

5 Discussion

We analyze the effects of reputational concerns on optimal disclosure rules from the point of view of a monopolistic certifier. Our main finding is that if capture is an issue, a certifier

[16] Honest certification yields an expected payoff $\alpha v + (1-\alpha)V_2$. This value is reduced when α increases and V_2 decreases at the same time.

benefits from resorting to coarser certification in order to reduce the threat of capture. In particular, for medium discount factors, sustaining honest certification is impossible if information is fully disclosed whereas it is still possible if information disclosure is noisy.
Implications of our analysis are manifold. First of all we provide a novel explanation for the occurrence of imperfect testing. In many papers on e.g. rating agencies (examples include Mathis et al. (2009) and Bolton et al. (2012)) imperfect testing is exogenously given, whereas here it arises in equilibrium. An empirical implication is that for low discount factors we expect disclosure to be coarser. This is consistent with the casual observation that certification in markets with low volume, such as wine, technical inspections or eco-labels often involves only a few different certificates. On the other hand, the high volume rating market exhibits a rather wide variety of different but still coarse certificates per rating agency.
Our findings also have important policy implications. Politics tend to push certifiers to precisely reveal information. Our results suggest that doing so may lead to unforeseen consequences for the functioning of those markets, as it might become more difficult to build up a reputation and resist capture if certificates are required to be too precise. Similarly, regarding the current financial crisis, forcing rating agencies to issue more precise information might even exacerbate capture problems.
We demonstrate our results in a highly stylized model, but the intuition behind our results is general. In particular, they carry over to more than only two quality specifications. Such a specification is on the one hand actually simpler, as it can be shown that already coarse deterministic disclosure rules outperform full disclosure. On the other hand the analysis is complicated by the fact that full disclosure is not necessarily optimal anymore, when capture is ignored. The first point already becomes clear from a setting with three quality levels. Full disclosure can then entail both the highest and the medium quality producer demand certification. A coarse rule would specify one certificate awarded to all but low quality. Obviously, for both rules the same types of producer demand certification. In the latter case however the maximal bribe is strictly lower. For similar investment levels and fees, the critical discount factor is therefore strictly lower for the

coarse rule. The precise analysis is more complicated, since the coarse rule generates different investment incentives for producers. In Appendix B we offer an illustration for a special case of probability distributions.

We point out that our restriction to a particular class of noisy disclosure rules is without loss of generality. First, offering various coarse certificates generates incentives for the certifier to always offer the best among the noisy certificates in a bribing offer. This will be accepted (at least by low quality producers) in order to avoid a lottery that includes the worst certificates. As deviations of this kind remain undetected they will occur with certainty, that destroys the equilibrium. Second, disclosure rules that do not allow for unambiguous detection of deviations call for a different type of trigger beliefs. Consumers lose trust in the certifier whenever they first detect low quality sold with the best certificate. This leads to punishments even if collusion did not take place. The harsher punishments makes it impossible to sustain capture proof equilibria for low discount factors. Proposition 8 in the appendix makes this statement precise.

Finally we use a specific extensive form to model capture. More sophisticated forms to study imply non-uniform bribing offers, e.g. menus, to elicit the producers' private information. Also, later bribing, after the certifier learned q or giving producers the possibility to signal their private information are possible extensions. The exact extensive form may well affect parts of the analysis, but the main finding of the advantage of opacity does not depend on the specific extensive form.

A Proofs

Proof of Lemma 1. Follows immediately from the arguments given in the text. □

Proof of Lemma 2. Following the arguments given in the text the certifier maximizes (3.2). Recall that we assume $k'''(\cdot) \geq 0$, which ensures that this profit function is concave in e, thus the first-order condition is sufficient for an optimum. This first-order condition is $0 = v - k'(e) - ek''(e)$. Define $\Psi(e) = v - k'(e) - ek''(e)$. We have $\Psi(0) = v > 0$ and $\Psi(1) = v - k'(1) - k''(1) \leq 0$ by our assumptions on $k(\cdot)$. Furthermore Ψ is strictly decreasing due to strict concavity of $k(\cdot)$. Hence there exists a unique e^{FD} such that $\Psi(e^{FD}) = 0$, which consequently is the unique maximizer of the certifier profit. The formulas for e^{FD} and f^{FD} follow easily from the formulas above. □

Proof of Proposition 1. First of all a disclosure rule can potentially lead to four different subgames: (1) no producer demands certification, (2) only low quality producers demand certification, (3) only high quality producers demand certification, and (4) all producers demand certification. Note that we do not explicitly consider mixed strategies by producers. The reason is that any outcome where some producers randomize their certification decision can be replicated by a disclosure rule that adds the respective probabilities for not certifying to the probabilities of remaining uncertified though paying for certification. To see this, assume type i chooses to certify with probability $\gamma \in (0,1)$. Now multiply every α^i by γ and increase the probability of remaining uncertified appropriately. After changing the fee from f to γf, it is easy to see that this adjusted disclosure with the reduced fee leads to the same investment incentives and also to the same equilibrium prices for (un-)certified products and the certifiers profit is unchanged.
Case (1) trivially leads to zero profits and the claim is proven.
Case (2) leads to consumers paying zero in equilibrium for certified products.[17] To make low quality producers "pay" for certification we consequently must have $f = 0$ which leads to zero profits and proves our claim also in this case.
Case (3) can be analyzed as follows: If only high types certify, rational behavior by consumers dictates that a certified product is sold at a price v. Uncertified products however can be of either high or low quality and have some price $V^{un} \in [0, v)$.
A producer's investment decision is given by the solution of

$$\max_e \; e\Big(\sum_k \alpha_k^1 v + (1 - \sum_k \alpha_k^1) V^{un} - f\Big) + (1-e)V^{un} - k(e),$$

[17] A disclosure leading to this particular subgame is given by $\mathcal{C} = \{C\}, \alpha^l = 1$ and $\alpha^h = 0$.

which yields the following first-order condition for producer investment:

$$\Big(\sum_k \alpha_k^1 (v - V^{un}) - f\Big) = k'(e).$$

Rewriting this constraint in terms of induced investment yields $f = v - k'(e) - (1 - \sum_k \alpha_k^1)(v - V^{un}) - V^{un}$. Now we have for the certifier profit

$$\pi^D(f) = e(f, D) \cdot f = e \cdot \big(v - k'(e) - (1 - \sum_k \alpha_k^1)(v - V^{un}) - V^{un}\big) \leq e \cdot (v - k'(e)) \leq \pi^{FD}.$$

This proves the claim for case (3).

Finally consider case (4): When both producer types demand certification, the resulting certifier profit in the subgame is $\pi^D(f) = f$. The price V^i of a product sold with certificate C^i is

$$V^i = v \cdot \frac{e\alpha_i^h}{e\alpha_i^h + (1-e)\alpha_i^l}.$$

Uncertified products are sold at price $V^{un} = v \cdot \frac{e\big(1 - \sum_i \alpha_i^h\big)}{e\big(1-\sum_i \alpha_i^h\big) + (1-e)\big(1 - \sum_i \alpha_i^l\big)}$. A producer's investment decision follows from maximizing his expected payoff from certification, given by

$$e \cdot \left(\sum_i \alpha_i^h V^i + \left(1 - \sum_i \alpha_i^h\right) V^{un}\right) + (1-e) \cdot \left(\sum_i \alpha_i^l V^i + \left(1 - \sum_i \alpha_i^l\right) V^{un}\right) - f - k(e).$$

The resulting investment constraint is

$$k'(e) = \sum_i (\alpha_i^h - \alpha_i^l)(V^i - V^{un}). \tag{A.1}$$

On the other hand, from the formula given for V^i we have $e\alpha_i^h V^i + (1-e)\alpha_i^l V^i = ev\alpha_i^h$. Similarly $e(1 - \sum_i \alpha_i^h)V^{un} + (1-e)(1 - \sum_i \alpha_i^l)V^{un} = ev(1 - \sum_i \alpha_i^h)$. Summing those expressions yields

$$\sum_i \left(e\alpha_i^h V^i + (1-e)\alpha_i^l V^i\right) + e(1 - \sum_i \alpha_i^h)V^{un} + (1-e)(1 - \sum_i \alpha_i^l)V^{un} = ev. \tag{A.2}$$

Rewriting the left hand side of equation (A.2) yields

$$e\sum_i(\alpha_i^h-\alpha_i^l)(V^i-V^{un})+\sum_i\alpha_i^lV^i+\left(1-\sum_i\alpha_i^l\right)V^{un}=ev. \tag{A.3}$$

Finally, to make all producer types demand certification we must have in particular

$$f\le\sum_i\alpha_i^lV^i+\left(1-\sum_i\alpha_i^l\right)V^{un} \tag{A.4}$$

i.e. low quality producers ecpected payoff from certification must be non-negative.[18] From this we can derive an upper bound on certifier profits:

$$\begin{aligned}\pi^D(f)=f &\overset{(A.4)}{\le} \sum_i\alpha_i^lV^i+\left(1-\sum_i\alpha_i^l\right)V^{un}\\ &\overset{(A.3)}{=} ev-e\sum_i(\alpha_i^h-\alpha_i^l)(V^i-V^{un})\\ &\overset{(A.1)}{=} ev-ek'(e)=e(v-k'(e)).\end{aligned}$$

But $e\,(v-k'(e))$ is the profit from implementing effort level e optimally with a full disclosure rule, therefore we have proven $\pi^D(f)\le\pi^{FD}$.

□

Proof of Proposition 2. In any equilibrium in which Assumption 1 holds capture may not take place, since otherwise the beliefs of consumers are not consistent with the behavior of the certifier. Hence, condition (4.2) must be satisfied for all b. As mentioned in the text, certifier profits from deviating $\widehat{\Pi}^{FD}(b|f)$ are largest for b approaching v. Taking this limit yields

$$\begin{aligned}\lim_{b\nearrow v}\widehat{\Pi}^{FD}(b|f)=&\ (1-e(f))v+e(f)\cdot\left(f+\delta\Pi^{FD}(f)\right)\\ =&\ (1-e(f))v+\pi^{FD}(f)+\frac{\delta}{1-\delta}e(f)\pi^{FD}(f)\\ =&\ (1-e(f))v-\frac{\delta}{1-\delta}(1-e(f))\pi^{FD}(f)+\Pi^{FD}(f).\end{aligned}$$

[18]More conditions are required in subgame where all producer types demand certification, but the one presented her is the only required for our proof.

Condition (4.2) is thus equivalent to

$$(1 - e(f))v \leq \frac{\delta}{1-\delta}(1 - e(f))\pi^{FD}(f).$$

Rearranging this expression yields that condition (4.2) is satisfied if and only if

$$\delta \geq \delta^{FD}(f) \equiv \frac{v}{v + \pi^{FD}(f)}.$$

□

Proof of Proposition 3. We first argue how condition (4.3) can be translated towards (4.5). Recall $\pi^{FD}(f) = e(f) \cdot f$ and optimal investment by producers requires $k'(e) = v - f$. Replacing f by $v - k'(e)$ yields (4.5). All other statements are straightforward reformulations of Proposition 2 and Corollary 1. □

Proof of Proposition 4. In any equilibrium in which Assumption 2 holds capture may not take place, since otherwise the beliefs of consumers are not consistent with the behavior of the certifier. Hence, condition (4.2) must be satisfied for all b. We compute the respective critical discount factors. Taking the limit of $\widehat{\Pi}^D(b|f)$ as b approaches $f + (1-\alpha)(v - V_2)$ we get

$$\begin{aligned}
\lim_{b \nearrow f+(1-\alpha)(v-V_2)} \widehat{\Pi}^D(b|f) &= f + (1-\alpha)(v - V_2) + e(\alpha)\delta\Pi^{PD}(f) \\
&= f + (1-\alpha)(v - V_2) + e(\alpha)\frac{\delta}{1-\delta}f \\
&= (1-\alpha)(v - V_2) - \frac{\delta}{1-\delta}(1 - e(\alpha))f + \Pi^{PD}(f).
\end{aligned}$$

Consequently this limit lies below $\Pi^{PD}(f)$ if and only if

$$(1-\alpha)(v - V_2) \leq \frac{\delta}{1-\delta}(1 - e(\alpha))f,$$

respectively whenever

$$\delta \geq \delta^{l,h}(\alpha, f) = \frac{(1-\alpha)(v - V_2)}{(1-\alpha)(v - V_2) + (1 - e(\alpha))f}.$$

Similarly the limit of $\widehat{\Pi}^D(b|f)$ as b approaches $f + (v - V^1)$ can be rewritten as follows

$$\lim_{b \nearrow f+(v-V_2)} \widehat{\Pi}^D(b|f) = (1 - e(\alpha)) \cdot (f + (v - V_2)) + e(\alpha)\big(f + \delta\Pi^{PD}(f)\big)$$
$$= (1 - e(\alpha))(v - V_2) - \frac{\delta}{1-\delta}(1 - e(\alpha))f + \Pi^{PD}(f).$$

Therefore $\lim_{b \nearrow f+(v-V_2)} \widehat{\Pi}^D(b|f) \leq \Pi^{PD}(f)$ if and only if

$$\delta \geq \delta^l(\alpha, f) = \frac{v - V_2}{f + v - V_2}.$$

Because capture-proofness requires $\widehat{\Pi}^D(b|f) \leq \Pi^{PD}(f)$ for all b, (4.6) follows. □

Proof of Corollary 2. As discussed in the text, the certifier may set $f = V_2$ to minimize the threat of capture. We consider $\delta^l(\alpha, f)$ first. Making use of $f = V_2$ allows us to simplify it to $(v - V_2)/v$. From the proof of Proposition 1 we get $V_2 = e(v - k'(e))$ and therefore

$$\delta^l(\alpha) = \frac{v - V_2}{v} = \frac{v - e(\alpha)(v - k'(e(\alpha)))}{v}.$$

Now consider $\delta^{l,h}(\alpha, f)$. With $f = V_2$ we may rewrite

$$\delta^{l,h}(\alpha, f) = \frac{(1-\alpha)(v - V_2)}{(1-\alpha)(v - V_2) + (1 - e(\alpha))V_2}$$

By Bayesian updating we have $V_2 = v \cdot \big((1-\alpha)e(\alpha)\big)/\big(1 - \alpha e(\alpha)\big)$ in equilibrium, which implies $v - V_2 = v \cdot \big(1 - e(\alpha)\big)(\big(1 - \alpha e(\alpha)\big)$. Replacing V_2 and $v - V_2$ accordingly yields

$$\frac{(1-\alpha)(v - V_2)}{(1-\alpha)(v - V_2) + (1 - e(\alpha))V_2} = \frac{1}{1 + e(\alpha)}.$$

□

Proof of Proposition 6. Recall, that with full disclosure the critical discount factor is $\delta^{FD}(e) = \frac{v}{v+\pi^{FD}(e)} = \frac{v}{v+e(v-k'(e))}$ and this term is minimized for the profit maximizing effort e, yielding $\min_e \delta^{FD}(e) = \frac{v}{v+\pi^{FD}}$. For all $e \in (0, e^*)$ we have $\frac{v-e(v-k'(e))}{v} < \delta^{FD}(e)$.

To see this:

$$\frac{v - e(v - k'(e))}{v} < \delta^{FD}(e) = \frac{v}{v + e(v - k'(e))} \quad \Leftrightarrow \quad \big(e(v - k'(e))\big)^2 > 0.$$

Also

$$\frac{1}{1+e} < \delta^{FD}(e) = \frac{v}{v + e(v - k'(e))} \quad \Leftrightarrow \quad ek'(e) > 0$$

Therefore also $\max\{\frac{1}{1+e}, \frac{v-e(v-k'(e))}{v}\} < \delta^{FD}(e)$ for all $e \in (0, e^*)$ and hence we can define

$$\delta^{PD} := \min_e \max\left\{\frac{1}{1+e}, \frac{v - e(v - k'(e))}{v}\right\}$$

and it follows that $\delta^{FD} > \delta^{PD}$. Since both $\delta^{PD,l}(e) < \delta^{FD}(e)$ and $\delta^{PD,l,h}(e) < \delta^{FD}(e)$ the last statement follows immediately. □

Proof of Proposition 7. When $\delta^{PD} < \tilde{\delta}(\pi^{FD})$ we must have $\tilde{\delta}(\pi^{FD}) = \delta^{PD}(e^{FD}) = \frac{1}{1+e^{FD}}$. Since $1/1+e$ decreases in e we have $\delta^{PD}(e) > \tilde{\delta}(\pi^{FD})$ for any $e < e^{FD}$. Consequently we must have $\delta^{PD}(e) < \tilde{\delta}(\pi^{FD})$ on some interval $[e^{FD}, \hat{e}]$. This proves our result. □

B Extensions

Example B.1.

Let quality levels be $\{0, 0.5, 1\}$ and $P(q = 0.5|e) = P(q = 1|e) = e/2$. Consequently $P(q = 0|e) = 1 - e$. The cost of effort is $k(e) = e^2/2$. If we restrict the analysis to deterministic disclosure rules, it is straightforward to show that full disclosure with a fee $f = 3/8$ maximizes certifier profits. With this fee both quality levels 0.5 and 1 get certified in equilibrium. Using the same line of argument as in the main text, this disclosure rule can be sustained as a capture-proof equilibrium whenever $\delta \geq \frac{16}{19}$.

A cut-off disclosure rule that certifies any product whose quality exceeds 0, but does not distinguish any further, achieves the same static profit as the mentioned full disclosure rule. However, the largest possible bribe is then not equal to 1 since no certificate which yields a price of one is available. Instead, the best certificate yields 3/4, the value of a certified product. Consequently, a capture-proof equilibrium with this disclosure rule exists whenever $\delta \geq \frac{16}{20}$. While profits remain the same, the largest acceptable bribe is lowered.

Proposition 8. *For any $\delta < \delta^{FD}$ and any disclosure rule which is such that the highest value certificate is different from v, no capture-proof equilibrium exists.*

Proof. We restrict the proof to the following simple disclosure rule[19]: there are two certificates, C_1 and C_2, where high quality always receives C_1 and low quality receives C_1 with probability $\alpha \in (0,1)$. Denote V the value of C_1, certificate C_2 is always worth zero (in equilibrium). The first-order condition for producer investment reads as

$$k'(e) = (1-\alpha)V$$

and from Baye's rule we have

$$V = v\frac{e}{e+\alpha(1-e)}.$$

Thus, to implement a particular e, the certifier has to set[20]

$$\alpha = \frac{e(v-k'(e))}{e(v-k'(e))+k'(e)}$$

The fee must be such that low quality producers are willing to get their product certified, i.e. $f \leq \alpha V$.

When a purchased product with certificate C_1 turns out to be of low quality, consumers cannot be sure whether this was due to bad luck or to a captures certifier. Appropriate trigger beliefs have to be such that the certifier is punished whenever low quality is sold with certificate C_1. This can well happen without any deviation by the certifier. The probability of entering punishment, absent any deviation, is $p = (1-e)\alpha$ and expected

[19]For any other rule, the argument is the same for selling the best certificate in a capture offer to the low quality producer. However, there are even more feasible bribing offers, which make it even harder to resist the threat of capture.

[20]Note that $\lim_{e\to 0}\alpha$ equals 1 whenever $k''(0) = 0$ and otherwise equals $\frac{v}{v+k''(0)\in(0,1)}$, that is in the latter case not all α are implementable.

profits from honest play are given by

$$\Pi^h(\alpha, f) = f + (1-p)\delta f + (1-p)^2\delta^2 f + \ldots = \frac{f}{1-(1-p)\delta}.$$

The maximal bribe is given by $b \approx (1-\alpha)V + f$, where only low quality producers accept it. The profit from making such an offer is

$$\Pi(b|f, \alpha) = (1-e)b + e(f + \delta\Pi^h(\alpha, f))$$

We have $\Pi(b|f,\alpha) \leq \Pi^h(\alpha, f)$ for $b \to (1-\alpha)V + f$ whenever

$$\delta \geq \frac{(1-e)b - (1-e)f}{(1-e)(1-p)b - epf} = \frac{b-f}{\big(1-(1-e)\alpha\big)b - e\alpha f}$$

This is both increasing in b and in f, such that the largest threat is exercised for $f = \alpha V$ and $b = V$, which results in the condition

$$\delta \geq \frac{1}{1+e\alpha}.$$

We have $\frac{1}{1+e\alpha} \geq \frac{v}{v+e(v-k'(e))}$ if and only if

$$v - k'(e) \geq v\alpha \quad \Leftrightarrow \quad 1 \geq e.$$

Hence, for all e to be implemented, this is only possible with a noisy rule without sure high quality certificate, when this is also possible using a full disclosure rule.

□

References

Albano, G.L. and A. Lizzeri, "Strategic certification and provision of quality," *International Economic Review*, 2001, *42* (1), 267–283.

Biglaiser, G., "Middlemen as Experts," *RAND Journal of Economics*, 1993, *24* (2), 212–223.

Bolton, P., X. Freixas, and J. Shapiro, "The Credit Ratings Game," *Journal of Finance*, 2012, *67* (1), 85–112.

Dubey, P. and J. Geanakoplos, "Grading exams: 100, 99,..., 1 or A, B, C? Incentives in Games of Status," *Games and Economic Behavior*, 2010, *69* (1), 72–94.

Farhi, E., J. Lerner, and J. Tirole, "Fear of Rejection? Tiered Certification and Transparency," 2012. http://idei.fr/doc/by/tirole/fear091212.pdf.

Faure-Grimaud, A., E. Peyrache, and L. Quesada, "The ownership of ratings," *RAND Journal of Economics*, 2009, *40* (2), 234–257.

Kartasheva, A. and B. Yilmaz, "Precision of Ratings," 2012. mimeo. University of Pennsylvania.

Kofman, F. and J. Lawarrée, "Collusion in Hierarchical Agency," *Econometrica*, 1993, *61* (3), 629–56.

Levin, J., "Relational incentive contracts," *American Economic Review*, 2003, *93* (3), 835–857.

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

Mailath, G.J. and L. Samuelson, *Repeated Games and Reputations: Long-run Relationships*, Oxford University Press, 2012.

Mathis, J., J. McAndrews, and J.C. Rochet, "Rating the raters: are reputation concerns powerful enough to discipline rating agencies?," *Journal of Monetary Economics*, 2009, *56* (5), 657–674.

Shapiro, C., "Premiums for high quality products as returns to reputations," *Quarterly Journal of Economics*, 1983, *98* (4), 659.

Strausz, R., "Honest certification and the threat of capture," *International Journal of Industrial Organization*, 2005, *23* (1-2), 45–62.

Tirole, J., "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations," *Journal of Law, Economics, and Organization*, 1986, *2* (2), 181–214.