Selling Tests under Moral Hazard*

Peiran Xiao†

August 2023

Abstract

I study how a monopoly certifier designs and prices the quality certification to maximize revenue when agents' quality is endogenous. The certifier designs the quality certification (i.e., Blackwell experiment) and charges agents a flat certification fee. Agents with heterogeneous costs choose their quality (and whether to take the test) and receive a payment equal to the expected quality conditional on the test result by the competitive market. By characterizing the feasibility condition for interim posterior means, I show the certifier can maximize revenue by committing to a noisy test. The interim approach applies to a more general class of problems, revenue maximization being a special case where the feasibility condition is equivalent to Bayesian plausibility.

I study how a monopoly certifier (e.g., credit rating agencies, standardized test providers) designs and prices quality certifications to maximize revenue when agents' quality is endogenous. The certifier designs the quality certification (i.e., Blackwell experiment) and charges agents a flat certification fee. Agents with heterogeneous costs choose their quality and whether to take the test. Based on the test result, the competitive market pays the agent the expected quality. I use an interim information design approach to study the problem. Specifically, I characterize the feasibility condition for *interim* posterior means to be inducible by Blackwell experiments. Thus, I consider an equivalent reduced-form problem where the certifier designs an interim posterior mean under the feasibility condition, rather than the experiment itself. I show that the

*Since the first draft, I have found my model essentially equivalent to Albano and Lizzeri (2001), although from an interim information design perspective. Moreover, Saeedi and Shourideh (2020) also use the same interim approach to a similar problem and prove the characterization theorem of feasible posterior means. Therefore, I have decided to summarize my findings in this note, previously titled "An interim information design approach to Albano and Lizzeri (2001)."

†Department of Economics, Boston University. Email: pxiao@bu.edu.

certifier can maximize revenue by committing to a noisy test that fully reveals quality with some probability and outputs the same signal for every participant otherwise. The interim approach applies to a more general class of problems, revenue maximization being a special case where the feasibility condition is equivalent to Bayesian plausibility. In this case, once Bayesian plausibility is substituted into the certifier's objective, the model essentially becomes the Laffont-Tirole cost-reimbursement model (see Laffont and Tirole, 1993, Chapter 1) with exclusion.

I focus on the *interim* posterior mean because quality is endogenously chosen by agents who privately know their types, and the competitive market offers a price equal to the agent's quality. Thus, the interim posterior mean, which is the agent's expected gain from the certification conditional on his privately known quality, captures the agent's incentive for quality provision.

A full characterization of feasible interim posterior means (i.e., can be induced by an experiment) has already been given by Saeedi and Shourideh (2020). Moreover, Saeedi and Shourideh (2022) provide a more general theorem when investing effort stochastically increases quality.[1] Doval and Smolin (2021) also study the information design problem from an interim perspective and characterizes the set of feasible interim payoffs in general (without the monotonicity constraint).

Apart from the interim approach, the model is essentially equivalent to Albano and Lizzeri (2001, henceforth AL) with one distinction—I limit transfers to a flat testing fee, so the test is the only channel to provide incentives. Because I model the certification as a Blackwell experiment rather than a disclosure after the certifier observes the quality, an upfront flat fee is more realistic and also ex-post incentive compatible in that the certifier has no incentive to tamper with the experiment (see Section 1.1). In AL, the fee is contingent on the agent's quality, so a wide range of test-fee structures achieve the maximal revenue, including a flat fee and noisy disclosure.[2]

In AL, agents are sellers who choose their quality provision, whereas the competitive market pays the seller a price equal to the expected quality. Alternatively, in a job market example, agents are students with heterogeneous costs of learning (human capital accumulation), and the job market offers a wage equal to the expected quality (human capital). In other words, learning is productive, in contrast to Spence's signaling model.

[1] Rodina (2020) shows a characterization theorem when investing effort stochastically increases quality but when the market values the exogenous ability rather than endogenous quality.

[2] As AL, I do not consider a test-fee structure that contains a disclosure fee, which the agent pays to disclose the test result to the market, in addition to the testing fee (see Ali et al., 2021).

1 The Model

1.1 Setup

An agent (he) chooses quality $\theta \in [0, \theta_{\max}] \equiv \Theta \subseteq \mathbb{R}_+$ at cost $c(\theta, t)$, where t is his type drawn according to a continuous CDF F with support $T = [\underline{t}, \bar{t}] \subseteq \mathbb{R}_+$. A monopoly certifier (she) commits to a test (Blackwell experiment) $\pi : \Theta \to \Delta(S)$ that takes the agent's quality $\theta \in \Theta$ as input and outputs a stochastic signal $s \in S$. The agent can choose whether to take the test or not. If he takes the test, the certifier charges him a price $P \geq 0$ and sends a signal s drawn from the distribution $\pi(\theta)$ to the market. Otherwise, he gets a null signal $s = \varnothing$ at no cost. The competitive market (e.g., two buyers bidding in a first-price auction) offers a price $p(s) = \mathbf{E}[\theta|s] = \mathbf{E}_{\mu_s}[\theta]$, where μ_s is the posterior belief induced by s.

The setup is the same as AL, except that I model certification as a Blackwell experiment and the certification fee as an upfront flat fee. By contrast, they model the certification as strategic disclosure (garbling) after the certifier perfectly observes the quality, so that the certification fee can be a function of the quality. First, as they have shown, an upfront flat fee is without loss of generality as it can also achieve maximal revenue. Second, compared to a fee schedule contingent on the agent's quality, an upfront fee is more realistic and also ex-post incentive compatible in that the certifier has no incentive to tamper with the certification (experiment).

For the convenience of comparison, I adopted the same notations whenever possible.[3] The equilibrium concept is sequential equilibrium as in their model. I also maintain their assumptions:

Assumption 1. For all $t \in T$ and $\theta \in \Theta$, $c_\theta > 0$, $c_t < 0$, $c_{tt} > 0$, $c_{t\theta} < 0$.

Assumption 2. There exists some $\theta \in \Theta$ such that $\theta - c(\theta, \bar{t}) > 0$.

Assumption 3. $c_{\theta\theta t} \leq 0$, $c_{\theta tt} \geq 0$.

Assumption 4. $\frac{1-F(t)}{f(t)}$ is decreasing.

Assumption 5. For all $t \in [\underline{t}, \bar{t}]$, $c(0, t) = 0$.[4]

The assumption can be weakened to:

Assumption 5'. For all $t \in [\underline{t}, \bar{t}]$, there exists $\theta^o(t) \in \Theta$ such that $c(\theta^o(t), t) = 0$.

[3] I would have used q for quality and θ for types, whereas AL use θ for quality and t for types.
[4] Although AL do not impose this assumption explicitly, it is necessary for Lemmas 1 and 2.

Assumption (A5') accommodates the additive case where $c(\theta, t) = c(\theta - t) \equiv c(e)$, that is, agents invest effort $e \geq 0$ at a cost $c(e)$ to increase $q = \theta + e$. Thus, higher types are endowed with higher quality without incurring any costs.

Denote the decision that a type-t agent takes the test by $\sigma(t) = \mathbf{1}[t \text{ takes the test}]$. We have the following lemmas:

Lemma 1. *If $\sigma(t) = 0$, then $\theta(t) = 0$, and the market offers $w_\varnothing = 0$.*

Remark 1. Under Assumption (A5'), Lemma 1 becomes: If $\sigma(t) = 0$, then $\theta(t) = \theta^o(t)$, and the market offers $w_\varnothing = \mathbf{E}[\theta^o(t) \mid \sigma(t) = 0]$.

Lemma 2. *There exists a cutoff type t_0 such that $\sigma(t) = 1$ if and only if $t \geq t_0$.*

1.2 Feasibility

From an interim perspective, I characterize a test π by the *interim posterior mean* $w(\theta) = \mathbf{E}_{s\sim\pi(\theta)}[p(s)]$ induced by π. Formally, say a test π *induces* $w(\theta)$ if $w(\theta) = \mathbf{E}_{s\sim\pi(\theta)}[\mathbf{E}[\tilde{\theta}|s]]$. I focus on increasing $w(\theta)$ due to incentive compatibility.

Denote the set of quality of agents who take the test in equilibrium by $\Theta^* \equiv \{\theta^*(t) \in \Theta : \sigma^*(t) = 1, t \in T\}$. Denote the market's prior on their quality by $\mu \in \Delta\Theta^*$. In equilibrium, we have $\mu(\theta^*(t)) = \bar{F}(t)$, where $\bar{F}(t) \equiv \Pr(\tilde{t} \leq t | \sigma^*(\tilde{t}) = 1)$ is a truncated distribution of F over the types who take the test. As will be shown later, $\theta^*(t)$ and $\sigma^*(t)$ are increasing, so $\operatorname{supp}(\mu) = \Theta^* = [\underline{\theta}, \bar{\theta}]$.

Definition 1. An interim posterior mean $w(\theta)$ is *feasible* if there exists a Blackwell experiment $\pi : \Theta \to \Delta S$ that induces $w(\theta)$, i.e., $w(\theta) = \mathbf{E}_{s\sim\pi(\theta)}[\mathbf{E}[\tilde{\theta}|s]]$.

Therefore, the problem can be reduced to an equivalent one where the certifier designs a *feasible* interim posterior mean $w(\theta)$ instead of a test π. If restricted to deterministic tests (i.e., monotone partitions), any increasing feasible $w(\theta)$ is a "truthful filter" (Rayo, 2013) that satisfies $w(\theta) = \mathbf{E}[\tilde{\theta}|s(\theta') = s(\theta)]$ for all θ—i.e., a piecewise function that is either $w(\theta) = \theta$ or constant on each interval. In general, we have the following characterization theorem, which can be found in Saeedi and Shourideh (2020, Proposition 1 and Theorem 1).[5]

Theorem 1. *If $w(\theta)$ is increasing, then it is feasible if and only if θ majorizes $w(\theta)$ in the quantile space,*[6] *that is,*

[5] See also Saeedi and Shourideh (2022, Lemma 1 and Theorem 1) for a more general characterization theorem when investing effort increases quality stochastically.

[6] Consider the quantile $\sigma = \mu(\theta)$, and define $\omega(\sigma) = w(\mu^{-1}(\sigma))$. Then, θ majorizes $w(\theta)$ in the quantile space means that $\mu^{-1}(\sigma)$ majorizes $\omega(\sigma)$ (denoted by $\mu^{-1}(\sigma) \succeq \omega(\sigma)$). Equivalently, say θ is a mean-preserving spread (MPS) of $w(\theta)$ in the quantile space, or $\omega \in \text{MPS}(\mu^{-1})$.

(i) $\int_0^t w(\theta)\,\mathrm{d}\mu(\theta) \geq \int_0^t \theta\,\mathrm{d}\mu(\theta)$ ***for all*** $t \in \operatorname{supp}(\mu)$ *(SOSD),*

(ii) $\int_0^{\bar{\theta}} w(\theta)\,\mathrm{d}\mu(\theta) = \int_0^{\bar{\theta}} \theta\,\mathrm{d}\mu(\theta)$ *(Bayesian plausibility).*

Proof. Define the quantile $\sigma = \mu(\theta)$ and $\omega(s) = w(\mu^{-1}(\sigma))$. ($\Rightarrow$) Recall that when the test is π, we denote $p(s) = \mathbf{E}_{\mu_s}[\theta]$ and $w(\theta) = \mathbf{E}[p(s)|\theta]$, where μ_s is the posterior belief induced by s. Consider the fully revealing test $\bar{\pi}$, denote $\bar{p}(s) = \mathbf{E}_{\bar{\mu}_s}[\theta]$, where $\bar{\mu}_s$ is the posterior belief induced by s. Fix any $\theta \in \operatorname{supp}(\mu)$, we have, for any $\tau \in \Theta$,

$$\mathbf{E}[p(s) \mid \theta \geq \tau] \leq \mathbf{E}[\bar{p}(s) \mid \theta \geq \tau].$$

Because

$$\mathbf{E}[p(s) \mid \theta \geq \tau] = \frac{1}{1-\mu(\tau)} \int_\tau^{\bar{\theta}} w(\theta)\,\mathrm{d}\mu(\theta) = \frac{1}{1-\sigma} \int_\sigma^1 \omega(\tilde{\sigma})\,\mathrm{d}\tilde{\sigma},$$
$$\mathbf{E}[\bar{p}(s) \mid \theta \geq \tau] = \frac{1}{1-\mu(\tau)} \int_\tau^{\bar{\theta}} \theta\,\mathrm{d}\mu(\theta) = \frac{1}{1-\sigma} \int_\sigma^1 \mu^{-1}(\tilde{\sigma})\,\mathrm{d}\tilde{\sigma},$$

and $\mathbf{E}[w(\theta)] = \mathbf{E}[\theta]$, we have $\mu^{-1} \succeq \omega$.

($\Leftarrow$) By Kleiner et al. (2021, Proposition 1), $\omega \in$ MPS(μ^{-1}) implies there exists a probability measure λ supported on the extreme points of MPS(μ^{-1}) such that

$$\omega = \mathbf{E}[\tilde{\omega}|\tilde{\omega} \sim \lambda],$$

that is, $w = \mathbf{E}[\tilde{w}|\tilde{w} \sim \ell]$ where $\ell(\theta) = \lambda(\mu(\theta))$. One can thus implement w by randomizing over monotone partitional signals that implement extreme points of MPS(μ^{-1}) (i.e., either fully revealing or pooling). □

Remark 2. This is reminiscent of the symmetric version of Border's theorem (Maskin-Riley-Matthews condition) in reduced-form mechanism design (see, e.g., Border, 1991, 2007; Maskin and Riley, 1984; Matthews, 1984; Kleiner et al., 2021). The proof is à la Kleiner et al. (2021, Theorem 3).

Indeed, Border's theorem characterizes the feasibility condition for interim allocation rules to be implementable by ex-post allocations, thus allowing us to optimize over interim allocations (i.e., reduced-form mechanisms). Analogously, Theorem 1 characterizes the feasibility condition for interim posterior means to be inducible by Blackwell experiments, thus allowing us to optimize over interim posterior means rather than experiments themselves.

In particular, when $w'(\theta) \leq 1$ on $[\underline{\theta}, \bar{\theta}]$, it is straightforward to show that feasibility is equivalent to (BP).

Corollary 1.1. *If $w'(\theta) \leq 1$, then (BP) $\Rightarrow$ (SOSD), so feasibility $\Leftrightarrow$ (BP).*

The proof follows from the fact that $w^*(\theta) - \theta$ is decreasing (i.e., $w(\theta)$ single-crosses the 45 degree line from above) if $w^{*\prime}(\theta) \leq 1$.

1.3 Illustrative Example

Assume there are only two types: $t \in \{2, 4\}$ and $\Pr(t = 4) = p_0$. The cost function is $c(\theta, t) = \theta^2/2t$. Under a fully revealing test, the interim posterior mean $w(\theta) = \theta$, and the agent's utility is $u(\theta, t) = \theta - \theta^2/2t$, so he produces $\theta^*(t) = t$, and his indirect utility $U(t) = t/2$. The certifier can set either $P = 2$ to earn an expected revenue of $R = 2p_0$ (only high-type takes the test) or $P = 1$ to earn $R = 1$ (both types take the test). When $p_0 = 0.55$, it is optimal to set $P = 2$.

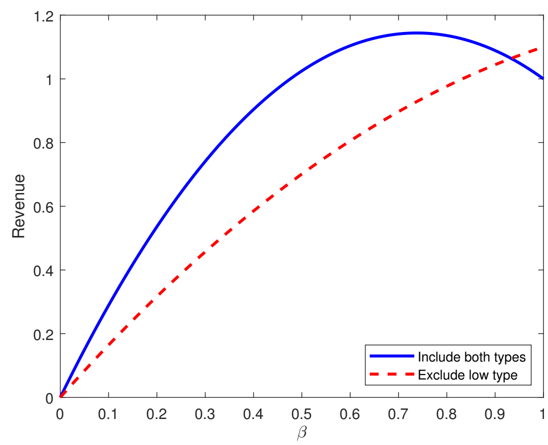


Figure 1: Including the low type can generate a higher revenue when the test is noisy

Now consider a *noisy* test that reveals θ with probability β and outputs a "pass" signal with probability $1 - \beta$. Now the interim posterior mean is $w(\theta) = \beta\theta + (1 - \beta)\,\mathbf{E}[\tilde{\theta}|\text{pass}]$ and $\theta^* = \arg\max w(\theta) - c(\theta, t)$. Therefore, in equilibrium, the quality $\theta^* = \beta t$, the interim posterior mean $w(\theta) = \beta\theta + 2(1 + p_0)\beta(1 - \beta)$, and the indirect utility $U(t) = \beta^2 t/2 + 2(1 + p_0)\beta(1 - \beta)$. When $p_0 = 0.55$, it is optimal to set $\beta = 0.74$ and $P = 1.14$ (both types take the test) to earn $R = 1.14 > 1$. Therefore, the certifier can profit from a noisy test.

The intuition is as follows. On the one hand, a more precise test provides the agent a higher incentive to improve quality and therefore a higher willingness to pay for the test. On the other hand, a noisier (i.e., less precise) test redistributes the payoffs from high-type (i.e., low-cost) agents to low-type agents. Because the price of the test cannot exceed the payoff of the agent who takes the test,[7] the redistribution allows the agent to

[7] The highest possible price equals the payoff of the lowest type among those who take the test.

charge a higher price while including lower types. Consequently, a noisy test maximizes the certifier's revenue.

2 Optimal Test Design

The certifier's problem, reformulated as choosing a feasible interim posterior mean $w(\theta)$ instead of a test π, is

$$\max_{P,w(\theta),\theta(t),\sigma(t)} \int_{\underline{t}}^{\bar{t}} \sigma(t) P \,\mathrm{d}F(t) \tag{1}$$

subject to

$$w(\theta) = \mathbf{E}_{s\sim\pi(\theta)}[\mathbf{E}[\tilde{\theta}|s]] \text{ is feasible,} \tag{2}$$

$$\theta(t), \sigma(t) \in \underset{\hat{\theta}\in\Theta,\, \hat{\sigma}\in\{0,1\}}{\arg\max}\ [w(\hat{\theta}) - c(\hat{\theta}, t) - P]\hat{\sigma}, \qquad \text{(IC)} \tag{3}$$

$$U(t) \equiv [w(\theta(t)) - c(\theta(t), t) - P]\sigma(t) \geq 0,\ \forall t \in [\underline{t}, \bar{t}] \qquad \text{(PC)} \tag{4}$$

By the standard argument, (IC) is equivalent to

$$U(t) = U(\underline{t}) - \int_{\underline{t}}^{t} \sigma(x) c_t(\theta(x), x) \,\mathrm{d}x, \tag{5}$$

$$\theta(t) \text{ and } \sigma(t) \text{ are increasing,} \tag{6}$$

which also implies $w(\theta)$ is increasing. By Theorem 1, feasibility is equivalent to

$$\int_{\underline{t}}^{t} w(\theta(t'))\sigma(t') \,\mathrm{d}F(t') \geq \int_{\underline{t}}^{t} \theta(t')\sigma(t') \,\mathrm{d}F(t'), \forall t \in [\underline{t}, \bar{t}] \qquad \text{(SOSD)}, \tag{7}$$

$$\int_{\underline{t}}^{\bar{t}} w(\theta(t))\sigma(t) \,\mathrm{d}F(t) = \int_{\underline{t}}^{\bar{t}} \theta(t)\sigma(t) \,\mathrm{d}F(t) \qquad \text{(BP)}. \tag{8}$$

Now I solve a relaxed program subject to constraints (4), (5), and (8).

Proposition 1. *There exists a threshold* $t_0^* \in T$ *such that* $\sigma^*(t) = 1$ *if and only if* $t \geq t_0^*$*.The revenue-maximizing* $\theta^*(t)$ *is given by*

$$\begin{cases} c_\theta(\theta^*(t), t) = 1 + \frac{1-F(t)}{f(t)} c_{t\theta}(\theta^*(t), t), & \textit{if } t \geq t_0^*, \\ \theta^*(t) = 0, & \textit{otherwise.}^{8} \end{cases}$$

[8] Under the relaxed Assumption (A5'), $\theta^*(t) = 0$ is replaced by $c(\theta^*(t), t) = 0$.

Proof. Using integration by parts, we solve the relaxed problem as follows.

$$\begin{aligned}\int_{\underline{t}}^{\bar{t}} \sigma(t)P \,\mathrm{d}F(t) &= \int_{\underline{t}}^{\bar{t}} [w(\theta(t)) - c(\theta(t),t)]\sigma(t) - U(t)\,\mathrm{d}F(t) \\ &= \int_{\underline{t}}^{\bar{t}} [\theta(t) - c(\theta(t),t)]\sigma(t) - U(t)\,\mathrm{d}F(t) \quad \text{(by Bayesian plausibility)} \\ &= \int_{\underline{t}}^{\bar{t}} \left([\theta(t) - c(\theta(t),t)]\sigma(t) + \int_{\underline{t}}^{t} \sigma(x)c_t(\theta(x),x)\,\mathrm{d}x \right) \mathrm{d}F(t) - U(\underline{t}) \\ &= \int_{\underline{t}}^{\bar{t}} [\theta(t) - c(\theta(t),t) + \frac{1-F(t)}{f(t)} c_t(\theta(t),t)]\sigma(t)\,\mathrm{d}F(t) - U(\underline{t})\end{aligned}$$

(IR) binds at $U(\underline{t}) = 0$. Pointwise maximization gives the optimal $\theta^*(t) = \arg\max_\theta\{\theta - c(\theta,t) + \frac{1-F(t)}{f(t)} c_t(\theta,t)\}$ and that $\sigma^*(t) = 1$ if and only if $y(t) \equiv \max_\theta\{\theta - c(\theta,t) + \frac{1-F(t)}{f(t)} c_t(\theta,t)\} \geq 0$. Because $y'(t) = ([\frac{1-F(t)}{f(t)}]' - 1)c_t(\theta^*(t),t) + \frac{1-F(t)}{f(t)} c_{tt}(\theta^*(t),t) > 0$ by Assumptions (A1) and (A4), a unique threshold $t_0^* = \{t_0 \in T : y(t) \geq 0$ if and only if $t \geq t_0\}$ exists such that $\sigma^*(t) = 1$ if and only if $t \geq t_0^*$.

Then, we need to verify that $\theta^*(t)$ is increasing. This is guaranteed by assumptions on third-order derivatives (A3) and the monotone hazard rate (A4). Moreover, $\theta^*(t)$ is strictly increasing when $t \geq t_0^*$.[9]

Finally, it remains to show that $\theta^*(t)$ can be implemented by a feasible interim posterior mean $w^*(\theta)$, which will be shown in Proposition 2. □

Remark 1. Once Bayesian plausibility ($\mathbf{E}[w(\theta(t))] = \mathbf{E}[\theta(t)]$) is substituted into the certifier's objective, the model essentially becomes a simplified version of the Laffont-Tirole cost-reimbursement model (see Laffont and Tirole, 1993, Chapter 1) with exclusion. The connection is particularly obvious in the additive case $\theta = t + e$, where effort e incurs a cost $c(e)$. Because types and (effort) costs determine quality deterministically, this can also be viewed as a "false moral hazard" model where inefficiency arises from pure adverse selection.

Recall that μ a distribution over $\Theta^* \equiv \{\theta^*(t) \in \Theta : \sigma^*(t) = 1,\, t \in T\}$. By (IC) and Proposition 1, $\mu(\theta^*(t)) = \bar{F}(t) = \frac{F(t)-F(t_0)}{1-F(t_0)}$ and $\operatorname{supp}(\mu) = [\underline{\theta}, \bar{\theta}] = [\theta^*(t_0^*), \theta^*(\bar{t})]$. Hence, the optimal price and the interim posterior mean can be expressed as an expectation over μ.

Proposition 2. *The optimal price is*

$$P^* = \mathbf{E}\left[\theta^*(t) - c(\theta^*(t),t) + \frac{1-F(t)}{f(t)} c_t(\theta^*(t),t) \mid t \geq t_0^*\right] = \mathbf{E}_{\theta\sim\mu}\left[\theta - \frac{1-\mu(\theta)}{\mu'(\theta)} c_\theta(\theta,(\theta^*)^{-1}(\theta))\right] - c(\underline{\theta}, t_0^*),$$

[9] For the additive case where $c(\theta,t) = c(\theta - t)$, $c''' \geq 0$ guarantees $\theta'(t) > 0$.

and the interim posterior mean is

$$w^*(\theta) = \begin{cases} \int_{\underline{\theta}}^{\theta} c_\theta(u, (\theta^*)^{-1}(u))\,\mathrm{d}u + \mathbf{E}_{\theta\sim\mu}\left[\theta - \frac{1-\mu(\theta)}{\mu'(\theta)} c_\theta(\theta, (\theta^*)^{-1}(\theta))\right], & \textit{if}\,\theta \in [\underline{\theta}, \bar{\theta}], \\ w_\varnothing = 0, & \textit{otherwise.} \end{cases}$$

Proof. First, because $\mu(\theta^*(t)) = \bar{F}(t) = \frac{F(t)-F(t_0)}{1-F(t_0)}$ and $\theta^{*\prime}(t) > 0$,

$$\begin{aligned} P^* = \frac{\int_{t_0^*}^{\bar{t}} P^* \,\mathrm{d}F(t)}{1-F(t_0^*)} &= \int_{t_0^*}^{\bar{t}} [\theta^*(t) - c(\theta^*(t), t) + \frac{1-F(t)}{f(t)} c_t(\theta^*(t), t)]\,\mathrm{d}\bar{F}(t) \\ &= \int_{t_0^*}^{\bar{t}} [\theta^*(t) - c(\theta^*(t), t) - \frac{1-F(t)}{f(t)} c_\theta(\theta^*(t), t)\theta^{*\prime}(t)]\,\mathrm{d}\bar{F}(t) - c(\underline{\theta}, t_0^*) \\ &= \int_{\underline{\theta}}^{\bar{\theta}} [\theta - c(\theta, \theta^{*-1}(\theta)) - \frac{1-\mu(\theta)}{\mu'(\theta)} c_\theta(\theta, (\theta^*)^{-1}(\theta))]\,\mathrm{d}\mu(\theta) - c(\underline{\theta}, t_0^*), \end{aligned}$$

where the second line follows from

$$c(\theta^*(t), t) - c(\underline{\theta}, t_0^*) = \int_{t_0^*}^{t} c_\theta(\theta^*(\tilde{t}), \tilde{t})\theta^{*\prime}(\tilde{t})\,\mathrm{d}\tilde{t} + \int_{t_0^*}^{t} c_t(\theta^*(\tilde{t}), \tilde{t})\,\mathrm{d}\tilde{t}. \tag{9}$$

Then, $w^{*\prime}(\theta) = c_\theta(\theta^*(t), t) = 1 + \frac{1-F(t)}{f(t)} c_{t\theta}(\theta^*(t), t)$ and $w(\underline{\theta}) - c(\underline{\theta}, t_0^*) = P^*$ imply

$$w^*(\theta) = \int_{\underline{\theta}}^{\theta} c_\theta(u, (\theta^*)^{-1}(u))\,\mathrm{d}u + \mathbf{E}_{\theta\sim\mu}\left[\theta - \frac{1-\mu(\theta)}{\mu'(\theta)} c_\theta(\theta, (\theta^*)^{-1}(\theta))\right].$$

Finally, because $w^{*\prime}(\theta) = 1 + \frac{1-F(t)}{f(t)} c_{t\theta}(\theta^*(t), t) < 1$, its feasibility follows from Corollary 1.1.

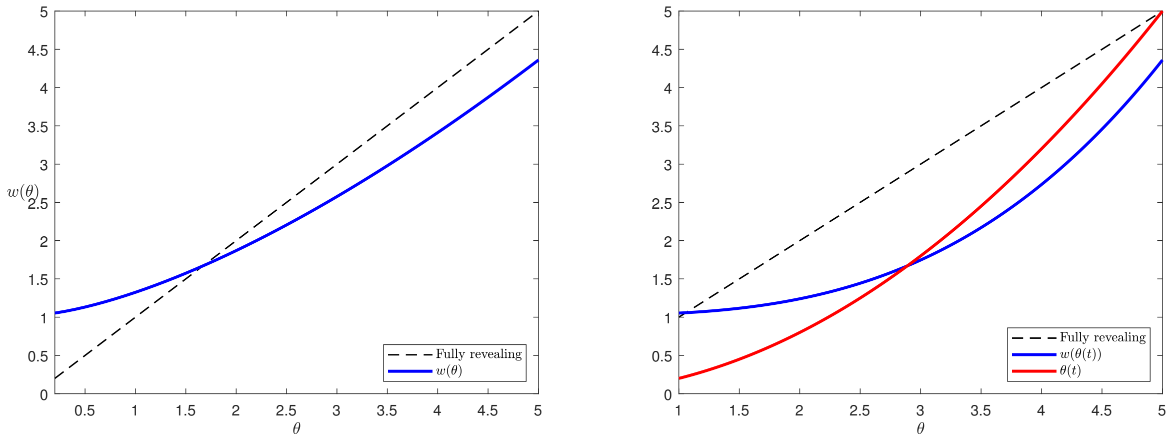
□

Alternatively, one can prove feasibility by constructing the optimal test π^* à la Albano and Lizzeri (2001, Proposition 5) that induces $w^*(\theta)$. The optimal test has a minimum standard and mixes between full revelation and an (almost) uninformative signal.

Proposition 3. *Define $\hat{\theta}$ as the (unique) fixed point of $w^*(\theta)$, and $\beta(\theta) = \frac{\theta - w^*(\theta)}{\theta - \hat{\theta}}$ for $\theta \neq \hat{\theta}$ and $\beta(\hat{\theta}) = 0$. The optimal test (π^*, S) is given by $S = [\underline{\theta}, \bar{\theta}] \cup \{$pass, fail$\}$ and*

$$\pi^*(\theta) = \begin{cases} \theta & \textit{w.p. } 1 - \beta(\theta) \\ \textit{pass} & \textit{w.p. } \beta(\theta) \end{cases}, \quad \forall \theta \in [\underline{\theta}, \bar{\theta}],$$

$$\pi^*(\theta) = \textit{fail w.p. } 1, \quad \forall \theta \notin [\underline{\theta}, \bar{\theta}].$$



(a) $w^*(\theta)$

(b) $w^*(\theta^*(t))$ and $\theta^*(t)$

Figure 2: Interim posterior mean and quality when $c(\theta,t)=\theta^2/2t$ and $t\sim$ Unif$[1,5]$

Proof. Because $w^{*\prime}(\theta)=c_\theta(\theta^*(t),t)\in(0,1)$, $w^*(\theta)$ has a unique fixed point $\hat{\theta}$ such that $w(\theta)\geq\theta$ if and only if $\theta\geq\hat{\theta}$. Therefore, $\beta(\theta)\equiv\frac{\theta-w^*(\theta)}{\theta-\hat{\theta}}\in[0,1]$. Finally, (BP) implies $\mathbf{E}[\theta|\text{pass}]=\hat{\theta}$ and therefore $w^*(\theta)=\beta(\theta)\hat{\theta}+(1-\beta(\theta))\theta=\mathbf{E}_{s\sim\pi^*(\theta)}[\mathbf{E}[\tilde{\theta}|s]]$. □

Remark 2. The construction of π^* is only possible because $w^{*\prime}(\theta)=c_\theta(\theta^*(t),t)\leq 1$. Otherwise, if the solution to the relaxed problem is such that $c_\theta(\theta^*(t),t)>1$, such construction is impossible, and the interim posterior mean $w^*(\theta)$ we need to implement $\theta^*(t)$ may not be feasible.

3 Discussion

I revisit AL's revenue maximization problem using an information design approach and working with interim posterior mean. It turns out that in this particular case, the problem reduces to a classic mechanism design problem because Bayesian plausibility is sufficient for feasibility. In other words, the solution to the relaxed problem (subject to Bayesian plausibility) happens to be feasible—there exists an experiment that induces the posterior mean. This is true because the optimal policy involves underinvestment in quality compared to the agent's first-best (under a fully revealing test), i.e., $w^{*\prime}(\theta)=c_\theta(\theta^*(t),t)\leq 1$.

With general objective functions, the solution to the relaxed problem is not necessarily feasible, so we need to add the condition (SOSD) in Theorem 1 to the constraints to solve the problem. For example, if the certifier is a regulator who maximizes a weighted sum of the certification fee (corresponding to firms' profits) and the average quality in the market (Bizzotto and Harstad, 2023), that is, $\mathbf{E}[(\lambda P + (1-\lambda)\theta) \cdot \mathbf{1}[\theta \text{ takes the test}]] = \int_{t_0}^{\bar{t}} \theta - \lambda c(\theta, t) + \lambda \frac{1-F(t)}{f(t)} c_t(\theta, t)\, \mathrm{d}F(t)$, where $\lambda \in (0,1]$. Solving the relaxed problem yields $c_\theta(\theta, t) = \frac{1}{\lambda} + \frac{1-F(t)}{f(t)} c_{\theta t}(\theta(t), t)$. When $\lambda < 1$, the optimal policy $\theta^*(t)$ entails $c_\theta(\theta^*(t), t) > 1$, so the interim posterior mean $w^*(\theta)$ that implements the policy may be infeasible. A similar example is an information certifier (teacher) maximizing agents' (students') average quality (human capital) without internalizing their cost of quality provision, while the market values the agents' quality (cf. Zubrickas, 2015).

References

Albano, Gian Luigi and Alessandro Lizzeri (2001), "Strategic Certification and Provision of Quality." *International Economic Review*, 42, 267–283. 10.1111/1468-2354.00110. [1, 2, 9]

Ali, S. Nageeb, Nima Haghpanah, Xiao Lin, and Ron Siegel (2021), "How to Sell Hard Information." *The Quarterly Journal of Economics*, 137, 619–678. 10.1093/qje/qjab024. [2]

Bizzotto, Jacopo and Bård Harstad (2023), "The Certifier for the Long Run." *International Journal of Industrial Organization*, 87, 102920. 10.1016/j.ijindorg.2023.102920. [11]

Border, Kim C. (1991), "Implementation of Reduced Form Auctions: A Geometric Approach." *Econometrica*, 59, 1175–1187. 10.2307/2938181. [5]

Border, Kim C. (2007), "Reduced Form Auctions Revisited." *Economic Theory*, 31, 167–181. 10.1007/s00199-006-0080-z. [5]

Doval, Laura and Alex Smolin (2021), "Information Payoffs: An Interim Perspective." September, Working paper. https://www.tse-fr.eu/sites/default/files/TSE/documents/doc/wp/2021/wp_tse_1247.pdf. [2]

Kleiner, Andreas, Benny Moldovanu, and Philipp Strack (2021), "Extreme Points and Majorization: Economic Applications." *Econometrica*, 89, 1557–1593. 10.3982/ECTA18312. [5]

Laffont, Jean-Jacques and Jean Tirole (1993), *A Theory of Incentives in Procurement and Regulation*: MIT Press. [2, 8]

Maskin, Eric and John Riley (1984), "Optimal Auctions with Risk Averse Buyers." *Econometrica*, 52, 1473–1518. 10.2307/1913516. [5]

Matthews, Steven A. (1984), "On the Implementability of Reduced Form Auctions." *Econometrica*, 52, 1519–1522. 10.2307/1913517. [5]

Rayo, Luis (2013), "Monopolistic Signal Provision." *The B.E. Journal of Theoretical Economics*, 13, 27–58. 10.1515/bejte-2012-0003. [4]

Rodina, David (2020), "Information Design and Career Concerns." *CRC TR 224 Discussion Paper Series*. [2]

Saeedi, Maryam and Ali Shourideh (2020), "Optimal Rating Design." September, arXiv: 2008.09529. 10.48550/arXiv.2008.09529. [1, 2, 4]

Saeedi, Maryam and Ali Shourideh (2022), "Optimal Rating Design." Unpublished. https://www.andrew.cmu.edu/user/msaeedi/optimal_rating_design.pdf. [2, 4]

Zubrickas, Robertas (2015), "Optimal Grading." *International Economic Review*, 56, 751–776. 10.1111/iere.12121. [11]