Test design under voluntary participation

(This version: March 18, 2017)

Frank Rosar[a]

[a]*Department of Economics, University of Bonn, Lennéstr. 37, 53113 Bonn, Germany. Tel.: + 49 228 73 6192. Fax: + 49 228 73 7940. E-mail: email@frankrosar.de.*

Abstract

An agent who is imperfectly informed about his binary quality can voluntarily participate in a test that generates a public signal. I study the design of the test that allows for optimal learning of the agent's quality when the agent strives for a high perception of his quality but is averse towards perception risk. For a large class of reduced-form utility functions that reflect these properties, the optimal test is binary and not subject to false positives. I uncover the forces that drive this result and show how the problem with endogenous participation can be transformed into a problem to that the concavification approach from the Bayesian persuasion literature applies. Furthermore, for a non-reduced version of my model where the designer estimates the agent's quality but suffers either more from false positives or from false negatives, I show that the same type of test is optimal.

JEL classification: D02, D82, D83

Keywords: test design; Bayesian learning; concavification; false positive; asymmetric information; voluntary participation

1. Introduction

Technological advances have in many areas improved the capability to reveal information about an agent through testing procedures. In contrast to signaling and screening mechanisms that were subject of intense study in the past, such procedures are in many environments capable of revealing information that goes beyond what the agent knows himself. However, either because a regulating agency saw a need to protect the agent or for some technical reason, testing often requires the agreement of the agent. This article studies the design of the testing procedure in such environments.

To fix ideas, think of the following examples: A genetic test can determine whether a patient will develop a certain illness, but testing necessitates for legal reasons the formal agreement of the patient. Big data analysis can be used to assess whether a bank is viable in a scenario that is unknown to the bank, but the bank must grant access to the necessary data. Polygraphs and fMRI scanners can be used to judge about the eligibility of a potential employee, but using them requires the candidate's agreement. Similar problems arise also in more classical environments: The judge in an inquisitorial legal system controls the generation of information, but she requires an imperfectly informed prosecutor to bring cases to court.[1] Product testing can reveal information about the safety of a product, but the imperfectly informed producer has to apply for certification.[2]

In this kind of applications, the agent's quality is either good (i.e., he is healthy/viable/eligible/. . .) or bad (i.e., he is ill/nonviable/ineligible/. . .), but he is often only imperfectly informed about his true quality. A test is any device that, if used, issues a public signal that depends directly on the agent's true quality.[3] An accurate test perfectly reveals the agent's quality. It generates a *positive* result if the agent is good and a *negative* result if he is bad.[4] Any binary test that is not accurate is subject to false positives, false negatives, or both. Similar interpretations can be employed for non-binary tests. After the test is designed by the principal, the agent decides upon participation. The principal uses the generated test result or the agent's decision not to participate to draw a Bayesian inference about the agent's quality. I consider a version of the model with stylized reduced-form utility specifications that depend directly on this inference. In addition, I discuss a non-reduced version where the principal has to take a decision and a class of utility functions with standard properties that is interesting from an applied perspective.

In general, the principal strives for uncovering information about the agent's quality. She

[1] Kamenica and Gentzkow (2011) discuss in their leading example the complementary problem where information generation is controlled by an uninformed prosecutor.

[2] See Harbaugh and Rasmusen (2016) for a deeper discussion of the certification application.

[3] I am interested in the case where the generated information is always learned by the principal. Matthews and Postlewaite (1985) compare an agent's incentive to generate information through an accurate test for voluntary and mandatory disclosure of the generated information. See also Farhi et al. (2013).

[4] Note that I fix the wording such that "positive" refers to the result that is better for the agent. For some applications, the wording is interchanged. E.g., a positive HIV test is the result that is worse for the agent.

dislikes false positives and false negatives. However, depending on the application, she may dislike one type of error more than the other. The agent cares about how his quality is perceived by the principal. He benefits when the perception is higher but he is in many problems averse towards perception risk. For instance, such an aversion may be due to the Hirshleifer effect or it may arise for psychological reasons.[5]

What makes the test design problem interesting is that there is no full unraveling under an accurate test. The agent faces a trade-off: non-participation signals unfavorable private information about his quality but participation comes along with a perception risk. The usual unraveling argument fails as the imperfection of the agent's private information limits how adverse the inference associated to non-participation can be. Because of his aversion towards perception risk, the agent with the worst possible private information strictly prefers to be perceived as having the worst possible private information to his quality being perfectly revealed by an accurate test. The relevance of this kind of effect was already observed by Stiglitz (1975):

> "If individuals are very risk averse and not perfectly certain of their abilities, then they may prefer to be treated simply as average rather than to undertake the chance of being screened and labelled below average."

Thus, even if there are no technical constraints that limit the test accuracy and accuracy is costless, the principal may have an incentive to choose an inaccurate test to foster participation.

This gives rise to a number of questions: Is optimal learning achieved through an accurate test? If not, what is it optimal to test for? Is the optimal test subject to false positives, false negatives or both? How is this affected by whether the principal suffers more from false positives or from false negatives?

My article has two main contributions. The first contribution is theoretical. I use the version of my model with reduced-form utility specifications to uncover the forces that drive the optimal test design. For a large class of reduced-form utility functions that reflect the principal's interest in learning and perception risk-aversion on the agent's side, I find that a single test structure is optimal: the optimal test is binary and not subject to false positives.

In the literature on Bayesian persuasion (Kamenica and Gentzkow, 2011) the sender chooses a Bayes plausible distribution of beliefs to maximize her value function. Concavification methods can be used to solve this problem (see, e.g., Aumann and Maschler, 1995). Such an approach cannot be (directly) applied to my problem as two technical difficulties arise. First, the principal cannot completely control the distribution of beliefs about the agent's quality. This distribution is jointly determined by her test design and the agent's endogenous participation behavior. Second, as the agent is privately informed, two different distributions of posterior

[5]According to Hirshleifer (1971), such an aversion can arise because the revelation of public information takes away hedging opportunities. Moreover, in medical contexts, learning aversion arises often for psychological reasons. An indicator for such an aversion is that participation rates in medical tests are often low even when testing is costless. See Lyter et al. (1987) and Hull et al. (1988).

beliefs are relevant instead of one.[6] I show that it is possible to transform the problem in a way such that the concavification approach applies nevertheless.

The second contribution is more applied. I discuss the non-reduced version of my model where the principal has to estimate the agent's quality but suffers either more from a high estimate when it turns out that the agent is bad or from a low estimate when it turns out that he is good. That is, either false positives or false negative are more costly for the principal. Expectedly, tests that are not subject to false positives perform well when the principal suffers more from false positives; more surprisingly, such tests turn out to be also optimal when the principal suffers more from false negatives.

2. A model of test design with voluntary participation

2.1. The reduced model

A principal (she) wants to learn about the quality $\omega \in \Omega \equiv \{g, b\}$ of an agent (he). The agent's quality is either good ($\omega = g$) or bad ($\omega = b$), but the agent is only imperfectly informed about his quality. He knows that he is good with probability θ and bad with probability $1 - \theta$. The agent knows the value of θ whereas the principal knows only its distribution. θ is drawn from a cumulative distribution function $F(\theta)$ that admits a strictly positive density $f(\theta)$ on the support $\Theta \equiv [\underline{\theta}, \overline{\theta}]$ with $0 < \underline{\theta} < \overline{\theta} < 1$.[7] The assumption on the support of θ describes my notion of imperfection of private information. As $\underline{\theta} < \overline{\theta}$, the agent is endowed with some meaningful information about his quality. As $\underline{\theta} > 0$, he is never certain to be bad.[8]

A test π consists of a finite realization space S and a family of distributions $\{\pi(\cdot|\omega)\}_{\omega\in\Omega}$ over S. I can restrict without loss of generality attention to informative tests (i.e., tests with $\pi(s|g) \neq \pi(s|b)$ for some $s \in S$) that are minimal (i.e., tests with $\pi(s|g) + \pi(s|b) > 0$ for all $s \in S$). I call a test perfectly informative if for each $s \in S$ either $\pi(s|g) = 0$ or $\pi(s|b) = 0$.

The timing is as follows: First, the principal chooses a test π. Second, nature draws θ and ω. Third, the agent observes the principal's test choice and his private signal θ. He decides then between participating in the test ($x = Y$) and not participating ($x = N$). Fourth, if the agent decided to participate, nature draws a test result $s \in S$ according to $\pi(\cdot|\omega)$. I assume that the test result s and the agent's private signal θ are independently distributed conditional on ω. Fifth, either the agent's non-participation decision or the generated test result becomes publicly observable. The principal forms a Bayesian belief about the agent's quality and she takes an action $a \in A$. Finally, payoffs realize. The principal's payoff $v(a, \omega)$ depends on her

[6] In this respect, my problem is related to the Bayesian persuasion problem with heterogeneous priors in Alonso and Câmara (2016a).

[7] An extension of my analysis to the case where the distribution of θ includes atoms and holes is straightforward but it would require a much more cumbersome notation.

[8] By assuming $\underline{\theta} > 0$, I restrict attention to the interesting cases. The test design problem is trivial for $\underline{\theta} = 0$. I will discuss this in Subsection 4.1. The assumption that $\overline{\theta} < 1$ is not crucial but it simplifies the exposition of my analysis. See Weiss (1983) for a similar notion of imperfection of private information.

action and the agent's quality whereas the agent's payoff $u(a)$ depends only on the principal's action.[9,10,11]

Since quality is binary, I can identify a belief about the agent's quality with a probability $\mu \in [0,1]$ with that the principal believes the agent to be good. I will refer to this probability as the principal's quality perception. When the principal takes the action $a^*(\mu)$, reducing the last stage of the game yields payoffs $\hat{v}(\mu) = \mathbb{E}_\mu[v(a^*(\mu), \omega)]$ and $\hat{u}(\mu) = \mathbb{E}_\mu[u(a^*(\mu))]$.

My main results will shed light on the optimal test design for a class of reduced problem where I directly impose assumptions on $\hat{v}(\mu)$ and $\hat{u}(\mu)$. I assume that $\hat{u} : [0,1] \to \mathbb{R}$ is a smooth function with $\hat{u}' > 0$, $\hat{u}'' < 0$ and $\hat{u}''' \geq 0$, and that $\hat{v} : [0,1] \to \mathbb{R}$ is a smooth function with $\hat{v}'' > 0$ and $\hat{v}''' \geq 0$. I will give an example of a non-reduced problem that implies these properties in Subsection 2.3.

The assumptions on the first two derivatives of the agent's utility function $\hat{u}$ are essential for the problems in that I am interested in. They capture that the agent likes his quality to be perceived as better ($\hat{u}' > 0$) but dislikes perception risk ($\hat{u}'' < 0$). The assumption on the third derivative ($\hat{u}''' \geq 0$) is useful for getting clear-cut results. Intuitively, it reflects an aversion to downside risk in the sense that among any two quality perception lotteries with the same expected perception and the same variance of perception, the agent prefers the lottery with the better downside protection/greater potential for upside gain.[12] All three assumptions are satisfied for a large class of HARA utility functions including quadratic utility, cubic utility, exponential utility/CARA utility, logarithmic utility, and CRRA utility.[13] I provide in Subsection 4.3 a specific (non-trivial) micro-foundation of the assumptions on the agent's reduced form utility functions. The properties derive there (under some conditions) from the assumption that the agent is prudent in his monetary income.[14]

The assumption $\hat{v}'' > 0$ is essential for the problems in that I am interested in. Since updating is Bayesian, the release of any additional information about the agent's quality transforms the posterior quality perception distribution by a mean-preserving spread. The release of any additional information increases the principal's expected payoff if, and only if, $\hat{v}'' > 0$. Thus,

[9] The assumption that $u(a)$ depends only on a is analogous to the assumption in the important special case in Kamenica and Gentzkow (2011) where the sender's payoff depends only on the expected state. It will together with my assumptions on reduced-form utility functions that I introduce below imply that threshold participation behavior occurs (Proposition 1). My main analysis will extend to more general payoff specifications $u(a, \omega)$ as long as they do also induce threshold participation behavior. I will discuss this in more detail in Subsection 5.5.

[10] Implicit in the payoff specification is also that testing and getting tested is costless for the principal and the agent. It will be a direct consequence of my solution approach that this is without loss of generality. See the discussion in Subsections 5.2 and 5.4.

[11] Implicit in these payoff specifications is also that test accuracy is costless for the principal. This assumption allows me to focus on the use of inaccurate tests for strategic reasons. Such an assumption is standard in the information design literature. See Gentzkow and Kamenica (2014) for a notable exception.

[12] For a formal definition of downside risk and its relation to the third derivative, see Menezes et al. (1980).

[13] See Section 1 of the supplementary material for a parametrization of HARA utility and for a graphical illustration of the part of the parameter space to that my analysis applies.

[14] See Kimball (1990) for a motivation of prudence in financial applications.

$\hat{v}'' > 0$ captures that the principal is interested in learning. Like for the agent, the assumption on the third derivative ($\hat{v}''' \geq 0$) reflects an aversion to downside risk and it is useful for getting clear-cut results.[15] I will explain in Subsection 4.3 how I can use specific structure to extend my analysis to cases where this assumption is violated. An interesting special case that fits my assumptions is $\hat{v}(\mu) = (\mu - \mathbb{E}_F[\theta])^2$. The principal's ex ante expected payoff corresponds then to the variance of the induced quality perception distribution.

2.2. Solution concept

My solution concept is an adaptation of principal-preferred Perfect Bayesian Equilibrium to my reduced-form framework.

Updating. Given a test π and a participation strategy $\chi : \Theta \to \{N, Y\}$, the principal can draw inferences about the agent's quality from two sources of information. First, she can use the information revealed by the agent's participation decision x to update her belief about the agent's private signal. This allows her to form an updated quality perception $\mu_x \in \Theta$.[16] Second, when the agent participates, she can use the information revealed by the test result s to draw additional inferences. This leads to a quality perception $\mu_s \in [0, 1]$.

From the principal's and from the agent's interim perspective (after the agent chose $x = Y$ but before s is observed), the test leads to a distribution of quality perceptions. I denote a distribution of quality perceptions by τ.[17] The distribution τ faced by the principal differs from the distribution τ_θ faced by the agent with private signal θ as the agent is better informed. I say π and χ induce the distribution τ and the family of distributions $\{\tau_\theta\}_{\theta \in \Theta}$ if there exists some $\mu_Y \in \Theta$ that is consistent with χ[18] such that

$$\mu_s = \frac{\pi(s|g)\mu_Y}{\pi(s|g)\mu_Y + \pi(s|b)(1 - \mu_Y)} \text{ for all } s, \tag{1}$$

$$\text{Supp}(\tau) = \text{Supp}(\tau_\theta) = \{\mu_s\}_{s \in S} \text{ for all } \theta,$$

$$\tau(\mu) = \sum_{s:\mu_s = \mu} (\pi(s|g)\mu_Y + \pi(s|b)(1 - \mu_Y)) \text{ for all } \mu, \text{ and} \tag{2}$$

[15] See Gossner and Steiner (2016) for a micro-foundation of such third derivative assumptions. They identify non-reduced models that imply $\hat{v}''' < 0$; non-reduced models that imply $\hat{v}''' > 0$ can be identified in a similar fashion.

[16] Technically, a belief about the agent's private signal is an element of $\Delta(\Theta)$ that can be described by a cumulative distribution function F_x. Any F_x implies a quality perception $\mu_x = \mathbb{E}_{F_x}[\theta]$ that lies in Θ. Only this implied quality perception is relevant for my analysis.

[17] Technically, a belief about the agent's quality is a distribution on Ω (i.e., an element of $\Delta(\Omega)$) and the test induces a distribution on such distributions (i.e., an element of $\Delta(\Delta(\Omega))$). Since I can identify the former by a quality perception μ, I can identify the latter by a distribution of quality perceptions τ.

[18] I.e., if χ prescribes that the decision x is taken with positive probability, then $\mu_x = \mathbb{E}_F[\theta|\chi(\theta) = x]$; if x prescribes that the decision x is taken with probability zero, μ_x can assume any value in Θ.

$$\tau_\theta(\mu) \quad = \quad \sum_{s:\mu_s=\mu} (\pi(s|g)\theta + \pi(s|b)(1-\theta)) \text{ for all } \mu \text{ and all } \theta. \tag{3}$$

(1), (2) and (3) imply that τ_θ depends only through τ on the test design:

$$\tau_\theta(\mu) = \left(\frac{\theta}{\mu_Y}\mu + \frac{1-\theta}{1-\mu_Y}(1-\mu)\right)\tau(\mu).^{19} \tag{4}$$

In need thus only to specify which distribution τ is induced. I can then infer from this the family of distributions $\{\tau_\theta\}_{\theta\in\Theta}$.

Participation. I say χ is induced by π if there exists a distribution τ that is induced by π and χ and if there exists some $\mu_N \in \Theta$ that is consistent with χ such that

$$\chi(\theta) = Y \text{ implies } \mathbb{E}_{\tau_\theta}[\hat{u}(\mu)] \geq \hat{u}(\mu_N), \text{ and}$$

$$\chi(\theta) = N \text{ implies } \mathbb{E}_{\tau_\theta}[\hat{u}(\mu)] \leq \hat{u}(\mu_N).$$

Test design. The principal chooses a test π and a participation strategy χ that is induced by π to maximize her ex ante expected payoff

$$\mathbb{P}_F[\chi(\theta) = Y] \cdot \mathbb{E}_\tau[\hat{v}(\mu)] + \mathbb{P}_F[\chi(\theta) = N] \cdot \hat{v}(\mu_N). \tag{5}$$

Notice that (5) is completely pinned down by π and χ: The distribution τ that is induced π and χ may only be subject to a degree of freedom when χ prescribes participation with probability zero. μ_N may only be subject to a degree of freedom when χ prescribes non-participation with probability zero. It will follow from my analysis that the principal's maximization problem has a solution. I will call any test that is part of such a solution optimal.

2.3. A non-reduced example: Estimation with symmetric error cost

Before I proceed with the analysis of my reduced-form model, I provide a non-reduced example that reduces to an instance of my reduced-form model. Suppose the agent's quality ω determines whether his value (e.g., his productivity or the value of his product) is low or high, say $a_b^* \in \mathbb{R}_+$ or $a_g^* > a_b^*$. The principal is a certifier whose task it is to assess the agent's quality. She chooses an action $a \in \mathbb{R}_+$ that can be interpreted as an estimate of the agent's expected value. At some stage in the future, the agent's quality comes to light and the principal suffers a quadratic loss from inaccuracies in her estimate (e.g., due to reputation concerns):

$$v(a,\omega) = \begin{cases} -(a-a_g^*)^2 & \text{if } \omega = g \\ -(a-a_b^*)^2 & \text{if } \omega = b \end{cases}.$$

The agent benefits when he obtains a higher estimate (e.g., because he is paid the estimate by a third party as a wage or a price), but he is risk-averse and prudent: $u'(a) > 0$, $u''(a) < 0$ and

[19] A formal proof of this relationship can be found in the proof to Lemma 2 (a) where I use this property for the first time. See Alonso and Câmara (2016a) for a similar property.

$u'''(a) \geq 0$.

I have then $\mathbb{E}_\mu[v(a,\omega))] = -\mu(a-a_g^*)^2 - (1-\mu)(a-a_b^*)^2$. This implies that the principal's optimal estimate corresponds to the Bayesian estimate of the agent's value given the quality perception μ: $a^*(\mu) = \mu a_g^* + (1-\mu)a_b^*$. This gives rise to the reduced-form utility functions $\hat{u}(\mu) = u(a_b^* + (a_g^* - a_b^*)\mu)$ and $\hat{v}(\mu) = -\mu(1-\mu)(a_g^* - a_b^*)^2$ that represent an instance of my reduced model. In fact, the principal's reduced-form utility function induces the same preferences over tests as $\hat{v}(\mu) = (\mu - \mathbb{E}_F[\theta])^2$. Thus, the principal strives in this non-reduced example to maximize the variance of the posterior quality perception distribution.[20] I will discuss the more involved non-reduced problem where one type of error is relatively more costly for the principal in Subsection 4.3.

3. An illustrative example: Updating and trade-offs for binary tests that are not subject to false positives

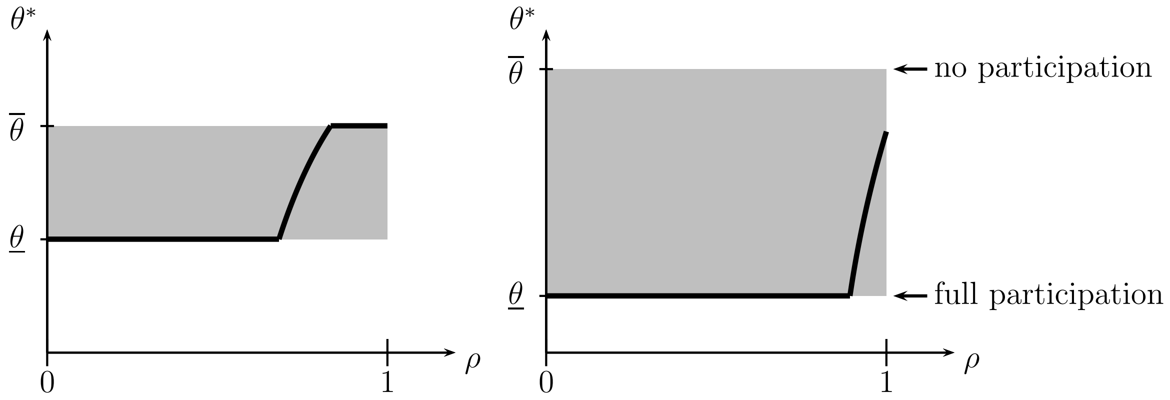
Binary tests that are not subject to false positives. For any binary test, I call the test result that is relatively more likely to be obtained by a good agent (bad agent) the positive result (negative result). A test is subject to false positives (false negatives) if a bad agent (good agent) can obtain the positive result (negative result). Any binary test with $\pi(s|b) = 0$ for some $s \in S$ is not subject to false positives. The positive result allows then for the inference that the agent is good. A parametrization of this class of tests is given by π_ρ^{NFP} with $S = \{1,2\}$, $\pi_\rho^{\text{NFP}}(1|g) = 1-\rho$, $\pi_\rho^{\text{NFP}}(2|g) = \rho$, $\pi_\rho^{\text{NFP}}(1|b) = 1$ and $\pi_\rho^{\text{NFP}}(2|b) = 0$. I will refer to this class of tests as NFP tests and to the parameter ρ as the test accuracy. $\rho = 0$ describes a completely noninformative test whereas $\rho = 1$ describes a perfectly informative test. The class of binary tests that is not subject to false negatives can be introduced in an analogous fashion. I will refer to this class as NFN tests. In the remainder of this section, I will motivate effects and trade-offs that arise for NFP tests. It will follow from my analysis later on why this class of tests is of particular interest.

Updating. When the principal chooses the test π_ρ^{NFP}, her quality perception is $\mu_2 = 1$ after observing the positive result and

$$\mu_1 = \frac{(1-\rho)\cdot\mu_Y}{(1-\rho)\cdot\mu_Y + 1\cdot(1-\mu_Y)} < 1 \tag{6}$$

with $\mu_Y = \mathbb{E}_F[\theta|\text{"participation"}]$ after observing the negative result. When the test is not used, she can still draw inferences about the agent's quality through his participation behavior: $\mu_N = \mathbb{E}_F[\theta|\text{"non-participation"}]$.

[20] Harbaugh and Rasmusen (2016) measure information generation by applying a loss function to the error of the public's quality estimate. This example shows that the special case of my reduced model where the principal strives to maximize the variance of the posterior quality perception distribution can be interpreted as an adaptation of their utility specification with a quadratic loss function to my setting with a binary quality. Quadratic loss functions are a standard assumption in sender-receiver games.



(a) $\theta \sim \mathrm{U}[1/3, 2/3]$

(b) $\theta \sim \mathrm{U}[1/6, 5/6]$

Figure 1: Test accuracy and induced participation threshold $[\hat{u}(\mu) = -(1-\mu)^2,\ \pi = \pi_\rho^{\mathrm{NFP}}]$

Trade-off associated with the participation decision. If the agent participates when he has the private signal θ, the two test results are generated with the probabilities $\tau_\theta(\mu_2) = \rho\theta$ and $\tau_\theta(\mu_1) = 1 - \rho\theta$. The higher his private signal, the more likely the positive test result is generated. If he does not participate, it does not depend on his private signal how he is perceived. Thus, the agent has a stronger incentive to participate the higher his private signal. Only threshold participation strategies can be induced.

When I denote the participation threshold by θ^*, non-participation allows the principal to draw the inference $\mu_N = \mathbb{E}_F[\theta|\theta \leq \theta^*]$ whereas participation allows for the inference $\mu_Y = \mathbb{E}_F[\theta|\theta \geq \theta^*]$. From the perspective of the agent, participation leads to an expected quality perception of $\mathbb{E}_{\tau_\theta}[\mu] = (1-\rho\theta)\mu_1 + \rho\theta\mu_2$. When the agent has the threshold signal $\theta = \theta^*$, this can be written as

$$\mathbb{E}_{\tau_{\theta^*}}[\mu] = (1-\rho)\mathbb{E}_F[\theta|\theta \geq \theta^*]\frac{1-\rho\theta^*}{1-\rho\mathbb{E}_F[\theta|\theta \geq \theta^*]} + \rho\theta^* \geq \theta^*. \tag{7}$$

Non-participation leads to a certain quality perception of $\mu_N \leq \theta^*$. When the agent has the threshold signal, he is on average perceived as better when he participates but he faces then a perception risk.

Trade-off associated with the test accuracy choice. The test accuracy ρ determines for which private signals the agent will participate. Figure 1 illustrates for two numerical examples how participation decreases (i.e., the participation threshold increases) as the test accuracy increases. To build some intuition for the effect of the test accuracy on the participation threshold, it is illuminative to consider the polar cases where the test is accurate ($\rho = 1$) and very inaccurate (ρ close to 0). Suppose first the test is accurate. Then, from the perspective of the agent with any private signal θ, participation induces a fair quality perception lottery: the quality perception is $\mu = 1$ with probability θ and $\mu = 0$ with probability $1-\theta$. The lottery is non-degenerate as the agent is uncertain about his quality. As the agent is averse to perception risk, participation leads to an expected payoff that is strictly smaller than $\hat{u}(\theta)$. When the agent does not participate, the principal cannot infer anything from his behavior that the agent does not know himself. This limits how bad the principal's perception of the agent's quality can be. The most adverse perception is that the agent has the worst possible private signal $\theta = \underline{\theta}$. As such a perception implies a certain utility of $\hat{u}(\underline{\theta})$, the agent has a strict incentive not to participate when his private signal θ is close to $\underline{\theta}$. Hence, an accurate test does either induce no participation (see Figure 1(a)) or only partial participation (see Figure 1(b)). On the other hand, when the test is very inaccurate, participation imposes almost no perception risk on the agent. However, it allows the agent for any supposed participation threshold θ^* to "significantly" improve how his quality is perceived on average at almost no "cost".[21] As a consequence, there is full participation in any informative but sufficiently inaccurate test (see Figures 1(a) and 1(b)). It follows that participation decreases—at least eventually—as the test accuracy increases. This gives rise to the interpretation that participation can be fostered by making the test subject to errors.[22]

4. Optimal test design

In the preceding section, I imposed an ad hoc restriction on the structure of the tests that I considered. This gave me a one-dimensional test design problem that was suitable for motivating effects. Most interesting is, however, the question of how the optimal test structure does actually look like. This is the main objective of this article. Once I know this structure, the principal's design problem reduces to a much simpler, one-dimensional problem that is like the problem that I discussed in the preceding section.

4.1. Inducible participation behavior

I come first to the question as to whether any given test induces some participation strategy and which strategies can be induced.

Proposition 1 (Only threshold participation is inducible) *Fix any informative test π. (a) There exists a participation strategy χ that is induced by π. (b) If χ is induced by π, then χ is a threshold participation strategy. More specifically, there exists then a $\theta^* \in \Theta$ such that $\chi(\theta) = Y$ if $\theta > \theta^*$ and $\chi(\theta) = N$ if $\theta < \theta^*$.*

Part (a) is not very surprising. The intuition for part (b) is as follows: Given any informative test, test results associated with higher quality perceptions are good news in the sense of

[21]Participation in a very inaccurate test increases the expected quality perception from the perspective of the agent by approximately $\mathbb{E}_F[\theta|\theta \geq \theta^*] - \mathbb{E}_F[\theta|\theta \leq \theta^*]$. The "significance" part of the statement derives from the fact that the minimal value of $\mathbb{E}_F[\theta|\theta \geq \theta^*] - \mathbb{E}_F[\theta|\theta \leq \theta^*]$ is bounded away from zero. Thus, no matter which participation threshold is induced for a given ρ, participation increases the expected quality perception from the perspective of the agent significantly. By contrast, the "cost" that the perception risk imposes on the risk-averse agent vanishes as the test accuracy ρ approaches 0.

[22]In a certification context, De and Nabar (1991) show that a similar trade-off can arise with a perfectly informed, risk-neutral seller who can voluntarily decide to get tested at a fee: an inaccurate testing technology can foster participation relative to an accurate one.

Milgrom (1981). As a higher private signal makes more favorable news more likely in the sense of first-order stochastic dominance, participation is more attractive the higher the agent's private signal. Thus, only threshold participation behavior where the agent participates if, and only if, his private signal is above some threshold $\theta^* \in \Theta$ can be induced.

A consequence of threshold participation behavior is that the inference that the principal draws from any participation decision that is supposed to occur with positive probability has a simple form: $\mu_Y = \mathbb{E}_F[\theta|\theta \geq \theta^*]$ and $\mu_N = \mathbb{E}_F[\theta|\theta \leq \theta^*]$. A specific class of threshold strategies that will be useful in the subsequent proofs is defined by

$$\chi_{\theta^*}(\theta) \equiv Y \text{ if } \theta \geq \theta^* \text{ and } \chi_{\theta^*}(\theta) \equiv N \text{ otherwise.}$$

The next question is which participation thresholds can be induced. I say a threshold θ^* is induced by a test if there exists a threshold strategy with this threshold that is induced by this test. Only the inducement of thresholds $\theta^* \in [\underline{\theta}, \overline{\theta})$ is interesting from a design perspective. For my subsequent analysis it will be useful to distinguish between thresholds that can be induced by a perfectly informative test and ones that can only be induced by tests that are not perfectly informative. I introduce the following notation:

$$\Theta^* \equiv \{\theta^* \in [\underline{\theta}, \overline{\theta})|\text{threshold } \theta^* \text{ is induced by some informative test}\}$$

$$\Theta^{**} \equiv \{\theta^* \in [\underline{\theta}, \overline{\theta})|\text{threshold } \theta^* \text{ is induced by some perfectly informative test}\}$$

The subsequent proposition establishes some useful properties of the sets Θ^* and Θ^{**}. The proof contains also a characterization of Θ^* and Θ^{**} in terms of the primitives of my model.[23] As the exact structure of these sets will not be important for my subsequent analysis, I omit it in the proposition.

Proposition 2 (Inducible thresholds) *(a) The set of all thresholds $\theta^* \in \Theta$ that is induced by some informative test is non-empty and compact. (b) $\underline{\theta} \notin \Theta^{**}$. (c) $\underline{\theta} \in \Theta^* \backslash \Theta^{**}$.*

Because the agent is imperfectly informed, perfect learning of his quality requires full participation in a perfectly informative test. Part (b) establishes that perfect learning is generally impossible (as was already motivated in the introduction). Part (c) has two important implications. First, it implies that some learning is generally possible. Second, it shows that the set of all inducible participation thresholds is strictly enlarged by allowing for tests that are not perfectly informative.

Two remarks on the role of my notion of imperfection of privation information for Proposition 2 are in order: It is crucial for the first implication of Part (c) that the agent has some meaningful private information (i.e., $\underline{\theta} \neq \overline{\theta}$). $\underline{\theta} = \overline{\theta}$ would imply that there is no signaling

[23]Section 2 of the supplementary material contains an illustration of the sets Θ^* and Θ^{**} for an example. The example is useful for building an intuition for what determines the sets Θ^* and Θ^{**}.

motive on the agent's side. The sole effect of participation would be that additional, unbiased information is generated. The risk-averse agent would not be willing to participate in any informative test rendering learning about his quality impossible. It is crucial for Part (b) that the agent's private signal is imperfect in the sense that he is never certain to be bad (i.e., $\underline{\theta} \neq 0$). $\underline{\theta} = 0$ would imply that a full unraveling equilibrium exists for any perfectly informative test.[24] Perfect learning would be possible.

4.2. Optimal test design in the reduced model

The ability to describe which participation thresholds are inducible allows me to tackle the test design problem in two steps. First, I will derive in this subsection the test that optimally induces any given participation threshold $\theta^* \in \Theta^*$. Then, I will briefly discuss in Subsection 5.1 the problem of which participation threshold from Θ^* it is optimal to induce. This two-step approach will convey a better understanding of the underlying effects as it will separate the parts of the derivation that rely on the primitives of the model from those that do not: A single test structure will turn out to be optimal for all primitives of the model whereas the optimal participation threshold will depend on the primitives.

Fix any $\theta^* \in \Theta^*$. If θ^* is inducible by a perfectly informative test (i.e., $\theta^* \in \Theta^{**}$), it is obvious that θ^* is optimally induced by such a test. It remains thus only to study the optimal inducement of thresholds that are not inducible by a perfectly informative test. The subsequent lemma states how this problem can be simplified.

Lemma 1 (The inducement problem) *Fix any $\theta^* \in \Theta^* \backslash \Theta^{**}$ and define $\overline{u}_{\theta^*} \equiv u(\mathbb{E}_F[\theta|\theta \leq \theta^*])$. A test is optimal among all tests that induce the participation threshold θ^* if it solves the following program:*

$$\begin{aligned} \max_{\pi} \quad & \mathbb{E}_{\tau}[\hat{v}(\mu)] \\ \text{s.t.} \quad & \mathbb{E}_{\tau_{\theta^*}}[\hat{u}(\mu)] = \overline{u}_{\theta^*} \\ & \pi \text{ and } \chi_{\theta^*} \text{ induce } \tau \end{aligned}$$

The lemma shows that three intuitive simplifications are without loss of generality. First, I only need to consider the principal's expected payoff conditional on participation. Second, I need only to consider $\mu_N = \mathbb{E}_F[\theta|\theta \leq \theta^*]$ even when full participation shall be induced. Third, only the participation constraint of the agent with the threshold signal is relevant and I can restrict attention to the case where it is binding. This allows me to interpret the simplified problem as a problem between the principal and the "threshold agent".

The program is reminiscent of the Bayesian persuasion problem in Kamenica and Gentzkow (2011) with a constraint added.[25] In the original Bayesian persuasion problem, a sender designs

[24] If $\underline{\theta} = 0$, the worst possible private information is verifiable by an accurate test. The reasoning in the literature on information transmission with hard information applies to this case. See Grossman and Hart (1980), Grossman (1981), Milgrom (1981), and Okuno-Fujiwara et al. (1990).

[25] In the previous working paper version of this article, I employed a less elegant strategy to solve the program in Lemma 1 (see Rosar, 2016). I am grateful to an anonymous referee for pointing out how the solution of the program can be simplified by transforming the problem into a concavification problem.

a signal π to maximize the expectation of her value function $\hat{v}(\mu)$ with respect to the distribution τ of posteriors μ that is induced by the signal.[26] Such problems can be solved by showing that it is possible to maximize $\mathbb{E}_\tau[\hat{v}(\mu)]$ directly over all Bayes plausible distributions of posteriors (i.e., distributions τ with $\mathbb{E}_\tau[\mu] = \mu_Y$) and by applying then concavification methods. The structure of the optimal signal depends on the properties of $\hat{v}$.

The principal in my model corresponds to the sender in the persuasion problem; for a given participation behavior, the test corresponds to the signal; a posterior corresponds to a quality perception; and the principal's reduced-form utility function corresponds to the sender's value function. The additional constraint in my problem does per se not cause any fundamental problems as it can be handle through a Lagrangian approach. What makes my problem more involved is that it depends on two different quality perception distributions: the objective function depends on the quality perception distribution as perceived by the principal (i.e., τ) whereas the participation constraint depends on the quality perception distribution as perceived by the threshold agent (i.e., τ_{θ^*}). The following result allows me to transform the more involved problem into a problem to which the techniques from Kamenica and Gentzkow (2011) can be applied.

Lemma 2 (Transformation of the participation constraint) *Fix any $\theta^* \in \Theta^* \backslash \Theta^{**}$. (a) If π and χ_{θ^*} induce τ, then $\mathbb{E}_{\tau_{\theta^*}}[\hat{u}(\mu)] = \mathbb{E}_\tau[\tilde{u}_{\theta^*}(\mu)]$ with*

$$\tilde{u}_{\theta^*}(\mu) \equiv \left(\frac{\theta^*}{\mu_Y}\mu + \frac{1-\theta^*}{1-\mu_Y}(1-\mu) \right) \hat{u}(\mu) \text{ and } \mu_Y = \mathbb{E}_F[\theta | \theta \geq \theta^*].$$

(b) $\tilde{u}_{\theta^*}(\mu)$ is a smooth function with $\tilde{u}''_{\theta^*} < 0$ and $\tilde{u}'''_{\theta^*} > 0$.

Part (a) is a direct consequence of (4). By Part (b), the transformed reduced-form utility function $\tilde{u}_{\theta^*}$ inherits important properties of the original reduced-form utility function $\hat{u}$. These properties will be crucial for the structure of the optimal test. To determine this structure, consider the Lagrangian with Lagrange multiplier λ on the participation constraint:

$$\mathcal{L}(\tau, \lambda) \equiv \mathbb{E}_\tau[\varphi_{\lambda,\theta^*}(\mu) - \lambda \overline{u}_{\theta^*}] \text{ with } \varphi_{\lambda,\theta^*}(\mu) \equiv \hat{v}(\mu) + \lambda \tilde{u}_{\theta^*}(\mu). \tag{8}$$

By the preceding lemmas, I need to solve the following Lagrangian problem with

$$\max_{\pi} \quad \mathcal{L}(\tau, \lambda) \qquad \text{s.t.} \quad \pi \text{ and } \chi_{\theta^*} \text{ induce } \tau$$

for any $\lambda \geq 0$. The optimal test will maximize a weighted sum of the principal's and the threshold agent's expected utility. By a reasoning analogous to that in the Bayesian persuasion literature, the test design problem has the same maximum as the following one-dimensional

[26] π, τ, μ and $\hat{u}$ are essentially the same objects in Kamenica and Gentzkow (2011) as in my model, but Kamenica and Gentzkow allow these objects to be defined for a non-binary state space Ω.

concavification problem:

$$\begin{aligned} \max_{\tau} \quad & \mathcal{L}(\tau, \lambda) \\ \text{s.t.} \quad & \mathbb{E}_\tau[\mu] = \mu_Y \end{aligned}$$

Since $\varphi'''_{\lambda,\theta^*} > 0$ for all $\lambda > 0$ by $\hat{v}''' \geq 0$ and Lemma 2 (b), it follows that either the distribution that is implied by a completely non-informative test or a binary distribution τ^* with $1 \in \text{Supp}(\tau^*)$ is optimal. Since I know already from the construction of Θ^* that θ^* can be induced by an informative test, the latter must be the case. As any Bayes plausible distribution τ^* with $1 \in \text{Supp}(\tau^*)$ is implied by some NFP test π_ρ^{NFP}, I obtain the following result:

Proposition 3 (Optimal test design in the reduced model) *Any participation threshold $\theta^* \in \Theta^*$ is optimally induced by a binary test that is not subject to false positives $\pi_{\rho^*}^{NFP}$. If $\theta^* \in \Theta^{**}$, the optimal test is perfectly informative (i.e., $\rho^* = 1$). If $\theta^* \in \Theta^* \backslash \Theta^{**}$, the optimal test is subject to false negatives (i.e., $\rho^* \in (0,1)$).*

To build a less technical intuition for the result, consider the case where $\hat{u}''' = 0$. The threshold agent cares then only about his expected value $\mathbb{E}_{\tau_{\theta^*}}[\mu]$ and the variance $\mathbb{V}_{\tau_{\theta^*}}[\mu]$ of the quality perception distribution that he faces. As his expected perception has a positive impact whereas his variance of perception has a negative impact on his expected payoff, expected value and variance must co-move for different distributions τ_{θ^*} for that the participation constraint is satisfied. Intuitively, when the expected perception of the threshold agent increases, there is more pooling between the good and the bad state and thus less information generation. If $\hat{v}''' = 0$, pooling as measured by $\mathbb{E}_{\tau_{\theta^*}}[\mu]$ is exactly what the principal strives to minimize.[27] The question is thus, which test that induces θ^* minimizes the threshold agent's expected perception.

As maximization of the the threshold agent's expected perception corresponds to the minimization of his variance of perception, I need only to identify the test structure that imposes the least variance on the threshold agent. As the threshold agent is good with a below average probability, this is just the test structure that imposes the least variance of perception on a bad agent. As a bad agent is not subject to any perception risk when he participates in a NFP test, such a test structure is optimal.

It remains to argue why a NFP test remains optimal when $\hat{u}''', \hat{v}''' \geq 0$. Principal and threshold agent are then both averse to downside risk/like upside potential (see Menezes et al., 1980). Roughly speaking, there is then an additional effect that increases the principal's objective function and which makes the agent's participation constraint less binding when the support of the quality perception distributions moves to the right. As all pooling occurs already on the negative test result and the highest possible quality perception is associated with the positive test result for a NFP test, this test structure remains optimal.

[27] If $\hat{v}''' = 0$, the principal strives to maximize $\mathbb{V}_\tau[\mu] = \mathbb{E}_\tau[\mu^2] - \mathbb{E}_\tau[\mu]^2$. Since Bayesian plausibility implies that $\mathbb{E}_\tau[\mu]$ is constant, she strives to maximize $\mathbb{E}_\tau[\mu^2]$. By (4), $\mathbb{E}_\tau[\mu^2]$ is a negative linear transformation of $\mathbb{E}_{\tau_{\theta^*}}[\mu]$.

When $\hat{u}''' \geq 0$ or $\hat{v}''' \geq 0$ is violated, there still exists an optimal test that is binary.[28] This is a direct consequence of the fact that I have to solve a one-dimensional concavification problem. However, the optimal test may then be subject to false positives. I show in the next subsection that a test that is not subject to false positives remains optimal for an interesting application for that the third derivative assumptions are sometimes violated. Moreover, if $\tilde{u}'''_{\theta^*} < 0$ and $\hat{v}''' \leq 0$ is both true, it follows immediately from my solution approach that a test that is not subject of false negatives is optimal. Yet as $\hat{u}''' < 0$ does not imply $\hat{u}'''_{\theta^*} < 0$ (see (A.6) in the proof of Lemma 2), it is less obvious for which primitives of the model this would be the case.

The following corollary establishes some properties of the optimal test.

Corollary 1 (Properties of the optimal test) *Fix any $\theta^* \in \Theta^*$ and consider all tests that induce θ^*. The following is true for the test that is optimal from the principal's perspective: (a) It minimizes the expected quality perception of the threshold agent, $\mathbb{E}_{\tau_{\theta^*}}[\mu]$. (b) If the Arrow-Prat measure of absolute risk aversion $-\hat{u}''(\mu)/\hat{u}'(\mu)$ is sufficiently large for all $\mu \in [0,1]$ and $\hat{u}''' = 0$, it minimizes the expected payoff of the agent conditional on that he has any above average signal $\theta^{**} > \mu_Y$, $\mathbb{E}_{\tau_{\theta^{**}}}[\hat{u}(\mu)]$.*

Part (a) does formally establish that the optimal test minimizes the expected perception of the threshold agent even when $\hat{u}''' \neq 0$ or $\hat{v}''' \neq 0$. Harbaugh and Rasmusen (2016) consider a related problem in that the agent is perception risk-neutral. The principal's only instrument to foster participation is to increase the expected perception of the threshold agent. The corollary shows that when the agent is imperfectly informed and perception risk-averse, non-trivial effects are added. The principal can then also foster participation by reducing the threshold agent's perception risk and doing so is better for her than increasing the threshold agent's expected perception.

Part (b) is an auxiliary result that will be useful for discussing what changes when the principal is able to offer a menu of tests (see Subsection 5.3). The result gives sufficient conditions under that the NFP test minimizes the expected payoff of the agent conditional on that he has any above average signal. To get an intuition, recall what the NFP test that induces θ^* does. There are three effects. First, it imposed little perception risk on the agent conditional that he has a below average signal. Second, it implies that information generation works quite well. Third, it implies that high quality perceptions are associated with both test results. Conditional on that the agent has an above average signal, he dislikes the first effect, but he likes the other two. His perception risk would, for instance, be reduced when θ^* was induced by a NFN test. Such a test imposes no perception risk on a good agent and thus the smallest possible perception risk on any agent with an above average signal. By contrast, the agent likes the second effect as it improves how well he is perceived on average and he also likes the inducement

[28]Kamenica and Gentzkow (2011) show in Proposition 4 of their Web Appendix that in their persuasion problem, in that the designer is only subject to Bayesian plausibility, the optimal information structure requires at most as many different signals as there exist states of the world. The optimality of a binary test here shows that a similar result is obtained for my test design problem with an interim participation constraint.

of high quality perceptions because of his aversion to downside risk. Under the condition in the corollary, the first effect dominates the second effect and the third effect is mute.

4.3. Optimal test design in a non-reduced example: Estimation with asymmetric error cost

In Subsection 2.3, I motivated my reduced-form model with a non-reduced quality estimation problem. In this problem, too high estimates (which follow from false positives) and too low estimates (which follow from false negatives) caused symmetric error cost for the principal. I relax now the symmetry assumption. Suppose that the principal's non-reduced payoff is

$$v(a,\omega) = \begin{cases} -\gamma_L(a-a_g^*)^2 & \text{if } \omega = g \\ -\gamma_H(a-a_b^*)^2 & \text{if } \omega = b \end{cases}.$$

with $\gamma_L, \gamma_H > 0$ and $\gamma_L \neq \gamma_H$. To simplify the exposition, I will set henceforth $a_g^* = 1$ and $a_b^* = 0$. As before, a can be interpreted as the principal's estimate of the agent's expected value. If $\gamma_H > \gamma_L$ ($\gamma_H < \gamma_L$), the principal suffers more from too high estimates (too low estimates). This allows for the interpretation that false positives (false negatives) are more costly.[29] Everything else remains as in Subsection 2.3.

For every μ, the principal chooses her estimate a to maximize $\mathbb{E}_\mu[v(a,\omega)] = -\mu\gamma_L(a-1)^2 - (1-\mu)\gamma_H(a-0)^2$. This leads to the optimal estimate

$$a^*(\mu) = \frac{\mu\gamma_L}{\mu\gamma_L + (1-\mu)\gamma_H}. \tag{9}$$

While the principal's optimal estimate corresponded to the Bayesian estimate of the agent's value μ when error cost are symmetric, her optimal estimate is biased when error cost are asymmetric. If too high estimates are relatively more costly ($\gamma_H > \gamma_L$), she is biased towards lower estimates (i.e., $a^*(\mu) < \mu$). If too low estimates are relatively more costly ($\gamma_L > \gamma_H$), she is biased towards higher estimates (i.e., $a^*(\mu) > \mu$). The bias determines the curvature of the optimal estimate.

Lemma 3 (Properties of the optimal estimate) *(a) If $\gamma_H > \gamma_L$, then $a^{*\prime}(\mu) > 0$, $a^{*\prime\prime}(\mu) > 0$ and $a^{*\prime\prime\prime}(\mu) > 0$. (b) If $\gamma_L > \gamma_H$, then $a^{*\prime}(\mu) > 0$, $a^{*\prime\prime}(\mu) < 0$ and $a^{*\prime\prime\prime}(\mu) > 0$.*

The optimal estimate implies the reduced-form utility functions

$$\hat{v}(\mu) = \mathbb{E}_\mu[v(a^*(\mu),\omega)] = -\frac{\gamma_L\gamma_H\mu(1-\mu)}{\mu\gamma_L + (1-\mu)\gamma_H} \text{ and} \tag{10}$$

$$\hat{u}(\mu) = u(a^*(\mu)).$$

The subsequent lemma shows that $\hat{v}(\mu)$ and $\hat{u}(\mu)$ do not always satisfy all the assumptions that I imposed on my reduced model.

[29] In Perez-Richet (2014), the receiver faces a related decision problem. However, his decision is binary and it is the perfectly informed sender who designs the information generation technology.

Lemma 4 (Properties of reduced-form utility functions) *(a) Suppose $\gamma_H > \gamma_L$. Then, $\hat{v}'' > 0$ and $\hat{v}''' > 0$. If u is a CARA or a CRRA utility function that exhibits sufficiently strong risk-aversion, then $\hat{u}' > 0$, $\hat{u}'' < 0$ and $\hat{u}''' > 0$. If u is a CARA utility function that exhibits sufficiently weak risk-aversion, then $\hat{u}' > 0$ and $\hat{u}'' > 0$. (b) Suppose $\gamma_L > \gamma_H$. Then, $\hat{v}'' > 0$ and $\hat{v}''' < 0$. Moreover, $\hat{u}' > 0$, $\hat{u}'' < 0$ and $\hat{u}''' > 0$.*

Surprisingly, even though I allow for the cases where false positives and where false negatives are relatively more costly, a NFP test performs well in both cases. In particular, a test that avoids false positives is generally optimal when false negatives are more costly.

Proposition 4 (Optimal test design in the non-reduced problem) *Under the conditions of Lemma 4 where $\hat{u}'' < 0$, any inducible participation behavior is optimally induced by some NFP test π_ρ^{NFP}. Under the conditions of Lemma 4 where $\hat{u}'' > 0$, perfect-information generation is possible with the perfectly informative NFP test π_1^{NFP}.*

Consider first the case where too high estimates/false positives are more costly (i.e., $\gamma_H > \gamma_L$). The principal's reduced-form utility function satisfies then the assumptions of my reduced model. However, the principal's bias towards low estimates implies that the agent gets a quite unfavorable estimate unless he is able to convince the principal that he is very likely to be good. As a consequence, if the agent is risk-neutral in the estimate a, he benefits from information-generation (i.e., $\hat{u}$ is convex). As long as risk-aversion is sufficiently weak, this effect does still dominate. Perfect information generation is then possible with a perfectly informative test. On the other hand, when risk-aversion is sufficiently strong, the results of my reduced model apply directly. That is, any inducible participation threshold is optimally induced by a NFP test.

Consider next the case where too low estimates/false negatives are more costly (i.e., $\gamma_L > \gamma_H$). Since the principal is biased towards high estimates, the agent gets a quite favorable estimate unless the test reveals that he is very likely to be bad. This does even strengthen the agent's inherent aversion towards risk (i.e., $\hat{u} = u \circ a^*$ is even more concave than u). Yet as I have $\hat{v}''' < 0$ in this case, my sufficient condition for the optimality of a NFP test is violated such that my analysis from the preceding subsection does not apply directly. I demonstrate in the proof of Proposition 4 how I can use the specific structure that I imposed on the principal's utility function to extend the result in Proposition 3. It turns out that the additional perception risk-aversion on the agent's side that arises through the concavity of a^* renders a NFP optimal no matter how strong the agent's initial risk-aversion is.

To get an impression of why the result extends form a somewhat more technical perspective, consider the role of the third derivative assumptions. It is important for the proof of Proposition 3 that

$$\varphi'''_{\lambda,\theta^*}(\mu) = \hat{v}'''(\mu) + \lambda \tilde{u}'''_{\theta^*}(\mu) > 0$$

for the "right" $\lambda > 0$ and for all μ. When $\hat{v}''' \geq 0$, both summands are for all $\lambda > 0$ and for all μ individually positive. For this reason, the assumption that $\hat{v}''' \geq 0$ simplifies the analysis. If this condition is violated, it depends on the relative size of $\hat{v}'''$ and $\hat{u}'''_{\theta^*}$ for the relevant value of λ whether $\varphi'''_{\lambda,\theta^*} > 0$ is true. $\hat{v}'''(\mu) < 0$ stays without consequences for the sign of $\varphi'''_{\lambda,\theta^*}(\mu)$ when $\tilde{u}'''_{\theta^*}(\mu)$ is large. This happens, in particular, to be the case when the agent is very averse towards perception risk (see (A.6) in the proof to Lemma 2).

5. Discussion

5.1. Optimal participation behavior

I know from my analysis in Section 4 that only threshold participation behavior can be induced (Proposition 1), which participation thresholds can be induced (see Lemma A2 in the proof of Proposition 2), and that any participation threshold is optimally induced by a NFP test (Proposition 3). If the principal wants to induce full participation for reasons that are not modeled here explicitly, the unique NFP test that induces the threshold $\theta^* = \underline{\theta}$ is optimal. If not, it remains to determine which participation threshold it is optimal for the principal to induce. My analysis so far reduces the test design problem to the following problem:

$$\begin{aligned}
\max_{\theta^* \in \Theta^*} \quad & (1 - F(\theta^*))\mathbb{E}_\tau[\hat{v}(\mu)] + F(\theta^*)\hat{v}(\mathbb{E}_F[\theta|\theta \leq \theta^*]) \\
\text{s.t.} \quad & \pi^{\text{NFP}}_{\rho(\theta^*)} \text{ and } \chi_{\theta^*} \text{ induce } \tau \\
& \rho(\theta^*) = 1 \text{ if } \theta^* \in \Theta^{**} \\
& \rho(\theta^*) \text{ solves } \mathbb{E}_{\tau_{\theta^*}}[\hat{u}(\mu)] = \hat{u}(\mathbb{E}_F[\theta|\theta \leq \theta^*]) \text{ if } \theta^* \in \Theta^* \setminus \Theta^{**}
\end{aligned}$$

Although the structure of the optimal test is quite simple, the effect of the participation threshold θ^* is complex. It determines simultaneously when the agent participates, with which probabilities the different test results are generated, and which inferences are drawn from the different possible observations. However, as the problem is one-dimensional, it can easily be solved numerically for any primitives of the model. In Section 2 of the supplementary material, I discuss the structure of the set of inducible participation thresholds and the optimal participation threshold for the uniform-quadratic case.

5.2. The principal can mandate participation but testing is costly

A key assumption in my model was that the agent's participation in the test is voluntary. While this assumption is reasonable for some applications, the principal may be able to mandate participation in others. Suppose that the principal can choose between voluntary and mandatory participation but that she faces a cost $\kappa > 0$ of conducting the test; that is, suppose that the principal strives to maximize

$$\mathbb{P}_F[\chi(\theta) = Y] \cdot (\mathbb{E}_\tau[\hat{v}(\mu)] - \kappa) + \mathbb{P}_F[\chi(\theta) = N] \cdot \hat{v}(\mu_N).$$

A completely informative test is obviously optimal under mandatory participation. It allows for perfect learning of the agent's quality. Nevertheless, voluntary participation may be better for the principal as it saves on testing costs. In particular, if κ is sufficiently large, voluntary participation is optimal. A nice feature of my two step solution approach is that the analysis that leads to my main result remains valid for any κ. Hence, for any κ and for any participation threshold that shall be induced, a NFP test is optimal.

5.3. The principal can offer a menu of tests

Suppose now that the principal can offer a menu of tests in that the agent can self-select himself.[30] This allows the principal to further exploit the agent's signaling motive. My aim in this subsection is to discuss what can learnt from my analysis of the problem with a single test and in what respects the menu problem is more complicated.

It will suffice for my purposes to consider the case with two tests. I use the following notation and wording: The menu of tests offered by the principal is $\{\pi^t\}_{t\in T}$ with $T \equiv \{1, 2\}$. The agent has to decide between non-participation ($x = N$) and participation in a test π^t with $t \in T$ ($x = t$). The principal's quality perception after observing the agent's participation in test π^t is μ_Y^t and her quality perception after observing the test result s in test π^t is μ_s^t. A menu of tests $\{\pi^t\}_{t\in T}$ and a participation strategy $\chi : \Theta \to \{N\} \cup T$ induce families of distributions of quality perceptions $\{\tau^t\}_{t\in T}$ and $\{\tau_\theta^t\}_{\theta\in\Theta,t\in T}$. Like in the model with a single test, I can infer τ_θ^t from τ^t. What this means and how all other elements of my model are defined extends in the obvious way.

Participation incentives. Suppose that χ is induced by the menu of tests $\{\pi^t\}_{t\in T}$ and suppose that $\{\tau^t\}_{t\in T}$ is induced by χ and $\{\pi^t\}_{t\in T}$. Like in the model with a single test, the agent's expected payoff from participating in any informative test π^t, $\mathbb{E}_{\tau_\theta^t}[\hat{u}(\mu)]$, is linear and strictly increasing in θ (see Lemma A1 in the appendix). When I denote the slope by l^t, this corresponds to $l^1, l^2 > 0$. Suppose without loss of generality that $l^1 \leq l^2$. Two implications of this are important. First, like before, there exists a threshold $\theta^* \in [\underline{\theta}, \overline{\theta}]$ such that the agent participates in some test if $\theta > \theta^*$ and does not participate if $\theta < \theta^*$. Second, if $l^1 < l^2$, the agent does strictly prefer participating in test π^1 (test π^2) when his private signal is smaller (larger) than some threshold θ^{**}; if $l^1 = l^2$, the agent is indifferent between participating in test π^1 and test π^2.[31]

Optimal inducement of a given threshold participation behavior. Suppose the thresholds θ^* and θ^{**} with $\underline{\theta} \leq \theta^* < \theta^{**} < \overline{\theta}$ are inducible by some menu of tests. This represents the most interesting case where both tests are used with positive probability for an interval of signals. My analysis in Subsection 4.2 allows for some conclusions about the structure of the test that optimally induces θ^* and θ^{**}. The two tests give rise to two participation constraints. The first constraint is as in the original problem:

$$\mathbb{E}_{\tau_{\theta^*}^1}[\hat{u}(\mu)] = \overline{u}^1 \text{ with } \overline{u}^1 \equiv \hat{u}(\mathbb{E}_F[\theta|\theta \leq \theta^*]).$$

[30] Bar et al. (2012) study the enrollment of students who differ in taste and quality in courses of different difficulty. They investigate the effect of providing information about the course difficulty to employers but not the design of course difficulties that maximizes information generation. Kolotilin et al. (2016) allow persuasion mechanisms to condition on announcements of a privately informed receiver in a Bayesian persuasion framework that has a different structure than my transformed problem.

[31] This case is analogous to the case where the principal offers a single completely uninformative test. A non-threshold participation behavior may then arise.

The second constraint corresponds to

$$\mathbb{E}_{\tau^2_{\theta^{**}}}[\hat{u}(\mu)] = \mathbb{E}_{\tau^1_{\theta^{**}}}[\hat{u}(\mu)].^{32}$$

If $\mathbb{E}_{\tau^1_{\theta^{**}}}[\hat{u}(\mu)]$ was constant, say $\overline{u}^2$, the problems of designing the first test π^1 and the second test π^2 were independent. It would immediately follow from my analysis in Subsection 4.2 that a menu of two NFP tests is optimal. However, the "outside option" of the second test design problem $\overline{u}^2$ depends typically on the design of the first test. If the outside option of the second problem is minimized by the NFP test that solves the first problem, a menu of two NFP tests would still be optimal. I give sufficient conditions for this in Corollary 1 (b). However, as I motivated at the end of Subsection 4.2, this will only be the case if the agent's aversion to perception risk is sufficiently stronger than his signaling motive and his downside risk aversion. If this is not the case, it may be optimal to choose a test that is subject to false positives as the first test to relax the outside option that is relevant for the second test design problem. It is always optimal to choose a NFP test as the second test.

5.4. The agent is perfectly informed and/or risk-neutral in conjunction with costly participation

The analysis of my main result in Subsection 4.2 does not rely on the assumption that the agent's private information is imperfect. It depends only on that the threshold agent has a below average signal. Moreover, it extends under mild additional assumptions on $\hat{v}$ to the case where the agent is perception risk-neutral.[33] However, the two assumptions rendered the analysis in Subsection 4.2 relevant. They were responsible for the emergence of an endogenous participation constraint that prevented full unraveling under a perfectly informative test. If a participation constraint arises for other reasons, my main analysis applies also to problems where the agent is perfectly informed and/or cares only about expected perception. For instance, such a constraint could arise like in Harbaugh and Rasmusen (2016) where the agent suffers an exogenously given cost from participating.

5.5. The agent's payoff depends directly on quality

Suppose next that the agent's non-reduced payoff is $u(a, \omega)$ instead of $u(a)$. Then the agent's reduced-form payoff is given by $\hat{u}_\theta(\mu) = \mathbb{E}_\mu[u(a^*(\mu), \omega)|\theta]$. In contrast to the original model, it may depend not only on the principal's quality perception but also on the agent's private signal θ as he can use this information to update his belief about ω. I can impose for each $\theta \in \Theta$ assumptions on reduced-form payoffs that are analogous to the assumption in my original model: $\hat{u}_\theta : [0,1] \to \mathbb{R}$ is a smooth function with $\hat{u}'_\theta > 0$, $\hat{u}''_\theta < 0$ and $\hat{u}'''_\theta \geq 0$. An interesting

[32] Notice that this is the only constraint that is necessary to induce that the agent strictly prefers π^1 over π^2 if $\theta < \theta^{**}$ and π^2 over π^1 if $\theta > \theta^{**}$. $l^1 < l^2$ follows automatically from the indifference condition, $\mu^1_Y < \theta^{**} < \mu^2_Y$, and strict monotonicity of $\mathbb{E}_{\tau^t_\theta}[\hat{u}(\mu)]$ in θ.

[33] This is because the auxiliary function $\tilde{u}_{\theta^*}$ that is key for the analysis is still strictly concave when $\hat{u}'' = 0$. However, since $\tilde{u}'''_{\theta^*} = 0$ when $\hat{u}'' = 0$, my reasoning in the proof of Proposition 3 would require assuming $\hat{v}''' > 0$ instead of $\hat{v}''' \geq 0$.

special case of the generalized framework reflects complementarities between the agent's actual quality and how good his quality is perceived by the principal. $\hat{u}_\theta$ is then increasing in θ.

When $\hat{u}_\theta$ depends on θ, the proof of Proposition 1 may be affected. I.e., the induced participation behavior may not be a threshold participation behavior anymore. However, if the specific structure of $\hat{u}_\theta$ is such the agent has still stronger incentives to participate in any informative test the higher his private signal, then the entire remaining analysis in Section 4 stays valid. This is a direct consequence of the fact that the remaining analysis depends only on the properties of $\hat{u}_\theta$ for a fixed θ and that these properties are assumed to be as in the original model. In particular, a NFP test remains optimal.

6. Literature

The classic literature on education observed already that signaling and testing are two important channels of learning. Spence (1973) focuses on the signaling role of education. Arrow (1973) and Stiglitz (1975) point to the role of schooling as an information generating device and a costly signal at the same time. Weiss (1983) and Alós-Ferrer and Prat (2012) take an information generation technology as given and study the incentives to engage in costly signaling prior to information generation. My article is concerned with the design of an information generation technology that optimally exploits signaling incentives.

My article contributes mainly to the literature on information design in sender-receiver games. I study a problem where the information structure is controlled by the receiver and a constraint derives from voluntary participation of the privately informed sender.

The literature on Bayesian persuasion studies the design of the information structure by an uninformed sender. In a seminal contribution, Kamenica and Gentzkow (2011) study this problem for a general class of environments.[34] Bayes' Law restricts the expected value of the posterior belief but the design of the information structure is not subject to any further constraints. As a concavification approach is feasible for this problem, it has a simple solution that can be extended into various directions.[35] My analysis requires techniques that can cope with an additional constraint that derives from the designer's inability to unilaterally impose an information structure and that depends on a different distribution of beliefs than the designer's objective function.

Some articles investigate persuasion problems where the sender is already informed at the design stage (e.g., Perez-Richet, 2014; Gill and Sgroi, 2012; Li and Li, 2013). The information

[34]See also Rayo and Segal (2010) and Kolotilin (2015). See Kolotilin (2016) and Kolotilin et al. (2016) for problems where the receiver possesses private information about which decision is optimal for the sender. Persuasion problems where the sender engages in specific kinds of market interaction after the disclosure of information are studied by Ostrovsky and Schwarz (2010) (matching markets) and Goldstein and Leitner (2015) (banking).

[35]This includes the disclosure of information by many senders (Gentzkow and Kamenica, 2017), the interpretation of information by many receivers (Wang, 2013; Alonso and Câmara, 2016b; Taneva, 2016), heterogeneity in priors (Alonso and Câmara, 2016a), and costly information structures (Gentzkow and Kamenica, 2014).

structure is then jointly determined by the sender's information generation technology choice and the signaling effect that is associated with it.[36] A related effect arises also in my article. The information structure is jointly determined by the information generation technology and the participation behavior that it implies.

In problems where the information structure is controlled by the receiver, participation of a privately informed sender is often an issue. Most closely related to my article is independent work by Harbaugh and Rasmusen (2016). Like me, they are interested in information generation when a constraint derives from voluntary participation. In contrast to me, they model the sender differently such that different incentive problems arise: he is perfectly informed about his continuous quality, risk-neutral with respect to how his quality is perceived, and he has to bear an exogenously given fixed cost when he participates. Manipulating how a participating sender's quality is perceived on average is the only instrument for setting participation incentives. In my article, risk-aversion makes participation endogenously costly and there are different ways to induce participation. Perception risk considerations, which are mute in their article, are the central theme in mine. In a less related article, Perez-Richet and Prady (2012) study similar channels of learning as I do. The sender, who is perfectly informed about his binary quality, chooses the cost of test accuracy ("complexity") and the receiver chooses the test accuracy ("the level of understanding"). In contrast to my article, signaling occurs before the test design and tests are restricted to a specific one-dimensional class.

Schweizer and Szech (2014) and Caplin and Eliaz (2003) study test design problems in medical contexts with anticipatory costs of learning. In Schweizer and Szech (2014), a doctor designs a test to maximize the expected utility of his patient. Since participation is no issue, the concavification approach does apply directly. For a reasonable class of preferences, an inaccurate binary test that is not subject to false positives is optimal. Caplin and Eliaz (2003) study the design of a pass-fail test that is capable of stopping the spread of HIV when participation is voluntary and agents condition their sexual matching behavior on the generated information. Unfavorable test results can be hidden and agents possess no private information. Setting participation incentives may require the test to be inaccurate. Stopping the spread of the disease requires that the test is never passed by an agent who is actually infected. The same test structure that is optimal in my article is for different reasons optimal in these articles.

Benoît and Dubra (2004) discuss the preferences of a sender with the median signal about two different imperfectly informative signals. They argue that he may prefer the signal that generates less information even though he would prefer his information to be perfectly revealed. There is no signaling effect and no test design in this paper, but it shares with me the effect

[36] For a (potentially) unconstrained and a constrained interim design problem where the sender is perfectly informed about his binary type, Perez-Richet (2014) and Gill and Sgroi (2012) find that only pooling equilibria can arise. By contrast, signaling can occur in Li and Li (2013) where the sender faces a constrained interim design problem and has imperfect binary information. Other constrained interim design problems are studied in Titman and Trueman (1986), Gill and Sgroi (2008) and in Section 4 of Daley and Green (2014).

that a party that likes information generation prefers a tests that is less accurate than possible.

Some strands of literature are related to my article but differ more strongly. One branch of the test design literature studies the incentive effects of tests on effort provision (e.g., Dubey and Geanakoplos, 2010). The literature on certification studies the design of the information structure by a profit maximizing intermediary. This design goal implies very different effects. For instance, in the seminal contribution by Lizzeri (1999), the certifier can capture the entire informational surplus in the market despite revealing no information. That endogenous participation affects the value of non-participation is also important in Tirole (2012) and Philippon and Skreta (2012), who analyze mechanism design problems in the context of government interventions in financial markets. Milgrom and Weber (1982) and Ottaviani and Prat (2001) identify environments in that some player has an interest in the public revelation of information. I take such an interest as given and study the problem where the information structure cannot be unilaterally imposed by such a player.

7. Conclusion

I have studied optimal test design under voluntary participation of an imperfectly informed, perception risk-averse agent when the principal benefits from information. Only threshold participation behavior where the agent participates when his information is sufficiently favorable can be induced. For a large class of reduced-form utility functions and for any participation threshold that shall be induced, my main result shows that a binary test that is not subject to false positives is optimal. The probability of false negatives serves as an instrument to foster participation. Learning about the agent's quality is generally imperfect either due to less than full participation, inaccuracy of the optimal test, or both. My main result can also be interpreted as a foundation for the use of simple testing procedures. Furthermore, I have shown that a binary test that is not subject to false positives is optimal for a class of non-reduced problems where the principal has to estimate the agent's quality and suffers from asymmetric error cost. Interestingly, even when the principal suffers more from false negatives, a test that avoids false positives is optimal.

A rough intuition for my main results is the following: Learning about the quality of an imperfectly informed agent comes along with imposing a perception risk on him. However, the risk faced by the agent differs depending on his private information. For example, a test that is not subject to false positives imposes little risk on the agent when it is relatively likely that he is bad; a test that is not subject to false negatives imposes little risk on the agent when it is relatively likely that he is good. By choosing a test structure that imposes less risk on the agent when he is bad, the participation constraint of the agent with the threshold signal, who is relatively likely to be bad, is relaxed. This allows it to increase the accuracy of the test strongly enough such that the agent with the threshold signal is perceived on average as worse. As this is just what improved learning corresponds to, a binary test that is not subject to false positives is optimal because it imposes the smallest possible risk on a bad agent.

I find two extensions of my model particularly interesting. In the first extension, quality is non-binary. It may then under certain conditions still be possible to reduce the test design problem into a concavification problem, but as this problem is then multi-dimensional, deriving the structure of the optimal test becomes more involved. Moreover, private information may then have a more subtle effect on participation incentives.[37] In the second extension, the principal can use monetary transfers between participants and non-participants as an additional instrument to set participation incentives and—if I allow also for the extension that I discussed in Subsection 5.3—also to set incentives for the agent to select into different tests. Studying the generalized problem would allow it to assess the relative importance of different instruments that can be used to foster participation.[38]

Acknowledgements

Previous versions of this article circulated under the titles "Test design under voluntary participation and conflicting preferences", "Optimal test design under imperfect private information and voluntary participation", and "Imperfect private information and the design of information-generating mechanisms." It elaborates ideas partly originating from joint work with Elisabeth Schulte. Therefore, I am deeply indebted to Elisabeth. I would like to thank Kfir Eliaz, Hans Peter Grüner, Johannes Hörner, Emir Kamenica, Johannes Koenen, Benny Moldovanu, Tymofiy Mylovanov, Volker Nocke, Marco Ottaviani, Martin Peitz, Sergei Severinov, Konrad Stahl, Stephan Lauermann, Nora Szech, Tymon Tatur, and various seminar participants. Moreover, I would like to thank the associate editor and two anonymous referees for very helpful suggestions. Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

Appendix A. Proofs

Proof of Proposition 1

I first prove an auxiliary result.

Lemma A1 *If τ is induced by an informative test π and a participation strategy χ, then $\mathbb{E}_{\tau_\theta}[\hat{u}(\mu)]$ is continuous and strictly increasing in θ.*

Proof. Fix any informative test π and any participation strategy χ. Let τ be induced by π and χ. Suppose without loss of generality that $S = \{1, \ldots, n\}$. By (3),

$$\mathbb{E}_{\tau_\theta}[\hat{u}(\mu)] \;=\; \theta \sum_{s\in S} (\pi(s|g) - \pi(s|b))\hat{u}(\mu_s) + \sum_{s\in S} \pi(s|b)\hat{u}(\mu_s).$$

[37] See Section 3 of the supplementary material for a discussion of the relation of my model to a problem where quality is non-binary but principal and agent care only about the expected quality as perceived by the principal.

[38] Such an extension would make the design problem more closely related to non-standard mechanism design problems that blend information design with mechanism design (see, e.g., Calzolari and Pavan, 2006a,b; Bergemann and Pesendorfer, 2007; Eső and Szentes, 2007; Pancs, 2014).

Since this expression is linear in θ, I need only to show that the coefficient of θ is strictly positive. I can rewrite this coefficient in the following way:

$$\begin{aligned}
\sum_s (\pi(s|g) - \pi(s|b))\hat{u}(\mu_s) &= \sum_{s'} \Big(\pi(s'|g) \sum_{s''} \pi(s''|b) - \pi(s'|b) \sum_{s''} \pi(s''|g))\hat{u}(\mu_{s'}\Big) \\
&= \sum_{s'} \sum_{s''>s'} (\pi(s'|g)\pi(s''|b) - \pi(s'|b)\pi(s''|g))\hat{u}(\mu_{s'}) \\
&\quad + \sum_{s'} \sum_{s''<s'} (\pi(s'|g)\pi(s''|b) - \pi(s'|b)\pi(s''|g))\hat{u}(\mu_{s'}) \\
&= \sum_{s'} \sum_{s''>s'} (\pi(s'|g)\pi(s''|b) - \pi(s'|b)\pi(s''|g))(\hat{u}(\mu_{s'}) - \hat{u}(\mu_{s''}))
\end{aligned}$$

It follows from (1) and $\hat{u}' > 0$ that the sign of $\hat{u}(\mu_{s'}) - \hat{u}(\mu_{s''})$ corresponds to the sign of $\pi(s'|g)\pi(s''|b) - \pi(s'|b)\pi(s''|g)$. This implies that each summand of the coefficient is weakly positive. Since π is informative, $\mu_{s'} \neq \mu_{s''}$ must be true for some s', s''. This implies that there exists at least one summand that is strictly positive. This proves the result. q.e.d.

(a) I define first a specific class of threshold participation strategies: For any $\theta^* \in [\underline{\theta}, \overline{\theta})$ define $\hat{\chi}_{\theta^*}(\theta) = Y$ if $\theta \geq \theta^*$ and $\hat{\chi}_{\theta^*}(\theta) = N$ if $\theta < \theta^*$; for $\theta^* = \overline{\theta}$ define $\hat{\chi}_{\theta^*}(\theta) = N$. Let $\{\tau_{\theta,\theta^*}\}_{\theta\in\Theta}$ denote the family of distributions that is induced by the given informative test π and the participation strategy $\hat{\chi}_{\theta^*}$ with $\mu_Y = \mathbb{E}_F[\theta|\theta \geq \theta^*]$ from the perspective of the agent. To prove the result, it suffices to show that there exists a θ^* such that $\hat{\chi}_{\theta^*}$ is induced by π with $\mu_N = \mathbb{E}_F[\theta|\theta \leq \theta^*]$ and $\mu_Y = \mathbb{E}_F[\theta|\theta \geq \theta^*]$. I distinguish three cases:

Case 1: $\mathbb{E}_{\tau_{\underline{\theta},\underline{\theta}}}[\hat{u}(\mu)] \geq \hat{u}(\mathbb{E}_F[\theta|\theta \leq \underline{\theta}])$. By the supposition, the agent has at least a weak incentive to participate when his private signal is $\theta = \underline{\theta}$ and the supposed participation strategy is $\hat{\chi}_{\underline{\theta}}$. This implies by Lemma A1 that he has a strict incentive to participate if $\theta > \underline{\theta}$. Hence, $\hat{\chi}_{\underline{\theta}}$ is induced by π.

Case 2: $\mathbb{E}_{\tau_{\overline{\theta},\overline{\theta}}}[\hat{u}(\mu)] \leq \hat{u}(\mathbb{E}_F[\theta|\theta \leq \overline{\theta}])$. By the supposition, the agent has at least a weak incentive not to participate when his private signal is $\theta = \overline{\theta}$ and the supposed participation strategy is $\hat{\chi}_{\overline{\theta}}$. This implies by Lemma A1 that he has a strict incentive not to participate if $\theta < \overline{\theta}$. Hence, $\hat{\chi}_{\overline{\theta}}$ is induced by π.

Case 3: $\mathbb{E}_{\tau_{\underline{\theta},\underline{\theta}}}[\hat{u}(\mu)] < \hat{u}(\mathbb{E}_F[\theta|\theta \leq \underline{\theta}])$ *and* $\mathbb{E}_{\tau_{\overline{\theta},\overline{\theta}}}[\hat{u}(\mu)] > \hat{u}(\mathbb{E}_F[\theta|\theta \leq \overline{\theta}])$. Continuity of F and of $\hat{u}$ implies that $\mathbb{E}_{\tau_{\theta^*,\theta^*}}[\hat{u}(\mu)]$ and $\hat{u}(\mathbb{E}_F[\theta|\theta \leq \theta^*])$ are both continuous in θ^*. This and the supposition allow me to apply an Intermediate Value Theorem. I obtain that there exists a $\theta^* \in (\underline{\theta}, \overline{\theta})$ such that $\mathbb{E}_{\tau_{\theta^*,\theta^*}}[\hat{u}(\mu)] = \hat{u}(\mathbb{E}_F[\theta|\theta \leq \theta^*])$. That is, the agent is indifferent between participating and not participating when his private signal is θ^* and the supposed participation strategy is $\hat{\chi}_{\theta^*}$. It follows from Lemma A1 that the agent has a strict incentive to participate for any signal $\theta > \theta^*$ and a strict incentive not to participate for any signal $\theta < \theta^*$. Hence, $\hat{\chi}_{\theta^*}$ is induced by π.

(b) Suppose τ is induced by π and χ. $\mathbb{E}_{\tau_\theta}[\hat{u}(\mu)]$ is by Lemma A1 strictly increasing in θ. Since $\hat{u}(\mu_N)$ does not depend on θ, the agent has a strictly larger incentive to participate the higher his private signal. Hence, if π induces χ, χ must be a threshold strategy where the agent participates for an upper interval of private signals. q.e.d.

Proof of Proposition 2

Before I prove the properties in the proposition, I characterize the sets Θ^* and Θ^{**} in terms of the primitives of the model.

Lemma A2 *(a)* $\Theta^* = \{\theta^* \in [\underline{\theta}, \overline{\theta}) | (1-\theta^*) \cdot \hat{u}(0) + \theta^* \cdot \hat{u}(1) \leq \hat{u}(\mathbb{E}_F[\theta|\theta \leq \theta^*])\}$. *(b)* $\Theta^{**} = \{\theta^* \in (\underline{\theta}, \overline{\theta}) | (1-\theta^*) \cdot \hat{u}(0) + \theta^* \cdot \hat{u}(1) = \hat{u}(\mathbb{E}_F[\theta|\theta \leq \theta^*])\}$.

Proof. (a) *Step 1: If* $(1-\theta^*)\cdot\hat{u}(0)+\theta^*\cdot\hat{u}(1) > \hat{u}(\mathbb{E}_F[\theta|\theta \leq \theta^*])$ *with* $\theta^* \in [\underline{\theta}, \overline{\theta})$*, then there exists no informative test that induces the threshold* θ^**.* Assume to the contrary that there exists an informative test π that induces a threshold participation strategy χ with the threshold θ^* even though

$$(1-\theta^*) \cdot \hat{u}(0) + \theta^* \cdot \hat{u}(1) > \hat{u}(\mathbb{E}_F[\theta|\theta \leq \theta^*]) \tag{A.1}$$

and $\theta^* \in [\underline{\theta}, \overline{\theta})$. Let τ be induced by π and χ. Notice that by $\hat{u}'' < 0$ and Jensen's inequality, (A.1) can only hold if $\theta^* > \underline{\theta}$. By consistency of μ_N and μ_Y with χ,

$$\mu_N = \mathbb{E}_F[\theta|\theta \leq \theta^*] \text{ and } \mu_Y = \mathbb{E}_F[\theta|\theta \geq \theta^*]. \tag{A.2}$$

The threshold agent's expected payoff from participation is $\mathbb{E}_{\tau_{\theta^*}}[\hat{u}(\mu)]$. By $\hat{u}'' < 0$ and Jensen's inequality, I obtain

$$\begin{aligned} \mathbb{E}_{\tau_{\theta^*}}[\hat{u}(\mu)] &\geq \mathbb{E}_{\tau_{\theta^*}}[(1-\mu)\hat{u}(0) + \mu\hat{u}(1)] \\ &= (1-\mathbb{E}_{\tau_{\theta^*}}[\mu])\hat{u}(0) + \mathbb{E}_{\tau_{\theta^*}}[\mu]\hat{u}(1). \end{aligned} \tag{A.3}$$

Furthermore, I have

$$\begin{aligned} \mathbb{E}_{\tau_{\theta^*}}[\mu] - \theta^* &= \sum_s (\pi(s|g)\theta^* + \pi(s|b)(1-\theta^*))\mu_s - \sum_s \pi(s|g)\theta^* \\ &= \sum_s (\pi(s|g)\theta^* + \pi(s|b)(1-\theta^*)) \frac{\pi(s|g)\mu_Y}{\pi(s|g)\mu_Y + \pi(s|b)(1-\mu_Y)} \\ &\quad - \sum_s \frac{\pi(s|g)\mu_Y + \pi(s|b)(1-\mu_Y)}{\pi(s|g)\mu_Y + \pi(s|b)(1-\mu_Y)} \pi(s|g)\theta^* \\ &= \sum_s \frac{\pi(s|g)\pi(s|b)(\mu_Y - \theta^*)}{\pi(s|g)\mu_Y + \pi(s|b)(1-\mu_Y)} \geq 0. \end{aligned} \tag{A.4}$$

The first equality follows from (3) and from writing θ^* in a more complicated way. The second equality follows from using (1). The third equality follows from simplifying. The inequality follows from (A.2). By $\hat{u}' > 0$ and (A.4), the right-hand side of (A.3) exceeds the left-hand side of (A.1) at least weakly. I obtain from this and from (A.2) that

$$\mathbb{E}_{\tau_{\theta^*}}[\hat{u}(\mu)] > \hat{u}(\mu_N).$$

Since $\mathbb{E}_{\tau_\theta}[\hat{u}(\mu)]$ is linear in θ and thus continuous, it follows from the strict inequality that the agent has a strict incentive to participate if his private signal is slightly smaller than θ^*. Contradiction.

Step 2: If $(1-\theta^*) \cdot \hat{u}(0) + \theta^* \cdot \hat{u}(1) \leq \hat{u}(\mathbb{E}_F[\theta|\theta \leq \theta^*])$ *with* $\theta^* \in [\underline{\theta}, \overline{\theta})$*, then there exists an informative test that induces* θ^**.* Consider any $\theta^* \in [\underline{\theta}, \overline{\theta})$ for that the inequality holds. Let $\{\tau_{\theta,\rho}\}_{\theta\in\Theta}$ denote the family of distributions that is induced by π_ρ^{NFP} and χ_{θ^*} from the agent's perspective. By consistency, $\mu_Y = \mathbb{E}_F[\theta|\theta \geq \theta^*]$. By Lemma A1, sufficient for the inducement of χ_{θ^*} by π_ρ^{NFP} is that there exists a test accuracy $\rho \in (0,1]$ and a $\mu_N \in \Theta$ that is consistent with χ_{θ^*} such that the participation constraint of the agent with the threshold signal $\theta^* \in [\underline{\theta}, \overline{\theta})$

is binding. Consider $\mu_N = \mathbb{E}_F[\theta|\theta \leq \theta^*]$. When the agent with the threshold signal participates, his expected payoff is

$$\mathbb{E}_{\tau_{\theta^*,\rho}}[\hat{u}(\mu)] \quad = \quad (1-\rho\theta^*) \cdot \hat{u}(\frac{\mu_Y - \rho\mu_Y}{1-\rho\mu_Y}) + \rho\theta^* \cdot \hat{u}(1)$$

by (1) and (3). When the test is perfectly informative (i.e., $\rho = 1$), this becomes

$$\mathbb{E}_{\tau_{\theta^*,1}}[\hat{u}(\mu)] \quad = \quad (1-\theta^*) \cdot \hat{u}(0) + \theta^* \cdot \hat{u}(1).$$

Since this is weakly smaller than $\hat{u}(\mu_N)$ by the supposition, the agent has at least a weak incentive not to participate in a perfectly informative test when he has the threshold signal. On the other hand, when the test is completely uninformative (i.e., $\rho = 0$), I get

$$\mathbb{E}_{\tau_{\theta^*,0}}[\hat{u}(\mu)] \quad = \quad \hat{u}(\mu_Y).$$

Since $\hat{u}(\mu_Y) > \hat{u}(\mu_N)$, the agent has a strict incentive to participate in a completely noninformative test when he has the threshold signal. These two observations together with the fact that $\mathbb{E}_{\tau_{\theta^*,\rho}}[\hat{u}(\mu)]$ is continuous in ρ allow me to apply an Intermediate Value Theorem. I obtain that there exists some $\rho^* \in (0,1]$ such that the participation constraint of the agent is binding when he has the threshold signal.

(b) *Step 1: There exists no perfectly informative test that induces $\theta^* = \underline{\theta}$.* Assume to the contrary that $\underline{\theta} \in \Theta^{**}$. There exists then a perfectly informative test π and a threshold participation strategy χ with threshold $\underline{\theta}$ such that π induces χ. This means that there exists a $\mu_N \in \Theta$ that is consistent with χ and a distribution τ that is induced by π and χ such that

$$\mathbb{E}_{\tau_{\underline{\theta}}}[\hat{u}(\mu)] \geq \hat{u}(\mu_N). \tag{A.5}$$

Since the test is perfectly informative, $\mathbb{E}_{\tau_{\underline{\theta}}}[\hat{u}(\mu)] = \underline{\theta}\hat{u}(1) + (1-\underline{\theta})\hat{u}(0)$. By $\hat{u}'' < 0$, $\underline{\theta} \in (0,1)$ and Jensen's inequality, $\mathbb{E}_{\tau_{\underline{\theta}}}[\hat{u}(\mu)] < \hat{u}(\underline{\theta})$. Since $\hat{u}(\underline{\theta}) \leq \hat{u}(\mu_N)$ for all $\mu_N \in \Theta$, I obtain a contradiction to (A.5).

Step 2: If $(1-\theta^*) \cdot \hat{u}(0) + \theta^* \cdot \hat{u}(1) \neq \hat{u}(\mathbb{E}_F[\theta|\theta \leq \theta^*])$ with $\theta^* \in (\underline{\theta}, \overline{\theta})$, then there exists no perfectly informative test that induces the threshold θ^*. It follows from Part (a) that there exists no perfectly informative test that induces $\theta^* \in (\underline{\theta}, \overline{\theta})$ if $(1-\theta^*) \cdot \hat{u}(0) + \theta^* \cdot \hat{u}(1) > \hat{u}(\mathbb{E}_F[\theta|\theta \leq \theta^*])$. It remains thus only to show that this is also not possible for the reverse inequality.

Assume to the contrary that there exists a perfectly informative test π that induces a threshold participation strategy χ with threshold $\theta^* \in (\underline{\theta}, \overline{\theta})$ even though

$$(1-\theta^*) \cdot \hat{u}(0) + \theta^* \cdot \hat{u}(1) < \hat{u}(\mathbb{E}_F[\theta|\theta \leq \theta^*]).$$

By consistency of μ_N with χ, $\mu_N = \mathbb{E}_F[\theta|\theta \leq \theta^*]$. Suppose τ is induced by π and χ. By Lemma A1 and $\theta^* \in (\underline{\theta}, \overline{\theta})$, necessary for the inducement of χ by π is that the agent's participation constraint is binding when he has the threshold signal. When the agent has the threshold signal, his expected payoff from participation is

$$\mathbb{E}_{\tau_{\theta^*}}[\hat{u}(\mu)] = (1-\theta^*)\hat{u}(0) + \theta^*\hat{u}(1).$$

Since the preceding observations and the supposition imply that the agent has a strict incentive not to participate when he has the threshold signal, I obtain a contradiction.

Step 3: If $(1-\theta^*) \cdot \hat{u}(0) + \theta^* \cdot \hat{u}(1) = \hat{u}(\mathbb{E}_F[\theta|\theta \leq \theta^*])$ for $\theta^* \in (\underline{\theta}, \overline{\theta})$, then there exists a *perfectly informative test that induces θ^*.* This is a direct consequence of the proof of Step 2 in Part (a). q.e.d.

I come now to the proof of the three statements in the proposition.

(b) This is Step 1 in the proof to Lemma A2 (b).

(c) By Lemma A2 (a), $\theta^* = \underline{\theta}$ is inducible if $(1-\underline{\theta}) \cdot \hat{u}(0) + \underline{\theta} \cdot \hat{u}(1) \leq \hat{u}(\underline{\theta})$. It follows from $\hat{u}'' < 0$ and Jensen's inequality that this is the case. Since it follows from Part (b) that $\underline{\theta}$ is not inducible by a perfectly informative test, I obtain the result.

(a) First observe that non-participation is always inducible. If the principal believes that the agent participates only if he has the worst possible private signal, $\chi(\theta) = N$ is induced by any test π_ρ^{NFP} with ρ sufficiently close to zero. Thus, the set of all thresholds $\theta^* \in \Theta$ that is inducible by some informative test is $\Theta^* \cup \{\overline{\theta}\}$. This set is obviously non-empty. Compactness follows directly from the characterization of Θ^* in Lemma A2 (a) and the fact that $\overline{\theta}$ is generally inducible. (Notice that Θ^* may only be non-closed because $\overline{\theta}$ is missing.) q.e.d.

Proof of Lemma 1

Fix any $\theta^* \in \Theta^* \backslash \Theta^{**}$. I distinguish two cases.

Case 1: $\theta^* \in (\underline{\theta}, \overline{\theta})$. It follows from Lemma A1 that bindingness of the participation constraint of the agent with the threshold signal is necessary and sufficient for the inducement of the participation threshold θ^*. Since the participation constraint must be binding, it is without loss of generality to restrict attention to tests that induce the specific threshold participation strategy χ_{θ^*}. Since $\theta^* > \underline{\theta}$, consistency implies $\mu_N = \mathbb{E}_F[\theta|\theta \leq \theta^*]$. Since for a given θ^* the probability terms and the payoff from non-participation in the principal's objective function (5) are constants, maximizing (5) is equivalent to maximizing $\mathbb{E}_\tau[\hat{v}(\mu)]$. This yields the optimization problem stated in the lemma.

Case 2: $\theta^* = \underline{\theta}$. Two things change in this case. First, necessary and sufficient for the inducement of threshold $\theta^* = \underline{\theta}$ is that the agent has a weak incentive to participate when he has the threshold signal. Second, μ_N is not pinned down by Bayes Law. However, since non-participation occurs with probability zero, it does not matter for (5) which μ_N is selected. Moreover, since a test induces the threshold θ^* if the threshold agent has a weak incentive to participate for some $\mu_N \in \Theta$, necessary and sufficient for the inducement of the participation threshold is that he has a weak incentive to participate for $\mu_N = \underline{\theta} = \mathbb{E}_F[\theta|\theta \leq \underline{\theta}]$. Since participation must be at least weakly optimal for the threshold agent, it is also in this case without loss of generality to restrict attention to tests that induce the specific threshold participation strategy χ_{θ^*}.

It remains for me to argue why it is without loss of generality to consider the case where the threshold agent's participation constraint is binding. Assume to the contrary that θ^* is optimally induced by a test π' for that the participation constraint holds strict. Consider the class of test $\{\pi''_\rho\}_{\rho \in [0,1]}$ that is constructed in the following way: with probability ρ the test π''_ρ does perfectly reveal the agent's quality and with probability $1-\rho$ the original test π' is executed. For all $\rho \in (0,1]$, any distribution τ''_ρ that is induced by π''_ρ and χ_{θ^*} is a mean-preserving spread of any distribution τ' induced by π' and χ_{θ^*}. Since $\hat{v}'' > 0$, $\mathbb{E}_{\tau''_\rho}[\hat{v}(\mu)] > \mathbb{E}_{\tau'}[\hat{v}(\mu)]$. Thus, to obtain a contradiction, it suffices to find a $\rho \in (0,1]$ such that the threshold agent's participation constraint is binding. By the supposition, the threshold agent has a strict incentive to participate if $\rho = 0$. By $\theta^* \in \Theta^* \backslash \Theta^{**}$ and Lemma A2, the threshold agent has a strict incentive not to participate if $\rho = 1$. Since the threshold agent's expected payoff from participating in the test π''_ρ when the supposed participation strategy is χ_{θ^*} is continuous in ρ, it follows from an Intermediate Value Theorem that there exists some $\rho^* \in (0, 1)$ such that his participation constraint is binding. Contradiction. q.e.d.

Proof of Lemma 2

(a) Notice that $\theta^* \in \Theta^* \backslash \Theta^{**}$ implies that only $\mu_Y = \mathbb{E}_F[\theta | \theta \geq \theta^*]$ is consistent with χ_{θ^*}. Thus, to prove the result, it remains only to show that (1), (2) and (3) imply (4). I have

$$
\begin{aligned}
& \left(\frac{\theta}{\mu_Y}\mu + \frac{1-\theta}{1-\mu_Y}(1-\mu)\right)\tau(\mu) \\
\stackrel{(2)}{=} \; & \sum_{s:\mu_s=\mu}\left(\frac{\theta}{\mu_Y}\mu + \frac{1-\theta}{1-\mu_Y}(1-\mu)\right)(\pi(s|g)\mu_Y + \pi(s|b)(1-\mu_Y)) \\
\stackrel{(1)}{=} \; & \sum_{s:\mu_s=\mu}\left(\frac{\theta}{\mu_Y}\pi(s|g)\mu_Y + \frac{1-\theta}{1-\mu_Y}\pi(s|b)(1-\mu_Y)\right) \\
= \; & \sum_{s:\mu_s=\mu}(\pi(s|g)\theta + \pi(s|b)(1-\theta)) \\
\stackrel{(3)}{=} \; & \tau_\theta.
\end{aligned}
$$

(b) Smoothness of $\tilde{u}_{\theta^*}$ follows from smoothness of $\hat{u}$ and linearity of the additional factor. It remains to investigate the signs of the derivatives of $\tilde{u}_{\theta^*}$. I get the following:

$$
\begin{aligned}
\tilde{u}'_{\theta^*}(\mu) &= \left(\frac{\theta^*}{\mu_Y} - \frac{1-\theta^*}{1-\mu_Y}\right)\hat{u}(\mu) + \left(\frac{\theta^*}{\mu_Y}\mu + \frac{1-\theta^*}{1-\mu_Y}(1-\mu)\right)\hat{u}'(\mu) \\
\tilde{u}''_{\theta^*}(\mu) &= 2\left(\frac{\theta^*}{\mu_Y} - \frac{1-\theta^*}{1-\mu_Y}\right)\hat{u}'(\mu) + \left(\frac{\theta^*}{\mu_Y}\mu + \frac{1-\theta^*}{1-\mu_Y}(1-\mu)\right)\hat{u}''(\mu) \\
\tilde{u}'''_{\theta^*}(\mu) &= 3\left(\frac{\theta^*}{\mu_Y} - \frac{1-\theta^*}{1-\mu_Y}\right)\hat{u}''(\mu) + \left(\frac{\theta^*}{\mu_Y}\mu + \frac{1-\theta^*}{1-\mu_Y}(1-\mu)\right)\hat{u}'''(\mu) \qquad \text{(A.6)}
\end{aligned}
$$

Notice that the first expression in parentheses is strictly negative by $\theta^* < \mu_Y$ whereas the second expression in parentheses is strictly positive by $\theta^* \in (0, 1)$. $\tilde{u}''_{\theta^*}(\mu) < 0$ follows from this, $\hat{u}' > 0$ and $\hat{u}'' < 0$. $\tilde{u}'''_{\theta^*}(\mu) > 0$ follows from the signs of the expressions in parentheses, $\hat{u}'' < 0$ and $\hat{u}''' \geq 0$. q.e.d.

Proof of Proposition 3

Case 1: $\theta^* \in \Theta^{**} \backslash \Theta^*$. By Lemma 1, I have to solve the program

$$
\begin{aligned}
\max_\pi \quad & \mathbb{E}_\tau[\hat{v}(\mu)] \\
\text{s.t.} \quad & \mathbb{E}_{\tau_{\theta^*}}[\hat{u}(\mu)] = \overline{u}_{\theta^*} \\
& \pi \text{ and } \chi_{\theta^*} \text{ induce } \tau
\end{aligned} \qquad \text{(A.7)}
$$

Because of Lemma 2 (a), I can rewrite the first constraint as $\mathbb{E}_\tau[\tilde{u}_{\theta^*}(\mu)] = \overline{u}_{\theta^*}$ such that the distribution τ_{θ^*} faced by the threshold agent disappears in the program (A.7). Because updating is Bayesian, the distribution τ induced by any π and χ_{θ^*} is Bayes plausible, i.e., $\mathbb{E}_\tau[\hat{v}(\mu)] = \mu_Y$ with $\mu_Y = \mathbb{E}_F[\theta|\theta \geq \theta^*]$. By a reasoning like in the proof of Proposition 1 in Kamenica and Gentzkow (2011), any Bayes plausible distribution τ is induced by some π and χ_{θ^*}. This allows me to pursue an approach where I first determine the distribution τ^* that the solves the program

$$\begin{array}{ll} \max\limits_{\tau} & \mathbb{E}_\tau[\hat{v}(\mu)] \\ \text{s.t.} & \mathbb{E}_\tau[\tilde{u}_{\theta^*}(\mu)] = \overline{u}_{\theta^*} \\ & \mathbb{E}_\tau[\mu] = \mu_Y \end{array} . \tag{A.8}$$

Any test π^* that together with χ_{θ^*} induces τ^* is then optimal.

I can now handle the first inequality constraint with a Lagrangian approach. By the Lagrangian Sufficiency Theorem, the following is true: If there exists $\lambda \geq 0$ such that τ^* solves the program

$$\begin{array}{ll} \max\limits_{\tau} & \mathcal{L}(\tau, \lambda) \text{ (where } \mathcal{L}(\tau, \lambda) \text{ is as defined in (8))} \\ \text{s.t.} & \mathbb{E}_\tau[\mu] = \mu_Y \end{array} \tag{A.9}$$

and

$$\mathbb{E}_\tau[\tilde{u}_{\theta^*}(\mu)] = \overline{u}_{\theta^*}, \tag{A.10}$$

then τ^* is a solution to the program (A.8).

Step 1. Consider first the program (A.9) for a given value $\lambda \geq 0$. This program possesses a solution because it is just a concavification problem (see, e.g., Aumann and Maschler, 1995; Kamenica and Gentzkow, 2011). If $\lambda = 0$, $\varphi_{\lambda,\theta^*} = \hat{v}$. $\hat{v}'' > 0$ implies that the distribution induced by a perfectly informative test and χ_{θ^*} is optimal. It remains to consider $\lambda > 0$. Because in this case $\tilde{u}'''_{\theta^*} > 0$ by Proposition 2 (b) and because $\hat{v}''' \geq 0$, I have $\varphi'''_{\lambda,\theta^*} > 0$. That is, the function to be concavified has no linear portion. This has three important implications: First, the maximizer is unique. Second, because the concavification problem is one-dimensional, the maximizer must have a binary support. Third, it implies that the Lagrangian $\mathcal{L}(\tau, \lambda)$ is either concave, convex or first concave and the convex. As a consequence, the maximizer is either the degenerate distribution that puts probability 1 on $\mu = \mu_Y$ or a binary distribution where $\mu = 1$ belongs to the support. I will denote the optimal distribution by $\tau^*(\lambda)$. To sum up, I know from the reasoning in this paragraph that if there exists $\lambda \geq 0$ such that (A.10) holds for $\tau = \tau^*(\lambda)$, the program (A.7) is solved by a degenerate distribution or by a binary distribution where $\mu = 1$ belongs to the support. That is, if an optimal test exists, it can only be a completely uninformative or a NFP test. Since I know already that information generation is possible for $\theta^* \in \Theta^* \backslash \Theta^{**}$, only a NFP test can be optimal.

Step 2. Consider next (A.10) with $\tau = \tau^*(\lambda)$. By the reasoning in the preceding paragraph, $\tau^*(0)$ is the distribution induced by a perfectly informative test and χ_{θ^*}. On the other hand, because $\tilde{u}''_{\theta^*} < 0$ by Lemma 2 (b), $\varphi_{\lambda,\theta^*}(\mu)$ is strictly concave at $\mu = \mu_Y$ for any sufficiently large λ. Hence, $\tau^*(\lambda)$ is for any sufficiently large λ the distribution induced by a completely uninformative test and χ_{θ^*}. Now recall that $\mathbb{E}_\tau[\tilde{u}_{\theta^*}(\mu)]$ describes the threshold agent's expected payoff from participating in a test that induces together with χ_{θ^*} the distribution τ. $\theta^* \in \Theta^* \backslash \Theta^{**}$ implies that θ^* can be induced by some informative test but not by a perfectly informative test. This together with the previous two observations implies

$$\mathbb{E}_{\tau^*(0)}[\tilde{u}_{\theta^*}(\mu)] < \overline{u}_{\theta^*} < \lim_{\lambda \to \infty} \mathbb{E}_{\tau^*(\lambda)}[\tilde{u}_{\theta^*}(\mu)]. \tag{A.11}$$

Moreover, because $\mathcal{L}(\tau, \lambda)$ is continuous in τ and the set of Bayes plausible distributions does not depend on λ, it follows from Berge's Maximum Theorem (see, e.g., Aliprantis and Border, 2006) that the set of maximizers of the program (A.9) is upper hemicontinuous. Since the solution to the program (A.9) is unique, $\tau^*(\lambda)$ is continuous. Because of this and (A.11), I can apply an Intermediate Value Theorem. I obtain that there exists $\lambda^* > 0$ such that (A.10) holds for $\tau = \tau^*(\lambda^*)$. This concludes the proof for Case 1.

Case 2: $\theta^* \in \Theta^{**}$. χ_{θ^*} is then induced by any perfectly informative test π'. Notice that any perfectly informative test π' and χ_{θ^*} induce the same the distribution τ'. Notice further that this distribution is a mean-preserving spread of any distribution τ'' that is induced by any test π'' that is not perfectly informative and χ_{θ^*}. By this and $\hat{v}'' > 0$, $\mathbb{E}_{\tau'}[\hat{v}(\mu)] > \mathbb{E}_{\tau''}[\hat{v}(\mu)]$. Hence, the perfectly informative test π_ρ^{NFP} with $\rho = 1$ is optimal. q.e.d.

Proof of Corollary 1

To prove this result, I have to solve the problem that I solved in the proof to Proposition 3 with a modified objective function. As the objective function mattered only through $\hat{v}'' > 0$ and $\hat{v}''' \geq 0$ in this proof, I only have to show that the modified objective functions in Part (a) and Part (b) possesses these properties.

(a) Here I am interested in maximizing the objective function $\mathbb{E}_{\tau_{\theta^*}}[-\mu]$. By (4), I can rewrite this as $\mathbb{E}_\tau[-\hat{\mu}(\mu)]$ with

$$\hat{\mu}(\mu) \equiv \left(\frac{\theta^*}{\mu_Y}\mu + \frac{1-\theta^*}{1-\mu_Y}(1-\mu)\right)\mu.$$

Consistency of μ_Y with the participation strategy implies $\mu_Y > \theta^*$ for any $\theta^* \in \Theta^*$. Since this implies $\hat{\mu}'' < 0$ and since $\hat{\mu}''' = 0$, I obtain the result.

(b) Here I am interested in maximizing the objective function $\mathbb{E}_{\tau_{\theta^{**}}}[-\hat{u}(\mu)]$ with $\theta^{**} > \mu_Y$. By (4), I can rewrite this as $\mathbb{E}_\tau[-\tilde{u}_{\theta^{**}}(\mu)]$ with

$$\tilde{u}_{\theta^{**}}(\mu) \equiv \left(\frac{\theta^{**}}{\mu_Y}\mu + \frac{1-\theta^{**}}{1-\mu_Y}(1-\mu)\right)\hat{u}(\mu).$$

I get,

$$\tilde{u}''_{\theta^{**}}(\mu) = \hat{u}'(\mu)\left[2\left(\frac{\theta^{**}}{\mu_Y} - \frac{1-\theta^{**}}{1-\mu_Y}\right) - \left(\frac{\theta^{**}}{\mu_Y}\mu + \frac{1-\theta^{**}}{1-\mu_Y}(1-\mu)\right)\left(-\frac{\hat{u}''(\mu)}{\hat{u}'(\mu)}\right)\right], \textit{and}$$

$$\tilde{u}'''_{\theta^{**}}(\mu) = 3\left(\frac{\theta^{**}}{\mu_Y} - \frac{1-\theta^{**}}{1-\mu_Y}\right)\hat{u}''(\mu) + \left(\frac{\theta^{**}}{\mu_Y}\mu + \frac{1-\theta^{**}}{1-\mu_Y}(1-\mu)\right)\hat{u}'''(\mu).$$

Notice that the first expression in parentheses is strictly positive by $\theta^{**} > \mu_Y$ and that the second expression is also obviously strictly positive. $\tilde{u}''_{\theta^{**}} < 0$ follows from this, $\hat{u}' > 0$ and the assumption imposed on the Arrow-Prat measure of absolute risk-aversion. Finally, $\tilde{u}'''_{\theta^{**}} < 0$ follows from positiveness of the first expression in parentheses, $\hat{u}'' < 0$ and the assumption that $\hat{u}''' = 0$. q.e.d.

Proof of Lemma 3

I have

$$a^{*\prime}(\mu) = \frac{\gamma_L \gamma_H}{(\mu\gamma_L + (1-\mu)\gamma_H)^2} > 0,$$

$$a^{*\prime\prime}(\mu) = \frac{2\gamma_L\gamma_H(\gamma_H - \gamma_L)}{(\mu\gamma_L + (1-\mu)\gamma_H)^2}\begin{cases} > 0 & \text{if } \gamma_H > \gamma_L \\ < 0 & \text{if } \gamma_H < \gamma_L \end{cases}, \text{ and}$$

$$a^{*\prime\prime\prime}(\mu) = \frac{6\gamma_L\gamma_H(\gamma_H - \gamma_L)^2}{(\mu\gamma_L + (1-\mu)\gamma_H)^2} > 0. \qquad \text{q.e.d.}$$

Proof of Lemma 4

Step 1: Properties of the principal's reduced-form utility function. By differentiating $\hat{v}$, I get

$$\hat{v}'(\mu) = \gamma_L\gamma_H\frac{\gamma_L\mu^2 - \gamma_H(1-\mu)^2}{(\gamma_L\mu + \gamma_H(1-\mu))^2},$$

$$\hat{v}''(\mu) = \frac{2(\gamma_L\gamma_H)^2}{(\gamma_L\mu + \gamma_H(1-\mu))^3}, \text{ and}$$

$$\hat{v}'''(\mu) = \frac{6(\gamma_L\gamma_H)^2(\gamma_H - \gamma_L)}{(\gamma_L\mu + \gamma_H(1-\mu))^4}.$$

The inequalities in the lemma follow directly from this.

Step 2: Properties of the agent's utility function that will be important in Step 4. Some properties of the agent's reduced-form utility function depend on the Arrow Prat coefficient of absolute risk-aversion. This coefficient at a is defined as $r_A(a) = -u''(a)/u'(a)$. For HARA utility (see Section 1 of the supplementary material), it is given by

$$r_A(a) = \frac{1}{c_1 a + c_2}.$$

CARA utility corresponds to $c_1 = 0$ and $c_2 > 0$; CRRA utility corresponds to $c_1 > 0$ and $c_2 = 0$. In both cases, stronger risk-aversion corresponds to smaller values of the non-zero parameter.

I say that r_A is sufficiently large if $r_A(a) \geq \overline{r}$ for all $a \in [0,1]$ and some sufficiently large $\overline{r}$. In particular, r_A is sufficiently large if u is a CARA or a CRRA utility function that exhibits sufficiently strong risk aversion. Moreover, I say r_A is sufficiently small if $r_A(a) \leq \overline{r}$ for all $a \in [0,1]$ and some sufficiently small $\overline{r} > 0$. In particular, r_A is sufficiently small if u is CARA utility function that exhibits sufficiently weak risk-aversion.

Notice further that I have $-(c_1 a + c_2)u''(a) = u'(a)$ when u is a HARA utility function. By differentiating both sides of this equation and rearranging, I obtain that the coefficient of absolute prudence is

$$-\frac{u'''(a)}{u''(a)} = (1 + c_1)r_A(a).$$

Step 3: Properties of the principal's optimal estimate that will be important in Step 4. It follows from the formulas in the proof of Lemma 3 that

$$\frac{a^{*\prime\prime}(\mu)}{(a^{*\prime}(\mu))^2} = \frac{2(\gamma_H - \gamma_L)}{\gamma_L^2\gamma_H^2} \equiv C(\gamma_H, \gamma_L).$$

That is, the quotient is constant with respect to μ; its sign depends on the relative size of γ_H and γ_L.

Step 4: Properties of the agent's reduced-form utility function. First, I have

$$\hat{u}'(\mu) = u'(a^*(\mu))a^{*\prime}(\mu).$$

Because $u' > 0$ and $a^{*\prime} > 0$, $\hat{u}' > 0$. Second, I have

$$\hat{u}''(\mu) = u''(a^*(\mu))(a^{*\prime}(\mu))^2 + u'(a^*(\mu))a^{*\prime\prime}(\mu).$$

By using the definitions in Step 2 and Step 3, I can write this as

$$\hat{u}''(\mu) = u'(a^*(\mu))(a^{*\prime}(\mu))^2(C(\gamma_H, \gamma_L) - r_A(a^*(\mu))).$$

Consider first $\gamma_H > \gamma_L$. $C(\gamma_H, \gamma_L)$ is then positive. It follows from this, $u' > 0$ and $a^{*\prime} > 0$ that $\hat{u}'' > 0$ if r_A is sufficiently small and that $\hat{u}'' < 0$ if r_A is sufficiently large. Consider next $\gamma_H < \gamma_L$. $C(\gamma_H, \gamma_L)$ is then negative. This implies $\hat{u}'' < 0$. Third, I have

$$\hat{u}'''(\mu) = u'''(a^*(\mu))(a^{*\prime}(\mu))^3 + 3u''(a^*(\mu))a^{*\prime}(\mu)a^{*\prime\prime}(\mu) + u'(a^*(\mu))a^{*\prime\prime\prime}(\mu).$$

Consider first $\gamma_L > \gamma_H$. By Lemma 3 and my assumptions on u, each of the three summands is weakly positive and the last two are strictly positive. Hence, $\hat{u}''' > 0$. It remains to consider $\gamma_H > \gamma_L$ and HARA utility. By using the definitions in Step 2 and Step 3, I can write

$$\hat{u}'''(\mu) = -u''(a^*(\mu))(a^{*\prime}(\mu))^3\left[(1 + c_1)r_A(a^*(\mu)) - 3C(\gamma_H, \gamma_L)\right] + u'(a^*(\mu))a^{*\prime\prime\prime}(\mu).$$

First notice that $-u'' \cdot (a^{*\prime})^3$ and $u' \cdot a^{*\prime\prime\prime}$ are both strictly positive by $u' > 0$, $u'' < 0$ and Lemma 3. As also the remaining term is positive when r_A is sufficiently large, I obtain also $\hat{u}''' > 0$ in this case. q.e.d.

Proof of Proposition 4

Case 1: $\gamma_H > \gamma_L$. Under the conditions of Lemma 4 where $\hat{u}'' > 0$, the principal and the agent benefit both from information-generation. Perfect information-generation is possible with a perfectly informative test. In particular, the test π_1^{NFP} is optimal.

Under the conditions of Lemma 4 where $\hat{u}'' < 0$ the implied reduced-form utility functions satisfy all the conditions that I imposed on my reduced form model. Hence, Proposition 3 applies directly.

Case 2: $\gamma_L > \gamma_H$. I have in this case $\hat{u}'' < 0$ by Lemma 4. If $\theta^* \in \Theta^{**}$, a perfectly informative test is obviously optimal. I need thus only to consider the case where the principal wants to induce a participation threshold $\theta^* \in \Theta^* \backslash \Theta^{**}$. Before I come to the actual proof, I will derive two auxiliary results. The first auxiliary result shows that when I increase the accuracy of a NFP test for a given participation threshold, the principal's expected payoff increases whereas the threshold agent's expected payoff decreases.

Lemma A3 *Fix any $\theta^* \in [\underline{\theta}, \overline{\theta})$. Let τ_ρ and $\{\tau_{\theta,\rho}\}_{\theta\in\Theta}$ be induced by π_ρ^{NFP} and χ_{θ^*}. If $\hat{u}'' < 0$, then $\mathbb{E}_{\tau_\rho}[\hat{v}(\mu)]$ increases continuously and $\mathbb{E}_{\tau_{\theta^*,\rho}}[\hat{u}(\mu)]$ decreases continuously in ρ.*

Proof. Consider a NFP test with accuracy ρ. The quality perception associated with the positive result is then $\mu_2 = 1$ and the quality perception associated with the negative result is

μ_1 as given in (6). First notice that continuity of $\mathbb{E}_{\tau_\rho}[\hat{v}(\mu)]$ and of $\mathbb{E}_{\tau_{\theta^*,\rho}}[\hat{u}(\mu)]$ in ρ follows from continuity of μ_1 in ρ and from continuity of u and v in μ. I have

$$\frac{\mathrm{d}\mu_1}{\mathrm{d}\rho} = \frac{-\mu_Y(1-\mu_Y)}{(1-\rho\mu_Y)^2} = -\frac{\mu_Y}{1-\rho\mu_Y}(1-\mu_1) < 0. \tag{A.12}$$

Because of Bayesian plausibility, ρ does not affect $\mathbb{E}_{\tau_\rho}[\mu]$. Thus, a higher ρ transforms the principal's quality perception distribution by a mean-preserving spread. Because $\hat{v}'' > 0$ by Lemma 4, I obtain that $\mathbb{E}_{\tau_\rho}[\hat{v}(\mu)]$ increases in ρ.

I obtain from (7) that the expected perception from the threshold agent's perspective is decreasing in the test accuracy:

$$\frac{\mathrm{d}\mathbb{E}_{\tau_{\theta^*,\rho}}[\mu]}{\mathrm{d}\rho} = \theta^*(1-\mu_1) + (1-\rho\theta^*)\frac{\mathrm{d}\mu_1}{\mathrm{d}\rho} = \left[\theta^* - \mu_Y\frac{1-\rho\theta^*}{1-\rho\mu_Y}\right](1-\mu_1) < 0$$

The second equality follows from (A.12); the inequality follows from $\mu_Y > \theta^*$. It follows from this that an increase in ρ transforms the threshold agent's quality perception distribution by a mean-decreasing spread. Because $\hat{u}' > 0$ and $\hat{u}'' < 0$, I obtain that $\mathbb{E}_{\tau_{\theta^*,\rho}}[\hat{u}(\mu)]$ decreases in ρ. q.e.d.

The second auxiliary result shows that $\mathbb{E}_\tau[\hat{v}(\mu)]$ is an affine transformation of $\mathbb{E}_\tau[a^*(\mu)]$, and that this is in turn an affine transformation of $\mathbb{E}_{\tau_\theta}[a^*(\mu)]$.

Lemma A4 *Fix any $\theta^* \in [\underline{\theta}, \overline{\theta})$. Let τ' be induced by a test π' and χ_{θ^*}. Let τ'' be induced by a test π'' and χ_{θ^*}. Then, the following statements are equivalent:*

(i) $\mathbb{E}_{\tau'}[\hat{v}(\mu)] = \mathbb{E}_{\tau''}[\hat{v}(\mu)]$

(ii) $\mathbb{E}_{\tau'}[a^*(\mu)] = \mathbb{E}_{\tau''}[a^*(\mu)]$

(iii) $\mathbb{E}_{\tau'_\theta}[a^*(\mu)] = \mathbb{E}_{\tau''_\theta}[a^*(\mu)]$ *for all* $\theta \in \Theta$

Proof. It follows from (9) that

$$\gamma_H a^*(\mu)(1-\mu) = \gamma_L\mu(1-a^*(\mu)).$$

After rearranging, this becomes

$$\mu a^*(\mu) = \frac{\gamma_L}{\gamma_L-\gamma_H}\mu - \frac{\gamma_H}{\gamma_L-\gamma_H}a^*(\mu). \tag{A.13}$$

That is, $\mu a^*(\mu)$ is linear in μ and in $a^*(\mu)$. By using (9) and (A.13) to rewrite (10), I obtain

$$\hat{v}(\mu) \quad = \quad -\frac{\gamma_L\gamma_H}{\gamma_L-\gamma_H}(a^*(\mu)-\mu).$$

Because $\mathbb{E}_{\tau'}[\mu] = \mathbb{E}_{\tau''}[\mu]$ by Bayesian plausibility, I obtain immediately that (i) and (ii) are equivalent. Next suppose that τ is induced by π and χ_{θ^*}. By (4), I have

$$\mathbb{E}_{\tau_\theta}[a^*(\mu)] \quad = \quad \mathbb{E}_\tau[\left(\frac{\theta}{\mu_Y}\mu + \frac{1-\theta}{1-\mu_Y}(1-\mu)\right)a^*(\mu)].$$

By (A.13), this is a linear function of $\mathbb{E}_\tau[a^*(\mu)]$ and of $\mathbb{E}_\tau[\mu]$. As the latter is constant by Bayesian plausibility, I obtain the equivalence between (ii) and (iii). q.e.d.

By Lemma 1 and the reasoning in the proof of Proposition 3, I have to solve the program

$$\begin{aligned} \max_{\tau} \quad & \mathbb{E}_\tau[\hat{v}(\mu)] \\ \text{s.t.} \quad & \mathbb{E}_\tau[\tilde{u}_{\theta^*}(\mu)] = \overline{u}_{\theta^*} \\ & \mathbb{E}_\tau[\mu] = \mu_Y \end{aligned} \tag{A.14}$$

with $\hat{u} = u \circ a^*$ and $\hat{v} = v \circ a^*$. I will call any distribution τ with a binary support that contains $\mu = 1$ a NFP distribution. Notice that θ^* is optimally induced by a NFP test if, and only if, a NFP distribution solves the program (A.14).

Assume to the contrary that the the program (A.14) is not solved by a NFP distribution but by some other distribution τ'. Define $\overline{a}' \equiv \mathbb{E}_{\tau'}[a^*(\mu)]$ and $\overline{a}'_{\theta^*} \equiv \mathbb{E}_{\tau'_{\theta^*}}[a^*(\mu)]$. Consider the following auxiliary program:

$$\begin{aligned} \max_{\tau} \quad & \mathbb{E}_\tau[\tilde{u}_{\theta^*}(\mu)] \\ \text{s.t.} \quad & \mathbb{E}_\tau[a^*(\mu)] = \overline{a}' \ . \\ & \mathbb{E}_\tau[\mu] = \mu_Y \end{aligned} \tag{A.15}$$

In this program, I strive to maximize the threshold agent's expected utility subject to Bayesian plausibility and to keeping the expected estimate constant. If this program is solved by a NFP distribution, say τ'', I know that there exists a NFP test that is better for the threshold agent and equally good for the principal. The latter property follows from Lemma A4. The former property implies that the solution of the auxiliary program does not satisfy the participation constraint (that I did not impose on the auxiliary program). However, by Lemma A3 and the fact that the threshold agent has for $\theta^* \in \Theta^* \backslash \Theta^{**}$ a strict incentive not to participate when the test is perfectly informative, there exists another NFP distribution τ''' that is Bayes plausible, satisfies the participation constraint and is better for the principal than τ'. That is, I obtain a contradiction to the optimality of the non-NFP distribution τ'.

It remains for me to show that some NFP distribution maximizes indeed the auxiliary program (A.15). By using (4) to transform the objective function and the second constraint, and by using Lemma A4 to replace the first constraint by an equivalent constraint that depends on the expected perception from the agent's perspective, I can rewrite the program (A.15) as

$$\begin{aligned} \max_{\tau_{\theta^*}} \quad & \mathbb{E}_{\tau_{\theta^*}}[u(a^*(\mu))] \\ \text{s.t.} \quad & \mathbb{E}_{\tau_{\theta^*}}[a^*(\mu)] = \overline{a}'_{\theta^*} \qquad . \\ & \mathbb{E}_{\tau_{\theta^*}}\Big[\frac{\mu}{\frac{\theta^*}{\mu_Y}\mu + \frac{1-\theta^*}{1-\mu_Y}(1-\mu)}\Big] = \mu_Y \end{aligned} \tag{A.16}$$

$a^*(\mu)$ is strictly increasing with range $[0, 1]$ and inverse function

$$\mu^*(a) \equiv \frac{\gamma_H a}{\gamma_H a + \gamma_L(1-a)}.$$

It follows that there exists a one-to-one mapping between distributions of estimates on $[0, 1]$, say $\hat{\tau}$, and distributions of quality perceptions τ_{θ^*} with

$$\hat{\tau}(a) \equiv \tau_{\theta^*}(\mu^*(a)).$$

Hence, after substituting $\mu^*(a)$ for μ and simplifying the second constraint, I can rewrite the program (A.16) as the following Lagrangian problem with Lagrange multiplier λ on the second constraint:

$$\begin{aligned} \max_{\hat{\tau}} \quad & \tilde{\mathcal{L}}(\hat{\tau}, \lambda) \equiv \mathbb{E}_{\hat{\tau}}[\tilde{\varphi}_\lambda(a) - \lambda \mu_Y] \\ \text{s.t.} \quad & \mathbb{E}_{\hat{\tau}}[a] = \overline{a}'_{\theta^*} \end{aligned} \tag{A.17}$$

with

$$\tilde{\varphi}_\lambda(a) \equiv u(a) + \lambda \tilde{\mu}(a) \text{ and } \tilde{\mu}(a) \equiv \frac{\gamma_H a}{\frac{\theta^*}{\mu_Y}\gamma_H a + \frac{1-\theta^*}{1-\mu_Y}\gamma_L(1-a)}.$$

Notice that the principal maximizes in this program over distributions of estimations where the support is a subset of $[0,1]$ and the first moment is constant. Interestingly, this program resembles the program (A.9) in the proof of Proposition 3. u assumes the role of $\hat{v}$; $\tilde{\mu}$ assumes the role of $\tilde{u}_{\theta^*}$; the assumption that the first moment of the estimate distribution is constant assumes the role of Bayesian plausibility; the original Bayesian plausibility constraint assumes the role of the participation constraint; the distribution $\hat{\tau}$ of estimates assumes the role of the distribution τ of quality perceptions. Thus, to conclude this proof, I need only to show that the properties that are crucial in the proof of Proposition 3 are satisfied here.

Translated in the notation of program (A.17), crucial for Step 1 in the proof of Proposition 3 was $u''' \geq 0$ and $\tilde{\mu}''' > 0$. The first property is what I imposed on the agent's non-reduced utility function. The second property can easily be checked. I obtain

$$\begin{aligned} \tilde{\mu}'(a) &= \frac{\gamma_H \gamma_L \frac{1-\theta^*}{1-\mu_Y}}{\left(\frac{\theta^*}{\mu_Y}\gamma_H a + \frac{1-\theta^*}{1-\mu_Y}\gamma_L(1-a)\right)^2}, \\ \tilde{\mu}''(a) &= \frac{-2\gamma_H \gamma_L \frac{1-\theta^*}{1-\mu_Y}\left(\frac{\theta^*}{\mu_Y}\gamma_H - \frac{1-\theta^*}{1-\mu_Y}\gamma_L\right)}{\left(\frac{\theta^*}{\mu_Y}\gamma_H a + \frac{1-\theta^*}{1-\mu_Y}\gamma_L(1-a)\right)^3}, \text{ and} \\ \tilde{\mu}'''(a) &= \frac{6\gamma_H \gamma_L \frac{1-\theta^*}{1-\mu_Y}\left(\frac{\theta^*}{\mu_Y}\gamma_H - \frac{1-\theta^*}{1-\mu_Y}\gamma_L\right)^2}{\left(\frac{\theta^*}{\mu_Y}\gamma_H a + \frac{1-\theta^*}{1-\mu_Y}\gamma_L(1-a)\right)^4}. \end{aligned}$$

Since $\gamma_L > \gamma_H$ and $\mu_Y > \theta^*$ imply

$$\left(\frac{\theta^*}{\mu_Y}\gamma_H - \frac{1-\theta^*}{1-\mu_Y}\gamma_L\right) < 0,$$

it follows that $\tilde{\mu}'(a) > 0$, $\tilde{\mu}'' > 0$ and $\tilde{\mu}''' > 0$.

Denote the distribution of estimates that solves the program (A.17) for a given value of $\lambda \geq 0$ by $\hat{\tau}^*(\lambda)$. Crucial for Step 2 in the proof of Proposition 3 was the sign of $\mathbb{E}_{\hat{\tau}^*(\lambda)}[\tilde{\mu}(a)] - \mu_Y$. By assumption, there exists a distribution of estimates $\hat{\tau}'$ that has neither the support $\{0,1\}$ nor the support $\{\overline{a}'_{\theta^*}\}$ such that $\mathbb{E}_{\hat{\tau}'}[a] = \overline{a}'_{\theta^*}$ and $\mathbb{E}_{\hat{\tau}'}[\tilde{\mu}(a)] = \mu_Y$. For λ sufficiently high, strict convexity of $\tilde{\mu}(a)$ implies that $\hat{\tau}^*(\lambda)$ is the distribution with support $\{0,1\}$. It follows that $\hat{\tau}^*(\lambda)$ is a mean-preserving spread of $\hat{\tau}'$. Hence, by Jensen's inequality,

$$\lim_{\lambda \to \infty} \mathbb{E}_{\hat{\tau}^*(\lambda)}[\tilde{\mu}(a)] > \mu_Y.$$

For $\lambda = 0$, strict concavity of u implies that $\hat{\tau}^*(0)$ is the degenerate distribution with support $\{\overline{a}'_{\theta^*}\}$. It follows that $\hat{\tau}'$ is a mean-preserving spread of $\hat{\tau}^*(0)$. Jensen's inequality implies in this case

$$\mathbb{E}_{\hat{\tau}^*(0)}[\tilde{\mu}(a)] < \mu_Y.$$

Since the crucial properties of Step 1 and Step 2 in the proof of Proposition 3 are also satisfied here, I obtain that a NFP test is optimal by an analogous reasoning. q.e.d.

References

Aliprantis, C. D. and K. C. Border (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer.

Alonso, R. and O. Câmara (2016a). Bayesian persuasion with heterogeneous priors. *Journal of Economic Theory 165*, 672–706.

Alonso, R. and O. Câmara (2016b). Persuading voters. *American Economic Review 106*, 3590–3605.

Alós-Ferrer, C. and J. Prat (2012). Job market signaling and employer learning. *Journal of Economic Theory 147*, 1787–1817.

Arrow, K. J. (1973). Higer education as a filter. *Journal of Public Economics 2*, 193–216.

Aumann, R. J. and M. Maschler (1995). *Repeated Games and Incomplete Informatio*. The MIT Press.

Bar, T., V. Kadiyali, and A. Zussman (2012). Putting grades in context. *Journal of Labor Economics 30*, 445–478.

Benoît, J.-P. and J. Dubra (2004). Why do good cops defend bad cops? *International Economic Review 45*, 787–809.

Bergemann, D. and M. Pesendorfer (2007). Information structures in optimal auctions. *Journal of Economic Theory 137*, 580–609.

Calzolari, G. and A. Pavan (2006a). Monopoly with resale. *The RAND Journal of Economics 37*, 362–375.

Calzolari, G. and A. Pavan (2006b). On the optimality of privacy in sequential contracting. *Journal of Economic Theory 130*, 168–204.

Caplin, A. and K. Eliaz (2003). Aids policy and psychology: A mechanism-design approach. *The RAND Journal of Economics 34*, 631–646.

Daley, B. and B. Green (2014). Market signaling with grades. *Journal of Economic Theory 151*, 114–145.

De, S. and P. Nabar (1991). Economic implications of imperfect quality certification. *Economics Letters 37*, 333–337.

Dubey, P. and J. Geanakoplos (2010). Grading exams: 100,99,98,... or a,b,c? *Games and Economic Behavior 69*, 72–94.

Eső, P. and B. Szentes (2007). Optimal information disclosure in auctions and the handicap auction. *The Review of Economic Studies 74*, 705–731.

Farhi, E., J. Lerner, and J. Tirole (2013). Fear of rejection? tiered certification and transparency. *The RAND Journal of Economics 44*(4), 610–631.

Gentzkow, M. and E. Kamenica (2014). Costly persuasion. *American Economic Review: Papers & Proceedings 104*, 457–462.

Gentzkow, M. and E. Kamenica (2017). Competition in persuasion. *Review of Economic Studies 84*, 300–322.

Gill, D. and D. Sgroi (2008). Sequential decisions with tests. *Games and Economic Behavior 63*, 663–678.

Gill, D. and D. Sgroi (2012). The optimal choice of pre-launch reviewer. *Journal of Economic Theory 147*, 1247–1260.

Goldstein, I. and Y. Leitner (2015). Stress tests and information disclosure. Mimeo, November 2015.

Gossner, O. and J. Steiner (2016). Optimal illusion of control and related perception biases. Mimeo, December 2016.

Grossman, S. J. (1981). The informational role of warranties and private disclosure about product quality. *Journal of Law and Economics 24*, 461–483.

Grossman, S. J. and O. D. Hart (1980). Disclosure laws and takeover bids. *The Journal of Finance 35*, 323–334.

Harbaugh, R. and E. Rasmusen (2016). Coarse grades: Informing the public by withholding information. *AEJ - Microeconomics*, forthcoming.

Hirshleifer, J. (1971). The private and social value of information and the reward to inventive activity. *American Economic Review 61*, 561–574.

Hull, H. F., C. J. Bettinger, M. M. Gallaher, N. M. Keller, J. Wilson, and G. J. Mertz (1988). Comparison of hiv-antibody prevalence in patients consenting to and declining hiv-antibody testing in an std clinic. *JAMA: the Journal of the American Medical Association 260*, 935–938.

Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. *American Economic Review 101*, 2590–2615.

Kimball, M. S. (1990). Precautionary saving in the small and in the large. *Econometrica 58*, 53–73.

Kolotilin, A. (2015). Experimental design to persuade. *Games and Economic Behavior 90*, 215–226.

Kolotilin, A. (2016). Optimal information disclosure: A linear programming approach. Mimeo, October 2016.

Kolotilin, A., T. Mylovanov, A. Zapechelnyuk, and M. Li (2016). Persuasion of a privately informed receiver. Mimeo, November 2016.

Li, H. and W. Li (2013). Misinformation. *International Economic Review 54*, 253–277.

Lizzeri, A. (1999). Information revelation and certification intermediaries. *The RAND Journal of Economics 30*, 214–231.

Lyter, D. W., R. O. Valdiserri, L. A. Kingsley, W. P. Amoroso, and C. R. Rinaldo (1987). The hiv antibody test: Why gay and bisexual men want or do not want to know their results. *Public Health Reports 102*, 468–474.

Matthews, S. and A. Postlewaite (1985). Quality testing and disclosure. *The RAND Journal of Economics 16*, 328–340.

Menezes, C., C. Geiss, and J. Tressler (1980). Increasing downside risk. *American Economic Review 70*, 921–932.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics 12*, 380–391.

Milgrom, P. R. and R. J. Weber (1982). A theory of auctions and competitive bidding. *Econometrica 50*, 1089–1122.

Okuno-Fujiwara, M., A. Postlewaite, and K. Suzumura (1990). Strategic information revelation. *The Review of Economic Studies 57*, 25–47.

Ostrovsky, M. and M. Schwarz (2010). Information disclosure and unraveling in matching markets. *American Economic Journal: Microeconomics 2*, 34–63.

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

Pancs, R. (2014). Designing order-book transparency in electronic networks. *Journal of the European Economic Association 12*, 702–723.

Perez-Richet, E. (2014). Interim bayesian persuasion: First steps. *American Economic Review: Papers & Proceedings 104*, 469–474.

Perez-Richet, E. and D. Prady (2012). Complicating to persuade? Mimeo, February 2012.

Philippon, T. and V. Skreta (2012). Optimal interventions in markets with adverse selection. *American Economic Review 102*, 1–30.

Rayo, L. and I. Segal (2010). Optimal information disclosure. *Journal of Political Economy 118*, 949–987.

Rosar, F. (2016). Test design under voluntary participation. Mimeo, April 2016.

Schweizer, N. and N. Szech (2014). Optimal revelation of life-changing information. Mimeo, November 2014.

Spence, M. (1973). Job market signaling. *The Quaterly Journal of Economics 87*, 355–374.

Stiglitz, J. E. (1975). The theory of "screening," education and the distribution of income. *American Economic Review 65*, 283–300.

Taneva, I. (2016). Information design. Mimeo, October 2016.

Tirole, J. (2012). Overcoming adverse selection: How public intervention can restore market functioning. *American Economic Review 102*, 29–59.

Titman, S. and B. Trueman (1986). Information quality and the valuation of new issues. *Journal of Accounting and Economics 8*, 159–172.

Wang, Y. (2013). Bayesian persuasion with multiple receivers. Mimeo, June 2013.

Weiss, A. (1983). A sorting-cum-learning model of education. *Journal of Political Economy 91*, 420–442.