

The Design and Price of Certification

MIKAEL MÄKIMATTILA

Department of Economics, University of Oslo

YUCHENG SHANG

Department of Economics, Massachusetts Institute of Technology

RYO SHIRAKAWA

Department of Economics, Massachusetts Institute of Technology

We consider the design and pricing of certification contracts. A certifier offers a menu of tests to a sender, who holds partial private information about an unknown state and seeks to persuade a receiver. The selling mechanism gives information to the receiver through two channels: the actual informativeness of the tests and the sender's choice. In the revenue-maximizing menu, senders whose beliefs exceed an upper threshold purchase a common test, while senders whose beliefs fall below a lower threshold purchase no test. Those with intermediate beliefs

Mikael Mäkimattila: mikael.makimattila@hotmail.com

Yucheng Shang: ycshang@mit.edu

Ryo Shirakawa: shira723@mit.edu

Initial draft: December 2024. We are grateful to Ian Ball, Alessandro Bonatti, Yifan Dai, Sulagna Dasgupta, Glenn Ellison, Drew Fudenberg, Tan Gan, Caroline Liqui Lung, Matti Liski, Daniel Luo, Stephen Morris, Michèle Müller-Itten, Pauli Murto, James Okun, Julia Salmi, Hiroto Sato, Sara Shahanaghi, Ludvig Sinander, Roland Strausz, Jean Tirole, Juuso Välimäki, Mike Whinston, Alex Wolitzky, Yangfan Zhou and audiences at the 26th ACM Conference on Economics and Computation (2025), Berlin Micro Theory and Behavioral Economics PhD Conference (2025), the 36th Stony Brook International Conference on Game Theory (2025), and at seminars of Helsinki GSE, MIT, and Universidad Carlos III de Madrid for their insightful comments. This paper was written while Mikael Mäkimattila was a doctoral student at Aalto University. Shirakawa acknowledges the financial support from Funai Foundation for Information Technology. All errors are our own.

purchase distinct tests, so test choice perfectly reveals the sender's private information. All sender types who purchase a test are indifferent across all options in the menu, and the receiver obtains zero surplus. We also analyze how the results change if the certifier can conceal which test the sender selects.

KEYWORDS. Certification, Mechanism Design, Information Design, Bayesian Persuasion.

JEL CLASSIFICATION. D82, D83, L15.

1. INTRODUCTION

In many economic settings, a privately informed agent ("sender") attempts to persuade another agent ("receiver") to take a specific action by disclosing decision-relevant information. In such settings, a model's predictions usually depend on whether the sender can credibly commit to an information disclosure rule. While senders in many practical applications lack this commitment power, they can delegate the generation and transmission of information to a third party to leverage its institutional credibility. For example, students take standardized examinations administered by independent testing agencies to signal their academic potential to admissions committees. Similarly, firms undergo third-party audits to obtain quality certificates and communicate product information to otherwise uninformed consumers.

The reliance of senders on third-party certification introduces a potential conflict of interest. Certifiers are often profit-seeking; even non-profit organizations selling certification services (e.g., Educational Testing Service) may raise revenue to fund other activities. What is the impact of this incentive on information generation and transmission? How does a profit-maximizing certifier design and price its services?

To provide answers to these questions, we consider a model with three agents: a certifier, a sender and a receiver. The sender has imperfect private information about a binary state of the world (e.g., student's ability) with this belief as her type. The certifier offers a menu of tests with different prices, from which the sender observably chooses her preferred option. The test generates a certificate

which is informative about the state and therefore guides the receiver's action (e.g., admission decision), also assumed to be binary. The sender prefers the receiver to take the high action, while the receiver wants to do so only if the state is high. We focus on profit-maximizing certification design.

We decompose the certifier's problem into two steps. In the first step, the certifier chooses a partition determining which types purchase the same test. In the second step, the certifier designs a revenue-maximizing menu of tests subject to the constraint that sender types in the same element of the partition are incentivized to buy the same test. We argue that it is without loss to restrict attention to tests that generate simple "action recommendation" certificates. The certifier's problem in the second step is then to design an incentive-compatible, individually rational and receiver-obedient mechanism that assigns a binary-signal test and a related transfer to each element of the partition.

The partition is critical to characterizing the receiver's obedience because each test provides information to the receiver through two channels: the certificates generated by the test, and the sender's observable choice of test, which is informative about her private information. The sender's incentive compatibility turns out to require senders who are more certain that the state is high select tests with higher *differential informativeness* (the difference between the probability of recommending high action in the high versus the low state). Intuitively, such senders expect the information generated by tests to work in their favor. Higher differential informativeness also raises the sender's information rents by making her private information more relevant for test selection.

The main result identifies an optimal menu, parameterized by two thresholds. All belief types above the upper threshold purchase the same "Bayesian persuasion" test: one that recommends the high action in the high state and does so in the low state with a probability that leaves the receiver indifferent upon receiving the recommendation. Intermediate belief types are fully screened and purchase distinct tests, which can be viewed as discounted variants of the highest types' test. These tests share the same differential informativeness, but they do not always recommend the high action even in the high state. Belief types below the lower

belief threshold purchase no test. Since the obedience constraint for the high action always binds, the receiver obtains zero surplus.

The optimal menu balances several economic forces. One trade-off is between creating value for the sender and limiting information rents: a common Bayesian persuasion test creates value, but including lower types worsens the receiver's posterior upon test selection, so maintaining obedience would require a more informative test. A profit-maximizing certifier therefore sells this test only to the highest types to limit the sender's information rents. The menu structure also reflects a tension between the sender's incentive compatibility and the receiver's obedience: incentive compatibility pushes higher types toward more informative tests, whereas when higher types select a test, that test need not be very informative for the receiver to obey a high-action recommendation. This tension gives rise to an optimal menu consisting of tests with constant differential informativeness.

Real certifiers face trade-offs described in this article. They must balance value creation and senders' information rents and consider how senders' selection to tests interacts with the receiver's inference and decision-making. Profit-driven certifiers may exclude less optimistic senders from the market with pricing. Often, the certifiers only offer one test when the test choice is observable: a special case of the revenue-maximizing menu we describe. Sometimes, they offer different versions of a test (e.g., expensive retake options or differentiated certification programs), which the revenue-maximizing menu we derive can also be interpreted to represent.

The optimality of the menu we describe relies on how the sender's test selection reveals additional information to the receiver. If the certifier could conceal which test the sender selects, we show that the optimal menu could look very different. Under certain conditions on the distribution of the sender's belief, it becomes optimal to segment the market by offering two tests with heterogeneous informativeness: a cheap test outputting a high-action recommendation only in the high state purchased by the highest sender types, and an expensive test that guarantees a high-action recommendation in both states purchased by lower sender types.

1.1 Literature

This paper contributes to the growing literature on the sale of information. The most relevant work is [Bergemann et al. \(2018\)](#), which studies a principal who designs a menu of experiments for a partially informed decision maker to purchase. [Yang \(2022\)](#), [Ichihashi and Smolin \(2022\)](#), and [Park and Evans \(2025\)](#) apply this framework to monopoly-market settings, where sellers or buyers obtain information about valuations to guide their market decisions.¹ In contrast to these settings, our model assumes that those who pay for information do so not to make an informed decision, but to persuade another decision maker. This distinction introduces a novel technical challenge: the information sender's choice of a test is informative in itself, influencing the belief of the end player.

More broadly, many studies investigate the role of an information intermediary in models of verifiable communication, such as [Grossman \(1981\)](#) and [Milgrom \(1981\)](#), and Bayesian persuasion, as in [Kamenica and Gentzkow \(2011\)](#).² For example, [Lizzeri \(1999\)](#) analyzes a platform selling a certificate to an informed seller, demonstrating that equilibrium certification may reveal no information yet extract full surplus. In our model, full surplus extraction is generally impossible: generating information is necessary to maximize sender–certifier surplus, but doing so gives rise to information rents. [Ali et al. \(2022\)](#) studies a related model where the seller of an asset is initially uninformed and strategically conceals a certificate.³ Other works, such as [Albano and Lizzeri \(2001\)](#) and [Zapechelnyuk \(2020\)](#), explore the endogenous quality choices of sellers, deriving profit-maximizing and consumer-optimal certification designs, respectively. [Stahl and Strausz \(2017\)](#) analyze the implications of who pays for product certificates. Compared to these papers, ours

¹ See also [Haupt et al. \(2023\)](#), which studies a related problem in a competitive market. In contrast to these studies, each signal realization is assumed to be deterministic, and a profit-maximizing designer posts a price for each signal realization rather than for the entire information structure.

² See also, for instance, [Grossman and Hart \(1980\)](#) for models of verifiable communication, and [Rayo and Segal \(2010\)](#) for a model of Bayesian persuasion.

³ See also [Hörner and Skrzypacz \(2016\)](#) and [Mekonnen et al. \(2023\)](#) for dynamic models of markets with an information seller, although their settings are quite different from ours.

is unique particularly in that we build on [Bergemann et al. \(2018\)](#) by allowing the certifier full flexibility to design any menu of tests and screen sender types.

Also related is the fast-growing literature at the intersection of mechanism design and information design. For instance, [Bergemann and Pesendorfer \(2007\)](#) and [Bergemann et al. \(2022\)](#) examine the joint design of a product-selling mechanism and an information structure that shapes buyers' beliefs about product quality. More relevant is [Corrao \(2023\)](#), which examines an intermediary service between an informed and an uninformed decision maker. Unlike in our paper, where the certifier has complete flexibility in designing tests, information in his paper is not verifiable. The optimization problem we address diverges from those in this literature in that it involves optimizing not only the allocation of information but also the partitioning of sender types, which shapes the receiver's beliefs and incentives. In this regard, [Rayo \(2013\)](#) is also relevant, which extends the monopoly screening problem of [Mussa and Rosen \(1978\)](#) to a setting in which goods are used purely as signals of private information. In his model, the value of the goods is determined by the receiver's equilibrium posterior beliefs, which relates to our first-step optimization problem of designing a partition. In our setting, the certifier sells and designs a Blackwell experiment that is itself informative about the state, after which the receiver updates her belief once more.

Recent concurrent studies explore how an intermediary designs a menu of tests.⁴ [Celik and Strausz \(2025\)](#) study a certifier in a market with a partially informed seller and an uninformed buyer, where the certifier offers a menu of tests to the seller. The key economic difference is that we consider a classic Bayesian persuasion setting with no transfers between the sender and the receiver, whereas they consider a traditional market setting with transfers between the seller and the buyer. This makes the seller's payoff convex in the buyer's posterior belief, resulting in their optimal menu fully revealing the seller's private information. [Weksler and Zik \(2023\)](#) study an analogous problem in which the certifier can sell to the sender

⁴ An earlier work by [Hancart \(2024\)](#) considers a similar problem, but in his setting, it is the information receiver who chooses a menu out of a limited selection of options, among which the information sender selects a test for free.

1 a test from a fixed, exogenously given finite set. [Boleslavsky and Kolb \(2025\)](#) 1
2 also analyze an intermediary providing a menu of tests to an information user 2
3 who leverages them to persuade a receiver. A key distinction from our model is 3
4 that they exclude the selection effect, meaning that the choice of tests does not 4
5 convey additional information beyond the content of the tests themselves, which 5
6 allows them to pointwise maximize the revenue function, an approach not generally 6
7 applicable in our setting. [Chopra and Ely \(2025\)](#) study a related model in which the 7
8 receiver does not observe the sender's test choice, an environment we also study 8
9 in Section 4 as a benchmark.⁵ Finally, [Ely \(2026\)](#) studies the optimal design of tests 9
10 that can also screen sender types. In his model, the sender's private information 10
11 is hard information such as résumés and is automatically revealed to the receiver; 11
12 nevertheless, he proves that an optimal menu exhibits a two-threshold structure 12
13 similar to ours. 13

14 The remainder of the paper is organized as follows. In Section 2, we introduce 14
15 the model and formulate the optimization problem we solve. The main theorem 15
16 and its implications are presented in Section 3. Section 4 discusses an alternative 16
17 model where the receiver does not observe the sender's test choice. Finally, Section 17
18 5 concludes the paper. Proofs omitted from the main text are provided in the 18
19 Appendix. 19

20 20

21 21

22 22

2. MODEL

23 There are three economic agents in the model: a certifier, a sender and a receiver. 23
24 The unknown state of the world is binary $\omega \in \Omega = \{L, H\}$, low or high, with a 24
25 common prior $\xi \in (0, 1)$ for the high state $\omega = H$. 25

26 In the beginning, the sender privately observes a signal that is informative about 26
27 the state and updates her belief according to Bayes' rule. This generates sender's 27
28 interim belief $\theta \in \Theta = [0, 1]$ about the probability that the state is high. Neither the 28
29 certifier nor the receiver can observe this belief, and θ is the sender's private type. 29
30 Let F be the induced cumulative distribution function of sender types. Assume F 30

31 31

32 ⁵ We discuss the connection to this paper further in Section 4. 32

admits an absolutely continuous density f that is strictly positive on $[0, 1]$. The interim belief is consistent with the prior distribution, i.e., $\mathbb{E}[\theta] = \xi$.

The certifier designs a menu of tests and associated prices, $\mathcal{M} = (\mathcal{E}, t)$, and offers it to the sender in order to maximize expected revenue. The first component $\mathcal{E}$ is a collection of tests, each of which is a pair of a signal space S and a stochastic map $\pi : \Omega \rightarrow \Delta(S)$, where we write $\pi(s|\omega) = \pi(\omega)(s)$ to represent the probability of signal realization s under state ω . Each test is associated with one price according to a map $t : \mathcal{E} \rightarrow \mathbb{R}$.⁶

Upon observing a menu of tests $\mathcal{M} = (\mathcal{E}, t)$, the sender chooses which test $E \in \mathcal{E}$ to purchase, if any, with payment $t(E)$. The test is then conducted, and a public signal is realized.

The receiver observes both which test is chosen and the signal realization. Then, she forms a posterior belief and takes action $a \in A = \{a_L, a_H\}$. The receiver's payoff $v_R(a, \omega)$ depends on her action and the state, and she wants to match the state: $v_R(a_H, H) > 0 > v_R(a_H, L)$ and we normalize that $v_R(a_L, L) = v_R(a_L, H) = 0$. Her payoff function determines threshold belief

$$\bar{\mu} = -\frac{v_R(a_H, L)}{v_R(a_H, H) - v_R(a_H, L)}$$

such that she weakly prefers to take the high action if and only if her posterior belief that the state is high weakly exceeds the threshold.⁷ In contrast, the sender prefers that she takes high action a_H independently of the state. The sender's payoff is $\mathbb{I}\{a = a_H\} - t(E)$, where $\mathbb{I}$ is an indicator function and $t(E)$ is a payment.

Assume that if the sender does not purchase any test, the receiver takes low action and the sender receives the payoff of zero. In the Appendix, we prove that this assumption entails no loss of generality.⁸

⁶ As in Bergemann et al. (2018), the price cannot be conditioned on the signal that results from the test.

⁷ If she has a belief μ , she takes high action if $\mu \cdot v_R(a_H, H) + (1 - \mu) \cdot v_R(a_H, L) \geq 0$. Rearranging terms, we get the expression for $\bar{\mu}$.

⁸ Intuitively, whenever some types choose the outside option, we can construct an alternative equilibrium in which the certifier offers those types an uninformative test, and choosing no test is off path and assigned a pessimistic belief.

In summary, the timing of the game is as follows:

1. The certifier posts a menu of tests $\mathcal{M} = (\mathcal{E}, t)$.
2. The sender observes private information θ .
3. The sender chooses whether to buy a test $E \in \mathcal{E}$ at price $t(E)$. If the sender does not purchase any test, the game ends and all players receive a payoff of zero.
4. The receiver observes the chosen test E and a signal realization s .
5. The receiver updates her belief about the state ω and takes an action a .

We analyze the optimal design of tests and its implications. In Section 4, we consider an alternative model in which the chosen test is not observable to the receiver.

Throughout the rest of the analysis, we focus on certifier-optimal perfect Bayesian equilibria in pure strategies. This restriction to pure strategies matters when we rewrite the certifier's problem in terms of partitions in the next subsection. In Appendix B, we show that for any finite menu, restricting attention to pure sender strategies is without loss of revenue.

The outcome is trivial when $\xi \geq \bar{\mu}$. To illustrate this point, consider a menu that consists of a single null test $E_\emptyset = (\{\emptyset\}, \pi_\emptyset)$ that always outputs the same message $\emptyset$. Let the price of $E_\emptyset$ be 1, i.e., $t(E_\emptyset) = 1$. Then, if $\xi \geq \bar{\mu}$, we have an equilibrium in which the sender always purchases the test and the receiver then always takes high action a_H . If the sender does not buy a test, the sender is punished with pessimistic beliefs and low action a_L . In other words, the certifier can extract the full surplus, even though persuasion has no value to the sender. Henceforth, we assume $\xi < \bar{\mu}$.

2.1 Preliminary Analysis

We decompose the certifier's problem into two steps: In the first stage, we choose and fix a partition $\mathcal{P}$ over the type space $\Theta = [0, 1]$. In the second stage, we have the certifier design a revenue-maximizing menu $\mathcal{M}$ subject to the constraint that sender types in the same element in $\mathcal{P}$ purchase a common test.

A revelation-principle style argument implies that we can restrict attention to direct incentive-compatible mechanisms

$$\mathcal{M} = \{(E(\theta), t(\theta)) | \theta \in [0, 1]\}$$

such that the sender of type θ chooses $E(\theta) = (S(\theta), \pi(\theta))$ at price $t(\theta)$.

Furthermore, the outcome of any menu can be attained by a menu where the signals from tests are action recommendations that the receiver wants to obey, that is, $S(\theta) = A = \{a_L, a_H\}$ for any θ . The reason is that if there are several signals that lead to the same receiver's action, the signals could be combined into one action recommendation without affecting the receiver's willingness to choose the same action after seeing this recommendation.⁹ Hence, we can represent any test $E(\theta)$ by a pair $(\pi^\omega(\theta))_{\omega \in \Omega}$, where

$$\pi^\omega(\theta) = \pi(\theta)(a_H | \omega)$$

is the probability that $E(\theta)$ recommends to take high action under state ω . We call $x(\theta) = \pi^H(\theta) - \pi^L(\theta)$ *differential informativeness* of the test, since larger $x(\theta)$ implies that a test differentiates better between the two states.¹⁰

Therefore, for each given partition $\mathcal{P}$, the certifier designs a direct mechanism consisting of $E : \Theta \rightarrow [0, 1]^2$ and $t : \Theta \rightarrow \mathbb{R}$ such that the sender of each type θ will optimally purchase $E(\theta)$ at price $t(\theta)$, the receiver will follow the action recommendations, and $E(\theta)$ and $t(\theta)$ are constant within each element $P \in \mathcal{P}$.

The receiver learns about the state through two channels: a test that the sender selects and the sender's selection to that test. On observing that the test for types in an element P is selected, the receiver believes that the sender has type $\theta \in P$. She

⁹ See, e.g., [Bergemann et al. \(2018\)](#). Here, to ensure that the receiver's interim belief about sender types remains unchanged even when this operation happens to merge two distinct tests into a common one, we allow a menu to include multiple identical tests.

¹⁰ Section 3 of [Bergemann et al. \(2018\)](#) takes a similar approach and summarizes a test by a one-dimensional variable, which they also term differential informativeness. While their definition differs from our definition of $x(\theta)$, its role for the test buyer's preferences over tests is analogous. See also [Weksler and Zik \(2023\)](#), who introduce a different informativeness parameter in a related context.

therefore updates her belief even without seeing any signal to

$$\mu_P = \mathbb{E}[\theta | \theta \in P].$$

The receiver will then observe a signal recommending an action. Given the interim belief μ_P , she has an incentive to follow the recommendation if

$$\frac{\mu_P(1 - \pi^H)}{\mu_P(1 - \pi^H) + (1 - \mu_P)(1 - \pi^L)} \leq \bar{\mu} \leq \frac{\mu_P\pi^H}{\mu_P\pi^H + (1 - \mu_P)\pi^L}, \quad (\text{OB})$$

where each inequality is required only when the denominator of the corresponding fraction is strictly positive. The first inequality is the so-called *obedience constraint* for taking low action, and the second inequality is that for high action. In (OB), we write π^H and π^L for the (common) test assigned to the partition element P , i.e., $\pi^H(\theta) = \pi^H$ and $\pi^L(\theta) = \pi^L$ for all $\theta \in P$.

To explore the sender's incentives, notice that the expected payoff for the sender of each type θ from reporting a type $\hat{\theta}$ can also be written as

$$V(\hat{\theta}; \theta) = \theta\pi^H(\hat{\theta}) + (1 - \theta)\pi^L(\hat{\theta}) - t(\hat{\theta}) = \theta x(\hat{\theta}) - \hat{t}(\hat{\theta}),$$

where $\hat{t}(\hat{\theta}) = t(\hat{\theta}) - \pi^L(\hat{\theta})$. Incentive compatibility requires this to be maximized at $\hat{\theta} = \theta$. Therefore, as in Myerson (1981), a mechanism is incentive compatible for the sender if and only if $x(\theta)$ is non-decreasing in θ and

$$V(\theta) = \theta x(\theta) - \hat{t}(\theta) = V(\tilde{\theta}) + \int_{\tilde{\theta}}^{\theta} x(s) ds, \quad (\text{EV})$$

for any θ and $\tilde{\theta}$, where we let $V(\theta) = V(\theta; \theta)$. Moreover, as in the classic mechanism-design literature, we can use this equation under $\tilde{\theta} = 0$ to rewrite the revenue function,

$$\mathbb{E}[t(\theta)] = \mathbb{E}[\hat{t}(\theta) + \pi^L(\theta)] = -V(0) + \mathbb{E}[\psi(\theta)x(\theta) + \pi^L(\theta)],$$

where $\psi(\theta) = \theta - [1 - F(\theta)]/f(\theta)$ is the so-called *virtual type function*.

In a revenue-maximizing menu, optimality requires that individual rationality, $V(\theta) \geq 0$ for each type θ , must hold with equality for some type θ_0 . Given that $V(\theta_0) = 0$, the individual rationality can be rephrased as

$$x(\theta) \geq 0 \quad \text{for } \theta > \theta_0, \quad \text{and} \quad x(\theta) \leq 0 \quad \text{for } \theta < \theta_0, \quad (\text{IR})$$

which follows from equation (EV) at $\tilde{\theta} = \theta_0$ and the monotonicity of $x(\theta)$. It is not a priori obvious that the individual-rationality constraint binds at the lowest type, because x may in general take negative values and hence V need not be increasing. That said, we show in the Appendix that (OB) implies that x must be non-negative and that (IR) binds at the lowest type.

Summarizing our discussion thus far, we arrive at the following observation concerning the optimal revenue. Note that the first term in the objective function is exactly equal to $-V(0)$.¹¹

LEMMA 1. The maximum revenue under a given partition $\mathcal{P}$ is equal to the value of the following problem:

$$\max_{\theta_0 \in [0,1], \pi^\omega: [0,1] \rightarrow [0,1]} \int_0^{\theta_0} x(\theta) d\theta + \int_0^1 (\psi(\theta)x(\theta) + \pi^L(\theta)) dF(\theta) \quad (\text{R})$$

subject to the constraints that

$$x(\theta) \text{ is non-decreasing in } \theta, \quad (\text{IC})$$

$$\pi^\omega(\theta) = \pi^\omega(\theta') \text{ for all } \theta, \theta' \in P \text{ and } P \in \mathcal{P}, \quad (\text{F})$$

and the individual rationality (IR) and the obedience conditions (OB).

Two remarks are in order. First, although the partition $\mathcal{P}$ does not appear in (IC), feasibility (F) requires that x be constant on each element $P \in \mathcal{P}$. Second, we allow the same test to be offered on two distinct elements of the partition. That is, we allow two tests with different labels to have the same properties.

¹¹ The formulation here is related to Bergemann et al. (2018), which analyzes the problem of selling experiments to a partially informed buyer who uses them to guide her *own* action. In particular, the incentive compatibility constraint is closely related, describing how buyers with different private information about the state select tests with different informational properties. The main difference is obedience: in Bergemann et al. (2018), obedience is imposed directly on the buyer, whereas here obedience is imposed on the receiver, who has different beliefs and incentives than the buyer.

A novel aspect of our model is that the sender's choice of a test carries a *signaling effect*, influencing the set of obedience constraints through the partition $\mathcal{P}$. In words, the choice of a test itself provides the receiver with information that the sender has about the state ω . Thus, the shape of $\mathcal{P}$ is not only critical for the certifier to screen the sender types but also determines the sender's flexibility in signaling her type to the receiver. Hence, the problem we study has features of mechanism design, Bayesian persuasion, and signaling.

We focus on measurable rules $\pi = (\pi^H, \pi^L) : [0, 1] \rightarrow [0, 1]^2$, which is necessary for the certifier's problem to be well-defined. We also need each element of the partitions $\mathcal{P}$ under consideration to be measurable. To ensure this, we take $\mathcal{P}$ to be the level-set partition induced by π ; that is, each element of $\mathcal{P}$ is of the form $\{\theta \in [0, 1] \mid \pi(\theta) = \bar{\pi}\}$ for some $\bar{\pi} \in [0, 1]^2$. This entails no loss, because whenever different elements of a partition are assigned the same test, we can pool these types and construct a coarser partition, which weakly relaxes the obedience conditions while leaving feasibility otherwise unchanged. Since π is measurable, the conditional law of θ given π admits a regular version, i.e., a measurable selection $m(\bar{\pi}) = \mathbb{E}[\theta \mid \pi = \bar{\pi}]$. We fix a version of the receiver's conditional beliefs and require the obedience conditions to hold almost surely with respect to the induced distribution of selected π .¹²

We say that a partition is *monotone* if each element in the partition is an interval, i.e., is of a form $\{a\}$, (a, b) , $[a, b]$, $(a, b]$, or $[a, b)$ for some $a \leq b$. Call a menu monotone if it is feasible under a monotone partition. For any partition $\mathcal{P}$, we write $\mathcal{P}^*$ to be the finest coarsening of $\mathcal{P}$ such that $\mathcal{P}^*$ is monotone. Namely, $\mathcal{P}^*$ is the finest monotone partition such that for any $P \in \mathcal{P}$, there exists $P^* \in \mathcal{P}^*$ such that $P \subset P^*$.

It might be worth noting that $\mathcal{P}^*$ is well-defined. Clearly, there exists a monotone coarsening of $\mathcal{P}$: $\{[0, 1]\}$. Let $(\mathcal{P}_i)_{i \in I}$ be the collection of all monotone coarsening of $\mathcal{P}$. Then, for each $P \in \mathcal{P}$, we can define a collection $(P_i)_{i \in I}$ such that $P \subset P_i$ and $P_i \in \mathcal{P}_i$ for each $i \in I$. Then, let $P^* = \bigcap_{i \in I} P_i$. One can see that the induced partition

¹² To simplify the exposition, we omit the almost-sure qualifier throughout the remainder of the paper.

$\mathcal{P}^* = \{P^* | P \in \mathcal{P}\}$ is also a monotone coarsening of $\mathcal{P}$. Moreover, it is a refinement of every partition in the collection $(P_i)_{i \in I}$.

3. OPTIMAL CERTIFICATION DESIGN

The main theorem of our study characterizes the optimal menu. Henceforth, we often abuse terminology and identify a menu with its tests alone, omitting payments. This is without loss by revenue equivalence: the tests and the envelope formula (EV) pin down payments up to a constant, which is fixed by binding individual rationality.

Fix any two numbers θ_* and θ^* with $0 \leq \theta_* \leq \theta^* \leq \bar{\mu}$. Let

$$x^* = \frac{\bar{\mu} - \mu^*}{\bar{\mu}(1 - \mu^*)} \quad \text{and} \quad \pi_\theta^L = \frac{\theta(1 - \bar{\mu})}{\bar{\mu} - \theta} x^*,$$

for each $\theta \in (\theta_*, \theta^*)$, where $\mu^* = \mathbb{E}[\theta | \theta \geq \theta^*]$.¹³ Then, consider a menu

$$(\pi^H(\theta), \pi^L(\theta)) = \begin{cases} (1, 1 - x^*) & \text{if } \theta \in [\theta^*, 1], \\ (x^* + \pi_\theta^L, \pi_\theta^L) & \text{if } \theta \in [\theta_*, \theta^*), \\ (0, 0) & \text{if } \theta \in [0, \theta_*), \end{cases} \quad (\text{OPT})$$

Write $\mathcal{M}(\theta_*, \theta^*)$ for the menu (OPT). In this menu, the highest types $\theta \in [\theta^*, 1]$ purchase a common test. Intermediate types $\theta \in [\theta_*, \theta^*)$ purchase different tests so that their test selection perfectly reveals their private information. The lowest types $\theta \in [0, \theta_*)$ purchase no test. We elaborate on the intuition behind the construction of each term later. Figure 1 illustrates the menu.

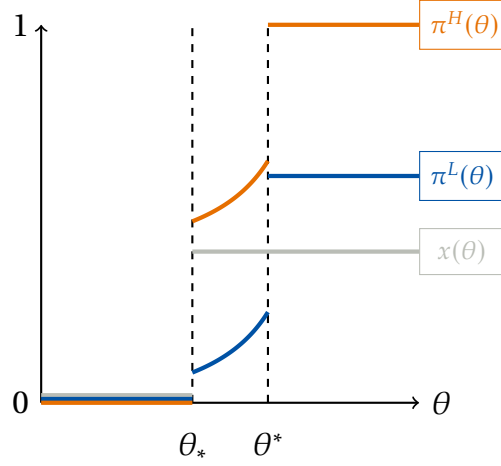
THEOREM 1. *Menu $\mathcal{M}(\theta_*, \theta^*)$ is optimal for some $0 < \theta_* \leq \theta^* < \bar{\mu}$.*

Before turning to the intuition behind Theorem 1, we provide several implications of the result. First, the optimal partition is monotone; that is, it consists of intervals.

COROLLARY 1. The optimal partition is monotone.

Second, it is natural to interpret the free test $(\pi^H, \pi^L) = (0, 0)$ as an opt-out option, i.e., purchasing no test. Under this interpretation, the optimal menu has a cutoff

¹³ Note that this menu is well-defined only when $\mu^* \leq \bar{\mu}$, which implies $x^* \in [0, 1]$.

FIGURE 1. Menu $\mathcal{M}(\theta_*, \theta^*)$.

structure: a sender of type θ purchases a test if and only if $\theta \geq \theta_*$. Moreover, the differential informativeness of every test in the menu is a common constant x^* . Since sender types' preferences differ only with respect to their valuation of differential informativeness, whenever the sender chooses to purchase a test, she is indifferent between all tests in the menu. The following corollary summarizes this point.

COROLLARY 2. The optimal menu has the same differential informativeness across all tests except the opt-out option $(0, 0)$. Every sender type that purchases a test is therefore indifferent among them.

Corollary 2 also suggests that equilibria under the optimal menu need not be unique. For example, holding all else equal, there exists another equilibrium in which all middle types $\theta \in A = [\theta_*, \theta^*)$ choose the test designed for type $\mathbb{E}[\theta | \theta \in A]$. By construction, the obedience constraint remains satisfied, i.e., the receiver optimally follows the recommendations. Moreover, sender types in A remain indifferent across all options in the menu, and hence their choices are still optimal.

This alternative equilibrium yields lower expected revenue than the truth-telling equilibrium.¹⁴

Third, both x^* and each π_θ^L are constructed such that the obedience constraint for taking high action binds conditional on each test being chosen. This means that, conditional on observing the sender's choice and signal realizations, the receiver always weakly prefers to take low action. Therefore, it is one of the optimal strategies for the receiver to act as if she receives no information, and thus, the receiver's expected payoff is minimized among all menus.

COROLLARY 3. Under the optimal menu, the receiver's expected payoff is zero conditional on any chosen test and any signal realization.

Fourth, the optimal menu fully reveals the private information of all intermediate types to the receiver while pooling the remaining two regions $[0, \theta_*)$ and $[\theta^*, 1]$. The test sold to the highest element is a classic Bayesian persuasion test that maximizes the sender's surplus (excluding transfers) conditional on the sender's type being in interval $[\theta^*, 1]$: this test is part of the menu to create value for the sender.¹⁵ The tests sold to lower types $\theta \in [\theta_*, \theta^*)$ are not classic Bayesian persuasion tests in that these tests do not always recommend the high action in the high state. If they did, the obedience constraint would require these tests to be more informative than the test sold to high sender types to counter the receiver's worse belief upon the sender choosing a lower type's test. This would conflict with incentive compatibility that requires the lower types' tests to have lower differential informativeness.

The optimal thresholds θ_* and θ^* balance the gains from increasing the sender's value, that is, the joint surplus of the certifier and the sender, against the gains from limiting her information rents. Higher differential informativeness increases the sender's information rent because it makes her private information more relevant

¹⁴ However, note that the certifier is able to achieve a revenue arbitrarily close to the optimal revenue using a menu where $x(\theta)$ strictly increases in θ on $[\theta_*, \theta^*]$ and each sender type therefore strictly prefers her equilibrium test over other tests.

¹⁵ By a "Bayesian persuasion test", we refer to a test that always recommends the high action in the high state and does so in the low state with a probability that leaves the receiver indifferent upon receiving the recommendation (see [Kamenica and Gentzkow, 2011](#)).

1 for test selection. The certifier can increase the sender's value by decreasing θ^* and 1
 2 thus selling the Bayesian persuasion test to a greater set of types. However, lowering 2
 3 θ^* worsens the receiver's belief conditional on the sender choosing the test of types 3
 4 $[\theta^*, 1]$ and therefore requires that test to be more differentially informative. The 4
 5 resulting increase in information rents can be so large that the certifier's revenue 5
 6 falls even though the sender's value increases. The certifier can also increase the 6
 7 sender's value by slightly decreasing θ_* and thereby selling "personalized" tests to 7
 8 a larger set of types. But this raises differential informativeness at θ_* from 0 to x^* , 8
 9 which in turn increases the information rents of all higher sender types. Indeed, 9
 10 the certifier's revenue from the menu $\mathcal{M}(\theta_*, \theta^*)$ can be written as a linear functional 10
 11 of $x(\theta)$, so that if $\psi(\theta_*) < 0$, a marginal decrease in the cutoff θ_* reduces revenue. 11

12 Fifth, note that $\mu^* = \mathbb{E}[\theta | \theta \geq \theta^*] > 0$ and therefore $x^* < 1$. This means that the 12
 13 optimal menu never contains a test that perfectly reveals the state, whose differential 13
 14 informativeness must equal one. This contrasts with the result in [Bergemann et al.](#) 14
 15 (2018), who show that the optimal menu always includes the fully informative test 15
 16 when the information buyer uses the experiment to guide her own action. However, 16
 17 our result is consistent with theirs in that the optimal menu contains a test that is 17
 18 optimal for the information buyer. In particular, the test for the highest element 18
 19 $[\theta^*, 1]$ is sender-optimal in the classic Bayesian persuasion setting if the common 19
 20 prior were given by μ^* . In this sense, we obtain the usual "no-distortion-at-the-top" 20
 21 result familiar in the mechanism design literature. 21

22 Sixth, the price schedule of the tests that the certifier optimally offers is such 22
 23 that the threshold type θ_* is indifferent whether to purchase a test. Given that the 23
 24 information rent type $\theta > \theta_*$ receives is 24

$$\int_0^\theta x(s)ds = \int_{\theta_*}^\theta x(s)ds = \int_{\theta_*}^\theta x^*ds,$$

27 types $\theta > \theta_*$ gain positive information rents if and only if $x^* > 0$. We therefore 27
 28 obtain the following corollary. 28

29
 30 **COROLLARY 4.** Under the optimal menu $\mathcal{M}(\theta_*, \theta^*)$, the sender's expected payoff is 30
 31 zero if and only if $x^* = 0$, i.e., if and only if $\mu^* = \bar{\mu}$. In this case, we have $\theta_* = \theta^*$, and 31
 32 $\mathcal{M}(\theta_*, \theta^*)$ consists of a single test that always recommends the high action. 32

The last observation in the corollary follows from noting that $\pi^\omega(\theta) = 0$ for all $\theta < \theta^*$ and $\omega \in \{L, H\}$ if $x^* = 0$.

When the certifier sells only a single test that always recommends the high action, certification does not generate any additional information beyond what the sender already knows. The sender's sole choice to take the test is then informative enough for the receiver to want to take the high action. Although the certifier may optimally choose not to generate new information through the test outcome, the optimal menu still always exploits the sender's selection as an information channel: only types above the strictly positive cutoff θ_* purchase a test.

Starting from a menu consisting of a single test that always recommends the high action, and computing the certifier's marginal revenue benefit from lowering the threshold $\theta^* = \theta_*$ and increasing the test's differential informativeness accordingly, we obtain a condition under which certification generates new information and the sender obtains a positive surplus.

PROPOSITION 1. Let $\hat{\theta}$ be such that $\mathbb{E}[\theta | \theta \geq \hat{\theta}] = \bar{\mu}$. If

$$\frac{1 - \hat{\theta}}{\bar{\mu}} < \frac{1 - \bar{\mu}}{\bar{\mu} - \hat{\theta}},$$

then the sender obtains a positive expected surplus under the certifier-optimal menu.

Proposition 1 gives a condition under which the sender receives a positive surplus. For example, in the case of uniform distribution, the condition of the proposition can be written as $\bar{\mu} > 2/3$.

Finally, it is insightful to compare the revenue-maximizing menu to menus designed with alternative objectives. Call a test $(\pi^\omega)_{\omega \in \{L, H\}}$ *sender-optimal* at belief μ if it maximizes sender μ 's payoff $\mu\pi^H + (1 - \mu)\pi^L$ subject to (OB) at μ .

PROPOSITION 2. A menu that maximizes the joint surplus of the certifier and the sender consists of a single test such that

$$\pi^H = 1 \quad \text{and} \quad \pi^L = \frac{\xi}{1 - \xi} \frac{1 - \bar{\mu}}{\bar{\mu}},$$

which is purchased by all types $\theta \in [0, 1]$. This test is sender-optimal at belief ξ . A menu that maximizes the receiver's expected payoff consists of a single test such that $\pi^H = 1$ and $\pi^L = 0$, which is purchased by all types $\theta \in [0, 1]$.

The certifier's and sender's total surplus is maximized by a menu that consists of a single Bayesian persuasion test. Intuitively, this menu maximizes their joint surplus because it maximizes the probability that the receiver takes the high action. This result also implies that the certifier could earn the highest revenue if she could sell this test to each sender type at a price equal to the sender's expected payoff. However, different sender types have different willingness to pay and the certifier cannot condition prices on private types; hence, full surplus extraction is infeasible in this sense. The revenue-maximizing certifier sells heterogeneous tests to a smaller set of types in order to screen private beliefs and charge a higher price.

A second natural benchmark is the receiver-optimal menu. This case is straightforward: because the receiver benefits from having as much information as possible, it is optimal from her perspective to sell a fully informative test to every type. Proposition 2 verifies this intuition.

The receiver's posterior belief distribution under the revenue-maximizing menu is a mean-preserving contraction of the posterior distributions under each of the two benchmarks.¹⁶ In this sense, the revenue-maximizing menu provides less information to the receiver than the two benchmarks, as the certifier seeks to limit the information rents of the sender.

3.1 Proof Sketch

The proof of Theorem 1 proceeds through a sequence of lemmas, some of which are technical, in part because we must handle partitions with a continuum of elements and some knife-edge cases. Accordingly, we provide a rough proof sketch here.

The first step is key, which shows that there exists an optimal partition that is monotone. To illustrate an idea, consider any partition $\mathcal{P}$ and any associated menu

¹⁶ Under the revenue-maximizing menu, the support of the receiver's posterior belief distribution is a subset of $[0, \bar{\mu}]$. Under the menu that maximizes the joint surplus of the certifier and the sender, the support is $\{0, \bar{\mu}\}$. Under the menu that maximizes the receiver's expected payoff, the support is $\{0, 1\}$.

(π^ω) $_{\omega \in \Omega}$. If it is not monotone, there exists an interval P^* that contains multiple elements of $\mathcal{P}$ whose convex hulls overlap. Hence, feasibility (F) and incentive compatibility (IC) imply that the differential informativeness $x = \pi^H - \pi^L$ has to be constant on P^* .

We then argue that the obedience constraint (OB) binds for every element $P \in \mathcal{P}$ contained in the interval P^* . The logic is as follows. If $\pi^H(\theta) < 1$ for some $\theta \in P^*$, we can increase both $\pi^H(\theta)$ and $\pi^L(\theta)$ by a constant, without changing $x(\theta)$. This does not violate (IC) and increases revenue (R). Hence, by optimality, (OB) must bind for the element that contains θ . If instead $\pi^H(\theta) = 1$, then unless (OB) binds, we can force some of the lowest types in P^* not currently choosing θ 's test to choose that test. Doing so weakly increases the conditional mean in every other interval and therefore does not violate (OB) for any element in P^* . Moreover, since $\pi^H(\theta) = 1$ implies that $\pi^L(\theta)$ is the highest among the tests offered to the types in P^* , this increases revenue.

The discussion thus far shows that if $\mathcal{P}$ is not monotone, then $x(\theta)$ has to be a constant x across overlapping elements and the obedience constraint for every such element binds. These observations yield an expression for the revenue generated from the interval P^* when $x > 0$:

$$x \cdot \left\{ \int_{\theta \in P^*} \psi(\theta) dF(\theta) + \mathbb{E}_{\mu \sim G} \left[\frac{\mu(1 - \bar{\mu})}{\bar{\mu} - \mu} \right] \cdot F(P^*) \right\},$$

where G is the distribution of interim beliefs $\mu_P = \mathbb{E}[\theta | \theta \in P]$ conditional on the set of elements $P \subset P^*$. The first term in the bracket does not depend on the shape of the partition, whereas the second term depends on how we partition the interval P^* .

The expression $\mu/(\bar{\mu} - \mu)$ inside the expectation operator is convex, because the observation that (OB) binds for every element $P \subset P^*$ suggests that the support of G is bounded above strictly by $\bar{\mu}$. Therefore, the certifier prefers a more dispersed distribution of interim beliefs. This is interpreted as follows: since constant differential informativeness fixes the sender's information rent, the certifier's problem

reduces to maximizing the sender's value. Given the constraints on tests, more dispersed interim beliefs imply more dispersed receiver posteriors, which increases the sender's value when those beliefs do not exceed $\bar{\mu}$.

We then show that a monotone partition of P^* achieves greater dispersion by fully separating bottom types and pooling the remaining top types so as to keep the induced interim belief strictly bounded by $\bar{\mu}$.¹⁷ Hence, the optimal partition is monotone.

Therefore, it is without loss of generality to focus on monotone partitions. We then begin with two small but important observations. First, for the highest element of a monotone partition $P \in \mathcal{P}$, we have $\pi^H(\theta) = 1$ for every $\theta \in P$. This is because we can show that $\mathbb{E}[\psi(\theta)|\theta \geq \theta^*]$ is non-negative for any θ^* , and therefore increasing $\pi^H(\theta)$ for the highest element up to the limit weakly increases revenue (R). Second, if $x(\theta)$ is constant over an interval $[\theta^*, 1]$, then there must exist an element of the partition $P \subset [\theta^*, 1]$ whose obedience constraint binds; otherwise, we can increase $\pi^L(\theta)$ for every type in the interval and raise revenue.¹⁸

The second step proves the bang-bang nature of differential informativeness. To do so, we consider an interval $[\theta_*, \theta^*]$ over which $x(\theta) \in (0, x(1))$. We first show that $\pi^H(\theta) < 1$ for every $\theta \in [\theta_*, \theta^*]$; otherwise, given that obedience binds within the top interval $[\theta^*, 1]$, having $x(\theta) < x(1)$ implies a violation of obedience constraint for the test offered to type θ . This in turn implies that (OB) must bind for every element $P \subset [\theta_*, \theta^*]$, because otherwise increasing $\pi^H(\theta)$ and $\pi^L(\theta)$ by the same amount increases revenue. With (OB) binding throughout, we can embed it into the revenue and reduce the problem to a linear program with a monotonicity constraint and other minor constraints such as feasibility. Then, a maximizer is an extreme point, which is a step function: $x(\theta) \in \{0, x(1)\}$ for all $\theta \in [\theta_*, \theta^*]$.

¹⁷ The induced conditional distribution takes the form of *upper censorship* as in Kolotilin et al. (2022). They show in general that upper censorship is optimal when the objective function is S-shaped. Here, though, the technical reason for pooling the top types is that it is optimal to have the obedience constraint binding for all tests, and this can only happen (with $x > 0$) if the interim beliefs fall below $\bar{\mu}$.

¹⁸ A slightly more involved argument is needed when the interval contains infinitely many elements and the obedience constraint binds only asymptotically in the limit; we treat this case in the Appendix.

In the third step, we show that the obedience binds for every element of a monotone partition. The idea is similar to the discussion above. If $\pi^H(\theta) < 1$, we can increase both $\pi^H(\theta)$ and $\pi^L(\theta)$ by the same amount unless the obedience binds, which increases revenue. If $\pi^H(\theta) = 1$ and obedience were slack, the certifier could profitably invite additional types to choose the corresponding (most expensive) test, as in the first step of the proof.

The discussion thus far shows that, in the optimal menu, the associated partition is monotone, no test is sold to types on $[0, \theta_*)$, and on $[\theta_*, 1]$ the differential informativeness is constant and the obedience constraint binds for every element of the partition. Hence, on $[\theta_*, 1]$ we again obtain the same expression for revenue as in the first step of this proof sketch, which implies that the optimal partition fully separates intermediate types $[\theta_*, \theta^*)$. The shape of the objective function further implies that we should exclude bottom types, as in standard auction problems, and hence $\theta_* > 0$. Finally, the top types must be pooled so that the obedience constraint binds, which in particular requires $\theta^* < \bar{\mu}$. These observations lead to the menu in Theorem 1.

4. DISCUSSION: UNOBSERVABLE TEST CHOICE

In the main analysis, we assumed the sender's test choice to be observable to the receiver. While this restriction mostly aligns with our leading applications, some details about the sender's selection might sometimes be concealed. For example, in contexts where a test can be taken multiple times, information about the number of attempts is not always disclosed to the receiver. Motivated by this observation, we now show that the certifier can benefit from hiding the sender's test choice and designing the menu accordingly.

In this model variation, it is without loss to focus on menus in which signals are action recommendations that the receiver wants to obey, containing no further information about the test choice.¹⁹ The constraint that the certifier faces when

¹⁹ If there were several signals containing arbitrary additional information about the sender's selection but leading to the same receiver's action, these could be combined into one action recommendation without affecting the receiver's willingness to choose the same action.

designing a menu that assigns tests (π^H, π^L) to different types is then the following aggregate obedience constraint:

$$\frac{\mathbb{E}[\theta(1 - \pi^H(\theta))]}{\mathbb{E}[\theta(1 - \pi^H(\theta)) + (1 - \theta)(1 - \pi^L(\theta))]} \leq \bar{\mu} \leq \frac{\mathbb{E}[\theta\pi^H(\theta)]}{\mathbb{E}[\theta\pi^H(\theta) + (1 - \theta)\pi^L(\theta)]}. \quad (\text{AOB})$$

The certifier solves the following problem. It is analogous to Lemma 1 in Section 2, but now has the aggregate obedience constraint in place of the obedience constraint for each partition element.

LEMMA 2. Suppose that the test choice is not observable to the receiver. The maximum revenue is equal to the value of the following problem:

$$\max_{\theta_0 \in [0,1], \pi^\omega: [0,1] \rightarrow [0,1]} \int_0^{\theta_0} x(\theta) d\theta + \int_0^1 (\psi(\theta)x(\theta) + \pi^L(\theta)) dF(\theta),$$

subject to (IC), (AOB), and (IR), where $x(\theta) = \pi^H(\theta) - \pi^L(\theta)$.

Note that any menu that satisfies (F) and (OB) also satisfies (AOB). Therefore, when the tests are not observed by the receiver, the maximum expected revenue is weakly higher compared to the original model wherein the tests are observed by the receiver.

We find that when virtual value ψ is concave, revenue-maximizing certification design indeed depends on the observability of the test selection. In this case, when test choice can be hidden, it becomes optimal to sell two tests: (i) the highest sender types purchase a fully informative test that recommends the high action only in the high state, and (ii) intermediate types purchase a test that guarantees a high-action recommendation in both states.

PROPOSITION 3. Suppose the sender's test choice is not observed by the receiver. If $\psi(\theta)$ is concave in θ , there exist $\underline{\theta}$ and $\bar{\theta}$ with $\underline{\theta} < \bar{\theta}$ such that

$$(\pi^H(\theta), \pi^L(\theta)) = \begin{cases} (1, 0) & \text{if } \theta \geq \bar{\theta}, \\ (1, 1) & \text{if } \bar{\theta} \geq \theta \geq \underline{\theta}, \\ (0, 0) & \text{if } \underline{\theta} \geq \theta \end{cases}$$

maximizes (R) subject to (IC), (AOB), and (IR). If $\psi(\theta)$ is strictly concave in θ , any optimal menu takes this form, differing by a measure-zero set.

We derive the solution by considering a relaxed problem without the incentive compatibility constraint (IC). However, the aggregate obedience constraint (AOB), which relates tests between different types, remains in place. In the proof, we adopt a general Lagrangian approach in Luenberger (1969) and construct an allocation that maximizes the Lagrangian pointwise.

The concavity assumption ensures that a Lagrange multiplier consistent with (AOB) can be found and that the pointwise maximizer satisfies (IC). Examples of distributions on the unit interval with a concave virtual value $\psi(\theta)$ include distributions with a linearly increasing density function as well as power distributions $F(\theta) = \theta^a$ for $\theta \in [0, 1]$ with parameter $a \geq 1$. Note that a limiting case of both of these classes of distributions is the uniform distribution, where $\psi(\theta)$ is linear.

The most expensive test in the menu of Proposition 3 is the one that guarantees a high-action recommendation in both states and is purchased by the intermediate types. These types choose it instead of the cheaper, fully informative test because they are not confident enough about the state being high. The aggregate obedience constraint is binding: upon receiving a high-action recommendation, the receiver is indifferent whether to obey. If she knew that the recommendation came from the test that always gives such a recommendation, she would not want to obey; if she knew that it came from the fully informative test, she would strictly prefer the high action. Figure 2 illustrates the optimal menu.

We can also regard the certifier as selling a *grade*, namely a recommendation that induces the high action, whose informational content is determined endogenously by equilibrium self-selection across menu options. Under the optimal menu in the proposition, the certifier sells a guaranteed high grade to middle types. To keep this grade credible to the receiver, the certifier relies on top types' demand for the test (1, 0) that gives the high grade only when the state is high.

An additional motivation for the model variation considered in this section is that even though the receiver might observe the test choice of the sender, she might face restrictions on the use of that information. For instance, consider a defendant

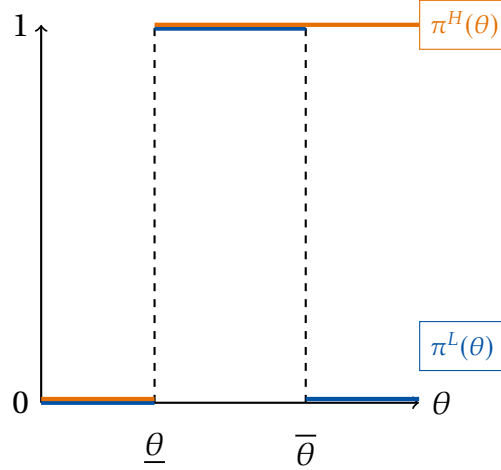


FIGURE 2. Menu in Proposition 3.

with private information about her guilt who hires an attorney to convince a jury of her innocence. The defendant's preference between an expensive service that might get a guilty defendant released and a less expensive service that is unlikely to achieve this could reasonably depend on her private information. However, the jury may not be allowed to consider such information when assessing guilt.

In two cases, the certifier could be indifferent whether the test choice is observable. First, the certifier is indifferent when the revenue-maximizing menu of the observed-choice model described in Theorem 1 coincides with the revenue-maximizing menu of the unobserved-choice model described in Proposition 3. However, this can happen only in the special case where the menu consists of a single test that always recommends the high action. Proposition 1 gives a sufficient condition under which such a menu is not optimal in the original model and hence the menus do not coincide. Second, the certifier might be indifferent when the menu described in Proposition 3 is not uniquely optimal, which can happen when the virtual value ψ is not strictly concave. For instance, with uniformly distributed θ and $\bar{\mu} > 2/3$, there are multiple optimal menus under unobservable test choice, one of which coincides with the optimal menu under observable test choice, and so the certifier does not strictly benefit from unobservability.

Outside of these cases, the certifier strictly benefits from unobservability of the sender's test choice. Methodologically, this means that one cannot solve the original problem by relaxing the obedience constraints (OB) to a single aggregate obedience condition (AOB) to remove the dependence of the constraints on the partition; that relaxation can yield a menu that is infeasible when test choice is observable.

Finally, it is worth noting that [Chopra and Ely \(2025\)](#) study a related model in which the receiver does not observe the sender's test choice, as we do here. In their college admissions application, they examine the same payoff specification as ours and, using an independent technique, obtain the same qualitative characterization of the optimal menu as our Proposition 3, under a slightly different assumption on the distribution of sender types.

5. CONCLUSION

Revenue-maximizing certification design is shaped by the fact that the receiver gains information through two channels: the intrinsic informativeness of the tests and the sender's test choice, which may reveal part of her private information. Consequently, the model combines elements of mechanism design, Bayesian persuasion, and signaling. We tackle this issue by decomposing the optimization problem into two steps, holding the signaling channel fixed in the second-stage optimization.

The interaction between the sender's incentive compatibility constraint and the receiver's obedience constraint results in the optimality of a menu with a monotone cutoff structure. The sender purchases a test if her belief exceeds a lower threshold. All sender types above an upper threshold purchase a common expensive test, whereas intermediate types choose fully separating tests. Every test in the menu has the same differential informativeness, which implies that those who choose a test are indifferent among all options in the menu. The receiver's obedience constraint for taking the high action binds conditional on any test being chosen, implying that the optimal menu minimizes the receiver's surplus.

The model is stylized compared to real-world situations that often involve many receivers, multiple actions, and senders who are heterogeneous in several dimensions. The advantage of our simple model is that it allows us to precisely characterize the optimal menu and highlight forces that are likely present in more complex settings. Yet, these limitations present possible avenues for future research on the topic.

APPENDIX A: THE OUTSIDE OPTION

In the model section, we assume that the value of the outside option is zero. However, in certain equilibria, the expected mean of the sender types who forgo test may exceed the threshold $\bar{\mu}$, in which case the sender's expected payoff from not purchasing any test is positive.

We demonstrate that this assumption entails no loss of optimality. Suppose that when the sender chooses not to purchase any test, this decision is observed by the receiver, who then updates her belief about the state and chooses an action accordingly. The value of the outside option is determined by which action the receiver will take.

To avoid confusion, we return to the basic setup and prove the lemma without recourse to a set of observations derived in subsection 2.1.

LEMMA 3. In any certifier-optimal equilibrium, the sender's expected payoff from not purchasing any test is zero.

PROOF. Consider an arbitrary menu $\mathcal{M} = \{(E_\lambda, t_\lambda)_{\lambda \in \Lambda}\}$ with some space Λ , along with any equilibrium associated with it. Let p denote the probability that the receiver takes the high action conditional on no test being chosen. Note that the sender's expected payoff from not purchasing any test is p , which is independent of the sender's type.

Suppose on the contrary that the expected payoff from not purchasing any test is strictly positive, that is, $p > 0$. Consider an alternative menu

$$\mathcal{M}^* = \{(E_\lambda, t_\lambda + p)_{\lambda \in \Lambda}\} \cup \{(E_\emptyset, p)\},$$

where $E_\emptyset = (\{\emptyset\}, \pi_\emptyset)$ with $\pi_\emptyset : \Omega \rightarrow \{\emptyset\}$ is a fully uninformative test. Under this menu, consider the following strategies: If a sender of type θ did not purchase any experiment in the original menu $\mathcal{M}$, we assign θ to select $(E_\emptyset, p)$ from the menu $\mathcal{M}^*$. If type θ selected some (E_λ, t_λ) in the original menu, we assign θ to choose $(E_\lambda, t_\lambda + p)$ in the new menu. The receiver's belief and strategy, conditional on observing an experiment E_λ from the original menu, are defined in a similar manner. If the receiver observes $E_\emptyset$, she takes the high action with probability p . Finally, if no test is chosen, the receiver is assigned a belief of 0 and never takes the high action.

Note that under the original menu, every sender type must receive an expected payoff that is at least p . Thus, under the alternative menu, no sender type has an incentive to deviate by choosing no test, as this would result in a payoff of zero. It is also straightforward to verify that no sender type has an incentive to purchase a different test from the menu. Moreover, since every sender type buys some test, the receiver's strategy is also optimal. Consequently, the alternative menu admits an equilibrium that generates an expected revenue higher than that of the original menu, leading to a contradiction. $\square$

In the proof, although we set the off-path belief following no test purchase to be degenerate at 0, the argument remains valid as long as the expected type under the off-path belief is below the threshold $\bar{\mu}$ so that the receiver takes the low action. In particular, this implies that our equilibrium outcome survives the D1 criterion of Banks and Sobel (1987). The logic is as follows. As Lemma 5 and the envelope condition (EV) imply, the sender's expected payoff $V(\theta)$ is non-decreasing in type in equilibrium. Therefore, the set of receiver actions following the off-path observation such that sender type θ would prefer to deviate is decreasing in θ under set inclusion. The D1 criterion thus requires the off-path belief to have support over $\arg \min_\theta V(\theta)$, which takes a form $[0, \theta_*]$ for some threshold θ_* . In particular, the belief equal to the prior conditional on $[0, \theta_*]$ satisfies the D1 criterion. By our assumption $\mathbb{E}[\theta] < \bar{\mu}$, we have $\mathbb{E}[\theta | \theta \leq \theta_*] < \bar{\mu}$ under such a off-path belief, and hence the receiver is still willing to take the low action on this off-path history.

APPENDIX B: MIXED-STRATEGY EQUILIBRIA

In the main analysis, we focus on equilibria in which the sender uses a pure strategy. We now show that, for any finite menu, any revenue generated by an equilibrium with sender randomization can also be generated by a pure-strategy equilibrium of the same menu.

LEMMA 4. Fix any finite menu and any associated equilibrium, possibly with sender randomization. There exists a pure-strategy equilibrium of the same menu that generates the same expected revenue.

PROOF. Let S be the set of options chosen by the sender with positive probability. Note that all options in S that share a differential informativeness must share an effective price $\hat{t} = t - \pi^L$, and every type is indifferent among them.

Partition S into finitely many classes $S_1, \dots, S_K$ by differential informativeness, with $x_1 < \dots < x_K$. Sender incentive compatibility yields cutoffs

$$0 = \theta_0 \leq \theta_1 \leq \dots \leq \theta_K = 1$$

such that for $\theta \in (\theta_{k-1}, \theta_k)$, only options in S_k are optimal and options outside S_k are strictly suboptimal, so any randomization by these types occurs only within S_k .

On each interval (θ_{k-1}, θ_k) , any randomized assignment of types to the finitely many options in S_k can be replaced by a deterministic assignment that preserves, for each option, both the mass and the mean type of the types choosing it, by Theorem 2.1 of [Dvoretzky et al. \(1951\)](#), applied with the two measures $dF(\theta)$ and $\theta dF(\theta)$. Boundary types $\{\theta_1, \dots, \theta_{K-1}\}$ have measure zero and can be assigned to any optimal option. The receiver's obedience, the sender's incentive compatibility, and the certifier's revenue are preserved. $\square$

Theorem 2.1 of [Dvoretzky et al. \(1951\)](#) applies to finite action sets, so the argument does not directly extend to arbitrary menus. We conjecture, however, that equilibria with sender randomization do not give the certifier a strictly greater revenue than pure-strategy equilibria even when infinite menus of tests are allowed.

APPENDIX C: PROOFS

C.1 Proof of Theorem 1

Since the proof is long, we provide a roadmap here: The first three lemmas, Lemmas 5, 6, and 7, are preliminary observations that rule out certain extreme cases. Lemma 8 is key, as it shows that a monotone partition is optimal. Lemmas 9 and 10 address top elements and establish that we can assume there exists at most one element in the partition whose conditional mean exceeds the threshold $\bar{\mu}$, and this element is assigned a test that always induces the high action in the high state. A second key result, Lemmas 11 and 12, establishes the bang-bang nature of differential informativeness. Then, Lemmas 13 and 14 show that the obedience constraint binds for every test. Finally, Lemma 15 finds the optimal shape of the partition and completes the proof.

Throughout the proof, we solve a relaxed problem. First, we exclude the first, negative term $-V(0)$ from the objective function. The solution will eventually satisfy $V(0) = 0$, and therefore, this assumption loses no generality. Second, we are also agnostic about the obedience constraint for taking low action. Again, it turns out to be without loss of generality.

In summary, the problem we solve in the proof is described as follows. We choose a partition $\mathcal{P}$ and a menu $(\pi^\omega)_{\omega \in \{L, H\}}$ that maximize

$$\begin{aligned} R &= \int_0^1 (\psi(\theta)x(\theta) + \pi^L(\theta))dF(\theta) \\ &= \int_0^1 (\psi(\theta)\pi^H(\theta) + (1 - \psi(\theta))\pi^L(\theta))dF(\theta), \end{aligned} \tag{R}$$

subject to the constraints that

$$x(\theta) \text{ is non-decreasing in } \theta, \tag{IC}$$

$$\pi^\omega(\theta) \in [0, 1], \tag{D}$$

$$\bar{\mu} \cdot (\mu_P \pi^H(\theta) + (1 - \mu_P)\pi^L(\theta)) \leq \mu_P \pi^H(\theta), \tag{OB-H}$$

$$\pi^\omega(\theta) \text{ is constant on } P, \tag{F}$$

for each $\theta \in P \in \mathcal{P}$ and $\omega \in \{L, H\}$, where $x(\theta) = \pi^H(\theta) - \pi^L(\theta)$ and $\mu_P = \mathbb{E}[\theta | \theta \in P]$. Note that revenue R is increasing in $\pi^L(\theta)$ for a fixed $x(\theta)$. Moreover, since $\psi(\theta) \leq \theta \leq 1$, revenue is also increasing in $\pi^H(\theta)$ for a fixed $\pi^L(\theta)$.

From now on, we simplify our terminology and call (OB-H) the obedience constraint. The obedience constraint *binds* if (OB-H) holds with equality. We say that a menu is *feasible* under partition $\mathcal{P}$ if it satisfies all these four constraints under $\mathcal{P}$. To simplify notation, for each menu $(\pi^\omega)_{\omega \in \{L, H\}}$ that is feasible under a partition $\mathcal{P}$, we write $\pi_P^\omega = \pi^\omega(\theta)$ for $\omega \in \{L, H\}$, $\theta \in P$, and $P \in \mathcal{P}$.

To further simplify the problem, we provide preliminary lemmas. By this lemma, we will focus on menus $(\pi^\omega)_{\omega \in \{L, H\}}$ under which $x(\theta) \in [0, 1]$ for all θ for the rest of the proof.

LEMMA 5. For any feasible $(\pi^\omega)_{\omega \in \{L, H\}}$, we have $\pi^H(\theta) - \pi^L(\theta) \geq 0$ for all θ .

PROOF. Consider any $\theta \in [0, 1]$. Suppose that $\pi^H(\theta) < \pi^L(\theta)$. Then, the obedience constraint (OB-H) implies that

$$\bar{\mu} \leq \frac{\mu_P \pi^H(\theta)}{\mu_P \pi^H(\theta) + (1 - \mu_P) \pi^L(\theta)} < \frac{\mu_P \pi^H(\theta)}{\mu_P \pi^H(\theta) + (1 - \mu_P) \pi^H(\theta)} = \mu_P,$$

Now, suppose that there exists $\theta \in [0, 1]$ such that $x(\theta) < 0$. Let $A \subset [0, 1]$ be the interval under which $x(\theta)$ takes the minimal value, which is well-defined as (IC) implies $\min x(\theta) = x(0)$. Summing up the above inequalities for elements in A with weight f , we obtain $\bar{\mu} < \mu_A$. However, since A is an interval including the left-end point 0, $\mu_A \leq \mathbb{E}[\theta]$ must hold. Therefore, we obtain $\bar{\mu} < \mathbb{E}[\theta]$, which is a contradiction to our assumption $\bar{\mu} \geq \mathbb{E}[\theta]$. $\square$

In the proofs of the remaining lemmas, we often pool elements of a given partition and construct a coarser partition. We briefly note here that this operation preserves measurability. Indeed, recall that the conditional mean $m(\bar{\pi}) = \mathbb{E}[\theta | \pi(\theta) = \bar{\pi}]$ admits a measurable version, and hence the function $g(\theta) = m(\pi(\theta))$ is measurable. The pooled sets we consider below will always be of the form $\bigcup \{P \in \mathcal{P} \mid \mu_P \in B\}$ for some Borel set B . Since we take $\mathcal{P}$ as the level sets of the test function π , this

set can be written as $g^{-1}(B)$ and is therefore measurable. Thus each element of the coarser partition admits a well-defined conditional mean.

The next lemma will be used to address the minor case in which a partition $\mathcal{P}$ contains multiple elements P whose conditional means exceed the threshold, i.e., $\mu_P \geq \bar{\mu}$. At a high level, it shows that, without loss of generality, we can pool these elements into a single element.

LEMMA 6. Take any feasible menu $(\pi^\omega)_{\omega \in \{L, H\}}$ under a partition $\mathcal{P}$. Suppose there exists $P^* \in \mathcal{P}^*$ such that $x(\theta) = \pi^H(\theta) - \pi^L(\theta) > 0$ is constant on P^* , and that there exists an element $P \in \mathcal{P}$ with $P \subset P^*$ and $\mu_P \geq \bar{\mu}$. Then there exists another menu that is feasible under the partition

$$\bar{\mathcal{P}} = \{P \in \mathcal{P} \mid P \cap P^* = \emptyset\} \cup \tilde{\mathcal{P}}$$

and yields weakly higher revenue, where either $\tilde{\mathcal{P}} = \{P^*\}$ or $\mu_Q < \bar{\mu}$ for almost every $Q \in \tilde{\mathcal{P}}$.²⁰

PROOF. Let $P \subset P^*$ be the union of all elements $Q \subset P^*$ such that $\mu_Q \geq \bar{\mu}$. If P^* is a singleton, then we have $P = P^*$ by assumption, and the statement holds with $\tilde{\mathcal{P}} = \{P^*\}$. Therefore, assume P^* is a non-singleton interval, which has a positive measure. The statement is also obvious if P^* and P differ by only a zero-measure set, in which case $\mu_{P^*} = \mu_P \geq \bar{\mu}$: In this case, the obedience constraint never binds even under an alternative menu such that, keeping all else equal, we let types $\theta \in P^*$ choose a test $(1, 1 - x(\theta))$. Since $x(\theta)$ is constant on P^* and $\pi^H(\theta) \leq 1$, this menu achieves a weakly higher revenue and is feasible under a partition $\bar{\mathcal{P}}$ with $\tilde{\mathcal{P}} = \{P^*\}$, completing the proof. Hence, assume that there exists a union of other elements $Q^* \subset P^*$ which has a positive measure.

Let $x = x(\theta)$ for any $\theta \in P^*$ (well-defined since $x(\theta)$ is constant on P^*). Consider any element $P_1 \subset P$. Note that increasing $\pi_{P_1}^L$ and $\pi_{P_1}^H$ by a common amount increases revenue. Moreover, $\mu_{P_1} \geq \bar{\mu}$ implies that (OB-H) for element P_1 never binds, and therefore, we can assume without loss of optimality that $\pi_{P_1}^H = 1$. Since

²⁰ When we say that a property holds for almost every element of a partition of a set S , we mean that the union U of the elements satisfying this property has full measure in S , i.e., $F(S \setminus U) = 0$.

1 $x(\theta)$ is constant on P , this means that every element in P chooses the same test 1
 2 $(1, 1 - x)$ and we can pool them. Therefore, assume $P \in \mathcal{P}$ and $\pi_P^H = 1, \pi_P^L = 1 - x$. 2
 3 If $F(P) = 0$, then we have $\mu_Q < \bar{\mu}$ for almost all $Q \in \mathcal{P}$ such that $Q \subset P^*$, which 3
 4 completes the proof. Therefore, assume $F(P) > 0$. 4

5 Take a small, positive-measure subset $\Delta \subset Q^*$ such that any point in Δ is smaller 5
 6 than any point in $Q^* \setminus \Delta$. Then, consider an alternative menu such that, keeping all 6
 7 else equal, we let types in Δ choose the same test as element P . Again, it is trivial 7
 8 to see that the resulting menu satisfies all constraints other than (OB-H) under the 8
 9 resulting partition. Given that $\pi_P^H = 1, \pi_P^L = 1 - x < 1$ and $\mu_P \geq \bar{\mu}$, the original menu 9
 10 satisfies (OB-H) for element P with strict inequality. Therefore, if Δ is sufficiently 10
 11 small, the change of μ_P is also small, hence (OB-H) for element $P \cup \Delta$. Moreover, 11
 12 (OB-H) for any element in $Q^* \setminus \Delta$ is also kept satisfied by the construction of Δ . 12

13 Finally, given that $x(\theta)$ is constant on P^* and $\pi_P^L \geq \pi_Q^L$ for all $Q \subset Q^*$, the resulting 13
 14 menu achieves a weakly higher revenue. We can do this procedure as long as 14
 15 $\mu_P \geq \bar{\mu}$. This completes the proof. $\square$ 15

16 We need another preliminary lemma which deals with an extreme case. 16
 17

18 LEMMA 7. Take any feasible menu $(\pi^\omega)_{\omega \in \{L, H\}}$ under partition $\mathcal{P}$. Suppose there 18
 19 exists $P^* \in \mathcal{P}^*$ such that $x(\theta) = \pi^H(\theta) - \pi^L(\theta) = 0$ for all $\theta \in P^*$. Then, there exists 19
 20 another menu that is feasible under partition $\bar{\mathcal{P}} = \{P \in \mathcal{P} | P \cap P^* = \emptyset\} \cup \tilde{\mathcal{P}}$ and 20
 21 achieves a weakly higher revenue, such that $\tilde{\mathcal{P}}$ is monotone. 21
 22

23 PROOF. Let P_1 be the union of all elements $P \subset P^*$ such that $\mu_P < \bar{\mu}$. Let $P_2 = P^* \setminus P_1$. 23
 24 Then, (OB-H) implies that $\pi^H(\theta) = \pi^L(\theta) = 0$ for all $\theta \in P_1$. This means that all 24
 25 types in P_1 choose the same test. Hence, assume without loss of generality that 25
 26 $P_1 \in \mathcal{P}$. Moreover, (OB-H) never binds for elements $P \subset P_2$. Therefore, we can 26
 27 weakly increase revenue by setting $\pi^H(\theta) = \pi^L(\theta) = 1$ for all $\theta \in P_2$. Hence, assume 27
 28 without loss of generality that $P_2 \in \mathcal{P}$. 28

29 It remains to show that we can replace P_1 and P_2 by intervals without lowering 29
 30 revenue. Note that the revenue obtained from region P^* is given by $F(P_2)$. If 30
 31 $\mu_{P^*} \geq \bar{\mu}$, we can let all types in P^* choose the test $(1, 1)$, which weakly increases 31
 32 revenue and completes the proof. Hence, assume $\mu_{P^*} < \bar{\mu}$. 32

Now consider the problem of choosing a function $g : P^* \rightarrow [0, 1]$ that solves

$$\max_g \int_{P^*} g(\theta) dF(\theta) \text{ subject to } \int_{P^*} g(\theta)(\theta - \bar{\mu}) dF(\theta) \geq 0.$$

Then, it follows from the necessity theorem of [Luenberger \(1969\)](#) that the solution g maximizes the Lagrangian²¹

$$\mathcal{L} = \int_{P^*} g(\theta) \{1 + \lambda \theta - \lambda \bar{\mu}\} dF(\theta).$$

Clearly, we have $g(\theta) = 1$ in the solution if and only if θ is greater than some threshold θ^* . Hence, the solution g is an indicator function: $g(\theta) = \mathbb{I}\{\theta \in P\}$, where $P = \{\theta \in P^* | \theta \geq \theta^*\}$. However, the function $g(\theta) = \mathbb{I}\{\theta \in P_2\}$ corresponding to the original menu is feasible for this problem. Therefore, the solution above gives weakly higher revenue and induces a monotone partition, completing the proof. $\square$

The next lemma shows that every non-monotone menu can be replaced with a monotone menu and raise revenue. Since the proof relies on a similar technique that appears in the proofs of the other results, we defer the proof of this lemma to the end for exposition.

For two cumulative distribution functions G and H over an interval $[\underline{\theta}, \bar{\theta}]$, we say that G is a *mean-preserving contraction* of H if they have a common mean and

$$\int_{\underline{\theta}}^x G(\theta) d\theta \leq \int_{\underline{\theta}}^x H(\theta) d\theta$$

for all $x \in [\underline{\theta}, \bar{\theta}]$. In this case, it is well-known that the expectation of any convex function is larger under distribution H than under distribution G : see, for example, Section 3.A in [Shaked and Shanthikumar \(2007\)](#).

LEMMA 8. For any non-monotone menu, there exists a monotone menu that achieves a weakly higher revenue.

²¹ See Lemma 16 below for a formal statement. The objective function is a bounded linear functional of g_P , and since $\mu_{P^*} < \bar{\mu}$, the conditions of the lemma clearly hold.

PROOF. Consider any non-monotone partition $\mathcal{P}$. Then $\mathcal{P} \neq \mathcal{P}^*$, and therefore there exists an interval $P^* \in \mathcal{P}^*$ that contains a non-convex set $P \in \mathcal{P}$. Let $\{P_\lambda\}_{\lambda \in \Lambda} \subset \mathcal{P}$ be a collection of elements such that

$$\bigcup_{\lambda \in \Lambda} P_\lambda = P^*,$$

where $\Lambda \subset \mathbb{R}$ is any index space. Note that P^* has a strictly positive measure by assumption.

Take any menu $(\pi^\omega)_{\omega \in \{L, H\}}$ that is feasible under a given partition $\mathcal{P}$. Let $x(\theta) = \pi^H(\theta) - \pi^L(\theta)$. For any $\theta \in P^*$, there must exist an element $P_\theta \neq \{\theta\}$ in $\mathcal{P}$ whose convex hull includes θ ; otherwise, every element of $\mathcal{P}$ would lie entirely either to the left or to the right of θ , contradicting the minimality of $\mathcal{P}^*$. Moreover, if θ is in the interior of P^* , then $P_\theta \neq \{\theta\}$ implies that there exists an open interval $I_\theta \subset P^*$ containing θ such that x is constant on I_θ .

Now we claim that x is constant on P^* . Let $(P^*)^\circ$ be the interior of P^* . Pick any point $\theta_0 \in (P^*)^\circ$ and let $I^* \subset (P^*)^\circ$ be the set of all types $\theta \in (P^*)^\circ$ such that $x(\theta) = x(\theta_0)$. Since x is non-decreasing by (IC), I^* is an interval. By the argument above, for any type $\theta \in I^*$ there exists an open interval I_θ with $I_\theta \subset I^*$, and therefore I^* is open. However, the same logic implies that $(P^*)^\circ \setminus I^*$ is also open. Hence, we must have $I^* = (P^*)^\circ$, i.e., x is constant on $(P^*)^\circ$. Moreover, for any point $\theta \in P^*$ on the boundary of P^* (if any), the argument in the above paragraph implies that the convex hull of an element $P_\theta \in \mathcal{P}$ that includes θ has a non-empty intersection with $(P^*)^\circ$. Hence, x is constant on P^* .

Let $\pi_\lambda^\omega = \pi^\omega(\theta)$, $x_\lambda = \pi_\lambda^H - \pi_\lambda^L$, and $\mu_\lambda = \mathbb{E}[\theta | \theta \in P_\lambda]$ for each $\theta \in P_\lambda$ and $\lambda \in \Lambda$. Assume without loss of generality that π_λ^L is increasing in λ . If $\pi_\lambda^L = \pi_\eta^L$ for some $\lambda, \eta \in \Lambda$, then $x_\lambda = x_\eta$ implies that the tests sold to these elements are identical, and therefore, we can pool these elements without changing revenue. Hence, assume that π_λ^L is strictly increasing in λ . From Lemma 7, we assume that $x_\lambda > 0$ for all $\lambda \in \Lambda$. Hence, by Lemma 6, we assume without loss of optimality that $\mu_\lambda < \bar{\mu}$ for almost all λ .

Let $\lambda^* = \sup \Lambda$. Consider any index $\lambda < \lambda^*$. Then, there exists a higher index $\lambda^{**} > \lambda$ in Λ . Since $\pi_\lambda^L < \pi_{\lambda^{**}}^L$ and $x_\lambda = x_{\lambda^{**}}$, we have $\pi_\lambda^H < \pi_{\lambda^{**}}^H \leq 1$. If (OB-H) for

element P_λ does not bind, we can therefore increase π_λ^L and π_λ^H by a small, common amount without violating any constraints. Since it increases revenue, assume that (OB-H) binds, hence (X-LB) for element P_λ .

Next, suppose that $\lambda^* \in \Lambda$ and P_{λ^*} have a positive measure. Suppose that (OB-H) does not bind. Let $A = P^* \setminus P_{\lambda^*}$. Take any small subset $\Delta \subset A$ such that any point in Δ is smaller than any point in $A \setminus \Delta$. Then, consider an alternative menu which lets types $\theta \in \Delta$ choose $(\pi_{\lambda^*}^\omega)_{\omega \in \{L,H\}}$, while keeping all else equal. By construction, (OB-H) for all elements in $A \setminus \Delta$ are still satisfied. Moreover, for sufficiently small Δ , (OB-H) for element $P_{\lambda^*} \cup \Delta$ is also satisfied. Since $x_\lambda = x_{\lambda^*}$, the alternative menu is also incentive compatible. Then, $\pi_\lambda^L < \pi_{\lambda^*}^L$ implies that it achieves a higher revenue. Hence, without loss of optimality, we assume that (OB-H) for the largest index element P_{λ^*} , if it exists with some positive measure, also binds.

In summary, we have that (OB-H) binds for each element $P \subset P^*$ and x_λ is a common constant x on Λ . Again, note that we may assume $x > 0$ from Lemma 7. Therefore, we have (L-LB) for each element $P \subset P^*$. Let $G \in \Delta(P^*)$ be the CDF of $\{\mu_\lambda\}_{\lambda \in \Lambda}$ conditional on P^* . Let F_{P^*} be the prior distribution conditional on P^* . Then, G is a mean-preserving contraction of F_{P^*} .

Then, using (L-LB), the revenue obtained from the interval P^* is given by

$$x \left\{ \int_{\theta \in P^*} \psi(\theta) dF(\theta) + \mathbb{E}_{\mu \sim G} \left[\frac{\mu(1 - \bar{\mu})}{\bar{\mu} - \mu} \right] \cdot F(P^*) \right\},$$

where $x + \pi_\lambda^L \leq 1$ for all $\lambda \in \Lambda$, i.e., (X-LB). Let $h(\mu) = \mu/(\bar{\mu} - \mu)$, which is convex in $\mu < \bar{\mu}$. We construct an alternative partition over P^* . Let μ^* be the essential supremum of $\{\mu_\lambda\}_{\lambda \in \Lambda}$. If $\mu^* = \bar{\mu}$, then (X-LB) for all elements $P \subset P^*$ implies that $x = 0$, which contradicts the claim $x > 0$ in the third paragraph. Hence, assume $\mu^* < \bar{\mu}$. Define $P^{**} = P^* \cap [\theta^*, 1]$, where θ^* is such that $\mathbb{E}[\theta | \theta \in P^{**}] = \mu^*$. Then, consider

$$\bar{\mathcal{P}} = \bigcup_{\theta \in P^* \setminus P^{**}} \{\{\theta\}\} \cup \{P^{**}\}.$$

Let $\bar{G} \in \Delta(P^*)$ be the CDF of conditional means of the elements in $\bar{\mathcal{P}}$. Note that $\bar{\mathcal{P}}$ is a monotone partition.

We claim that G is a mean-preserving contraction of $\bar{G}$. Let $\bar{F}$ be the conditional distribution of F on P^* . Suppose, by seeking a contradiction,

$$\int_{\inf P^*}^{\theta} G(s)ds > \int_{\inf P^*}^{\theta} \bar{G}(s)ds, \quad (\text{N-MPC})$$

for some $\theta \in [\inf P^*, \sup P^*]$. Since G is a mean-preserving contraction of $\bar{F}$ and $\bar{G}$ coincides with $\bar{F}$ on $\mu \in [\inf P^*, \theta^*)$, we must have $\theta > \theta^*$. Moreover, if $\theta \geq \mu^*$, integration by parts shows that both sides of (N-MPC) equals $\mu^* - \mu_{P^*}$, which is a contradiction. Hence, assume $\theta \in (\theta^*, \mu^*)$. Suppose $G(\theta) \leq \bar{G}(\theta)$. Then, since G is a non-decreasing function and $\bar{G}(s) = \bar{G}(\theta^*)$ for all $s \in [\theta^*, \mu^*)$, we have $G(s) \leq \bar{G}(s)$ for all $s \in [\theta^*, \theta)$, hence

$$\begin{aligned} \int_{\inf P^*}^{\theta} G(s)ds &= \int_{\inf P^*}^{\theta^*} G(s)ds + \int_{\theta^*}^{\theta} G(s)ds \\ &\leq \int_{\inf P^*}^{\theta^*} \bar{F}(s)ds + \int_{\theta^*}^{\theta} \bar{G}(s)ds = \int_{\inf P^*}^{\theta} \bar{G}(s)ds, \end{aligned}$$

which contradicts the assumption (N-MPC). Therefore, $G(\theta) > \bar{G}(\theta)$. However, we then have $G(s) \geq G(\theta) > \bar{G}(\theta) = \bar{G}(s)$ for all $s \in [\theta, \mu^*)$. This implies

$$\begin{aligned} \int_{\inf P^*}^{\mu^*} G(s)ds &= \int_{\inf P^*}^{\theta} G(s)ds + \int_{\theta}^{\mu^*} G(s)ds \\ &> \int_{\inf P^*}^{\theta} \bar{G}(s)ds + \int_{\theta}^{\mu^*} \bar{G}(s)ds = \int_{\inf P^*}^{\mu^*} \bar{G}(s)ds, \end{aligned}$$

where the inequality also uses (N-MPC). However, since both G and $\bar{G}$ have mean μ_{P^*} by construction, integration by parts implies that both sides of this inequality equal $\mu^* - \mu_{P^*}$. This is a contradiction. Hence, (N-MPC) does not hold, and G is a mean-preserving contraction of $\bar{G}$.

Then, consider a menu under $\bar{\mathcal{P}}$ such that $x(\theta) = x$ for all θ and $\pi^L(\theta)$ is defined such that the obedience constraint of every element in $\bar{\mathcal{P}}$ bind. This menu is well-defined by construction. Then, the revenue obtained from P^* is given as in the above expression, with G being replaced by $\bar{G}$. Since both G and $\bar{G}$ have support on $[0, \bar{\mu})$ and $h(\mu)$ is convex in this region, the revenue under this alternative partition $\bar{\mathcal{P}}$ is higher.

Therefore, for any interval P^* , if the associated sub-partition on P^* is not monotone, we can replace it with a monotone partition and weakly increase revenue. Applying this procedure to every interval in $\mathcal{P}^*$, we obtain a monotone partition on the unit interval $[0, 1]$. This completes the proof. $\square$

Then, we look for the optimal menu focusing on the class of monotone menus. We start with the following observation.

LEMMA 9. For an optimal menu $(\pi^\omega)_{\omega \in \{L, H\}}$ under any monotone partition $\mathcal{P}$, we have $\pi^H(1) = 1$.

PROOF. Let P be the highest element. Then, $1 \in P$. Since $\mathcal{P}$ is monotone, P is an interval between some $\theta \leq 1$ and 1. Therefore, if $\theta \neq 1$, we have

$$\begin{aligned} \int_P \psi(s) dF(s) &= \int_P s dF(s) - \int_P [1 - F(s)] ds \\ &= \int_P s dF(s) - \left[(0 - \theta F(P)) + \int_P s dF(s) \right] = \theta F(P), \end{aligned}$$

where the second equality follows from integration by parts. Since the last expression is clearly non-negative, it is optimal to set $\pi_P^H = 1$. Note that increasing π_P^H of the highest element P never violates any constraints. If $\theta = 1$, then $\psi(1) = 1 > 0$. Therefore, without loss of optimality, we can set $\pi_P^H = 1$. $\square$

In the proof of the remaining lemmas, for each element $P_1 \subset [0, 1]$ and $P_2 \subset [0, 1]$, we write $P_2 \geq P_1$ if any point in P_2 is greater than any point in P_1 . Formally, $P_2 \geq P_1$ if $\theta_2 \geq \theta_1$ for all $\theta_1 \in P_1$ and $\theta_2 \in P_2$.

LEMMA 10. There exists an optimal menu $(\pi^\omega)_{\omega \in \{L, H\}}$ under which there exists at most one element P such that $\mu_P \geq \bar{\mu}$.

PROOF. Note that increasing π_P^H and π_P^L by a common amount increases revenue (R). Therefore, optimality implies that we can assume either $\pi_P^H = 1$ or (OB-H) for element P binds.

Let $\bar{\mathcal{P}} \subset \mathcal{P}$ be the collection of elements which have a conditional mean greater than or equal to $\bar{\mu}$. Suppose that there exists an element $P \in \bar{\mathcal{P}}$ such that $x_P < x(1)$.

Then, consider increasing π_Q^L of every element $Q \geq P$ until we get $x_Q = x_P$. Since we have $\mu_Q \geq \bar{\mu}$ by assumption, this operation never violates (OB-H). Moreover, the resulting menu satisfies (IC). Revenue is increasing in $\pi^L(\theta)$, and therefore, optimality implies that x_P is constant on $\bar{\mathcal{P}}$.

Next, suppose that $\pi_P^L > \pi_Q^L$ for some elements $P, Q \in \bar{\mathcal{P}}$. In this case, consider an alternative menu such that, keeping all else equal, we let element Q choose the test for element P . Likewise, (OB-H) is kept satisfied. Moreover, $x_P = x_Q$ by the above paragraph, and therefore, this weakly increases revenue. Hence, without loss of optimality, we can assume $\pi_P^L = \pi_Q^L$.

In summary, both x_P and π_P^L are constant on $\bar{\mathcal{P}}$, which means that they choose a common test. Therefore, we can pool these elements without violating any constraints and keep the same revenue. This completes the proof. $\square$

The following two lemmas together show that an optimal menu has a cutoff structure such that the differential informativeness $x(\theta) = \pi^H(\theta) - \pi^L(\theta)$ takes one of two values, either 0 or $x(1)$. We say that (OB-H) *approximately binds* for a collection $\{P_n\}_{n \in \mathbb{N}}$ of elements if for any $\varepsilon > 0$,

$$x_{P_n} - \frac{\bar{\mu} - \mu_{P_n}}{\bar{\mu}(1 - \mu_{P_n})} \pi_{P_n}^H < \varepsilon,$$

for some $n \in \mathbb{N}$.

LEMMA 11. For an optimal menu $(\pi^\omega)_{\omega \in \{L, H\}}$ under a monotone partition $\mathcal{P}$, if (OB-H) for some collection $\{P_n\}_{n \in \mathbb{N}}$ of elements such that $x_{P_n} = x(1)$ for all $n \in \mathbb{N}$ approximately binds, then $x(\theta) \in \{0, x(1)\}$ for all θ .

PROOF. Assume (OB-H) for $\{P_n\}_{n \in \mathbb{N}}$ approximately binds. Take any element P such that $0 < x_P < x(1)$. Let ε be a small number such that $x(1) - \varepsilon > x_P$. Note that increasing π_P^H and π_P^L by a common amount increases revenue (R). Therefore, either $\pi_P^H = 1$ or (OB-H) for element P binds.

We claim that $\pi_P^H \neq 1$. Suppose not. Note that (OB-H) can be written as

$$x_P \geq \frac{\bar{\mu} - \mu_P}{\bar{\mu}(1 - \mu_P)} \pi_P^H.$$

Since we have $\pi_P^H = 1$ by assumption, this implies that there exists an element P^* with $x_{P^*} = x(1)$ such that

$$\frac{\bar{\mu} - \mu_P}{\bar{\mu}(1 - \mu_P)} \leq x_P < x_{P^*} - \varepsilon \leq \frac{\bar{\mu} - \mu_{P^*}}{\bar{\mu}(1 - \mu_{P^*})} \pi_{P^*}^H \leq \frac{\bar{\mu} - \mu_{P^*}}{\bar{\mu}(1 - \mu_{P^*})},$$

where the equality follows by assumption. Note that the fraction $[\bar{\mu} - \mu]/[1 - \mu]$ is decreasing in $\mu \in [0, 1]$. Therefore, this is a contradiction to $\mu_{P^*} \geq \mu_P$. Hence, we have $\pi_P^H \neq 1$, and therefore, (OB-H) for element P binds, for all P such that $0 < x_P < x(1)$.

Consider any element P such that $0 < x_P < x(1)$. Then, from the above discussion, the revenue conditional on element P can be written as

$$\mathbb{E} \left[\left\{ \psi(\theta) + \frac{\mu_P(1 - \bar{\mu})}{\bar{\mu} - \mu_P} \right\} x_P \middle| \theta \in P \right] = \mathbb{E} [\phi(\theta)x(\theta) | \theta \in P],$$

where we write $\phi(\theta) = \psi(\theta) + \mu_P(1 - \bar{\mu})/(\bar{\mu} - \mu_P)$ for each $\theta \in P$. In particular, the conditional revenue is linear in x .

Now, let $A \subset [0, 1]$ be the union of all elements $P \in \mathcal{P}$ such that $0 < x_P < x(1)$. Since $\mathcal{P}$ is monotone, (IC) implies that A is an interval. The above paragraph shows that the revenue conditional on types A is therefore bounded above by

$$\max_{x \in \mathcal{X}} \int_A \phi(\theta)x(\theta)dF(\theta),$$

where $\mathcal{X}$ is a set of non-decreasing functions $g : A \rightarrow [0, x(1)]$ such that g is constant within each element $P \in \mathcal{P}$.

Here, we claim that $\mathcal{X}$ is convex and compact in L^1 -norm and the objective function is a linear and continuous functional. To see the first claim, note first that convexity is trivial. Consider any sequence $\{g_n\}_{n \in \mathbb{N}} \subset \mathcal{X}$. For any $g \in \mathcal{X}$, the total variation is uniformly bounded by 1. Therefore, Helly's selection theorem implies that there exists a subsequence that converges pointwise, and therefore in L^1 -norm by dominated convergence theorem. It is clear that the limit g belongs to $\mathcal{X}$. Hence, $\mathcal{X}$ is sequentially compact in L^1 -norm and is therefore compact. For the latter claim, since the objective function is a linear functional, Proposition 2 in Luenberger (1969) Chapter 5.2 implies that it is continuous in L^1 -norm if $\phi(\theta)f(\theta)$

is uniformly bounded by a finite number, which is true because f is continuous over the unit interval $[0, 1]$ and therefore has a finite upper bound.

Therefore, Bauer's maximum principle implies that there exists a maximizer of the linear functional that is an extreme point of $\mathcal{X}$, that is, cannot be expressed as a convex combination of two distinct elements in $\mathcal{X}$. Note that every extreme point $x^* \in \mathcal{X}$ of this set must be a step function such that $x^*(\theta) \in \{0, x(1)\}$ for all $\theta \in A$.²²

Finally, consider an alternative menu such that, keeping all else equal, we assign a test $(\pi_P^{H^*}, \pi_P^{H^*} - x^*(\theta))$ for each type $\theta \in P \subset A$, where $\pi_P^{H^*}$ is defined so that (OB-H) binds for element P . It remains to check this menu is feasible under the same partition $\mathcal{P}$. By construction, it satisfies (OB-H), (IC), and (F). Note that (D) obviously holds for elements $P \subset A$ such that $x^*(\theta) = 0$ for $\theta \in P$. Finally, consider elements $P \subset A$ such that $x^*(\theta) = x(1)$ for $\theta \in P$. Then, since (OB-H) for $\{P_n\}_{n \in \mathbb{N}}$ approximately binds, for any $\varepsilon > 0$, there exists an element P^* such that

$$\frac{\bar{\mu} - \mu_P}{\bar{\mu}(1 - \mu_P)} \cdot \pi_P^{H^*} = x(1) = x_{P^*} \leq \frac{\bar{\mu} - \mu_{P^*}}{\bar{\mu}(1 - \mu_{P^*})} \cdot \pi_{P^*}^H + \varepsilon \leq \frac{\bar{\mu} - \mu_P}{\bar{\mu}(1 - \mu_P)} + \varepsilon,$$

where the last equation follows from $\mu_{P^*} \geq \mu_P$ as discussed in the second paragraph. Since this holds for every $\varepsilon > 0$, we must have $\pi_{P^*}^{H^*} \in [0, 1]$. We also have $\pi_P^{H^*} - x^*(\theta) \geq 0$, hence (D). Therefore, optimality implies $x(\theta) \in \{0, x(1)\}$ for all θ . $\square$

LEMMA 12. For an optimal menu $(\pi^\omega)_{\omega \in \{L, H\}}$ under a monotone partition $\mathcal{P}$, we have $x(\theta) \in \{0, x(1)\}$ for all θ .

PROOF. Suppose not. Below, we construct an alternative feasible menu $(\bar{\pi}^\omega)_{\omega \in \{L, H\}}$ that achieves higher revenue. For there to be some θ such that $x(\theta) \notin \{0, x(1)\}$, clearly we must have $x(1) > 0$. Fix any small number $\bar{\varepsilon} > 0$ such that $x(1) > \bar{\varepsilon}$. Then, take any small number $\varepsilon > 0$ such that $\bar{\varepsilon} \geq \varepsilon$. Let $\bar{\mathcal{P}} \subset \mathcal{P}$ be the set of elements P such that $x(1) > x_P \geq x(1) - \varepsilon$ and let $\tilde{\mathcal{P}} = \{P \in \mathcal{P} | x_P = x(1)\}$. For each element $P \in \bar{\mathcal{P}}$, we write $\varepsilon_P = x_P - (x(1) - \varepsilon) \geq 0$. Note that $\varepsilon_P \leq \varepsilon$.

Take any $P \in \bar{\mathcal{P}} \cup \tilde{\mathcal{P}}$. Then, we reduce $x(\theta)$ up to $x(1) - \varepsilon$ in two steps, as follows. First, increase π_P^L up to $\pi_P^L + \varepsilon_P$ as long as (OB-H) for element P does not bind. Then,

²² See, e.g., Yang and Yang (2025) and the references therein.

we obtain a menu $(\hat{\pi}_P^\omega)_{\omega \in \{L,H\}}$ for element P , where $\hat{\pi}_P^H = \pi_P^H$ and $\hat{\pi}_P^L$ is defined as the unique solution to the following:

$$\hat{\pi}_P^L = \min \left\{ \pi_P^L + \varepsilon_P, \frac{\mu_P(1 - \bar{\mu})}{\bar{\mu} - \mu_P} \hat{x}_P \right\},$$

where $\hat{x}_P = \hat{\pi}_P^H - \hat{\pi}_P^L$. If $\hat{\pi}_P^L = \pi_P^L + \varepsilon_P$, we set $\bar{\pi}_P^\omega = \hat{\pi}_P^\omega$ for each $\omega \in \{L, H\}$.

Suppose $(\hat{\pi}_P^\omega)_{\omega \in \{L,H\}}$ is such that $\hat{\pi}_P^L < \pi_P^L + \varepsilon_P$ and therefore (OB-H) binds for element P . Then, we set $\bar{\pi}_P^H = x_P - \varepsilon_P + \bar{\pi}_P^L$ and

$$\bar{\pi}_P^L = \frac{\mu_P(1 - \bar{\mu})}{\bar{\mu} - \mu_P} (x_P - \varepsilon_P).$$

Namely, we decrease $\hat{x}_P = \hat{\pi}_P^H - \hat{\pi}_P^L$ up to $x_P - \varepsilon_P = x(1) - \varepsilon$ while keeping (OB-H) for element P with equality.²³ Note that the resulting menu is feasible by construction.

Now, define $h_P = \mu_P[1 - \bar{\mu}]/[\bar{\mu} - \mu_P]$. Then, for any element $P \in \bar{\mathcal{P}}$ such that (OB-H) binds under the new menu $(\bar{\pi}^\omega)_{\omega \in \{L,H\}}$, we have

$$1 \geq \bar{\pi}_P^L \geq h_P(x_P - \varepsilon_P) = h_P(x(1) - \varepsilon) \geq h_P(x(1) - \bar{\varepsilon}),$$

which implies $h_P \leq \bar{h} = 1/[x(1) - \bar{\varepsilon}]$ for all such elements P . Moreover, note that for each such element P , our construction implies

$$\bar{\pi}_P^L \geq \hat{\pi}_P^L - h_P \varepsilon_P \geq \pi_P^L - \bar{h} \varepsilon \geq \pi_P^L - (\bar{h} + 1) \varepsilon,$$

$$\bar{\pi}_P^H = x_P - \varepsilon_P + \bar{\pi}_P^L \geq x_P - \varepsilon + \pi_P^L - \bar{h} \varepsilon = \pi_P^H - (\bar{h} + 1) \varepsilon,$$

which implies that the revenue loss which comes from such an element P under the new menu is bounded above by $(\bar{h} + 1) \varepsilon F(P)$.

²³ If $\mu_P \geq \bar{\mu}$, which may be the case only when P is the top interval, the above expression for $\bar{\pi}_P^L$ implies $\bar{\pi}_P^L < 0$. In this case, the binding obedience constraint suggests $x(1) = 0$. Then, (IC) implies $x(\theta) = 0$ for all θ , completing the proof.

Let P^* be the union of all elements in $\tilde{\mathcal{P}}$. Since $1 \in P^*$, Lemma 10 implies that we can assume $F(P^*) > 0$.²⁴ Take any small $\delta > 0$ such that $\delta < B$, where

$$B = \int_{P^*} (1 - \psi(\theta)) dF(\theta) > 0.$$

Then, by the definition of $\bar{\mathcal{P}}$, there exists a sufficiently small $\varepsilon < \bar{\varepsilon}$ such that

$$(\bar{h} + 1) \cdot F(\bar{P}) < \delta,$$

where $\bar{P}$ is the union of all elements in $\bar{\mathcal{P}}$. Suppose ε is small enough such that $(\pi_P^H, \pi_P^L + \varepsilon)$ does not violate (OB-H) for any element P in $\tilde{\mathcal{P}}$. Lemma 11 guarantees the existence of such an $\varepsilon > 0$; otherwise, (OB-H) for the collection of elements $\tilde{\mathcal{P}}$ approximately binds, and Lemma 11 completes the proof. Then, compared to the original menu, the revenue gained from P^* under the alternative menu is larger by an amount εB . Moreover, the revenue loss from the other elements under the alternative menu is smaller by at most $\varepsilon \delta$, which follows by the above discussion. This is a contradiction. $\square$

Then, we claim that the obedience constraint for all but the highest element binds under optimal menus.

LEMMA 13. For an optimal menu $(\pi^\omega)_{\omega \in \{L, H\}}$ under a monotone partition $\mathcal{P}$, the obedience constraint (OB-H) binds for every $P \in \mathcal{P}$ such that $1 \notin P$.

PROOF. Fix any element P such that $1 \notin P$. As discussed in the above proofs, either $\pi_P^H = 1$ or (OB-H) for element P binds. Suppose that (OB-H) for element P does not bind. Then, $\pi_P^H = 1$ and $\pi_P^L \neq 1$.

Let Q be any element such that $Q \geq P$. By assumption, we have $\pi_P^H = 1 \geq \pi_Q^H$, and therefore, (IC) implies that $\pi_P^L \geq \pi_Q^L$. We must also have $\mu_P < \mu_Q$ by the monotone

²⁴ Otherwise, $P^* = \{1\}$, and therefore, Lemma 10 implies that there is an element adjacent to 1 that has positive measure (otherwise, there would be two elements with conditional means of at least $\bar{\mu}$, contradicting Lemma 10). Then, we can reassign type 1 to that adjacent element, which affects neither the obedience constraint nor the revenue.

structure. We then obtain

$$\bar{\mu} < \frac{\mu_P \pi_P^H}{\mu_P \pi_P^H + (1 - \mu_P) \pi_P^L} \leq \frac{\mu_Q \pi_P^H}{\mu_Q \pi_P^H + (1 - \mu_Q) \pi_P^L},$$

because $\pi_P^H \geq \pi_P^L$ implies that the fraction is increasing in beliefs. This means that we can let element Q choose the test for element P without violating (OB-H).

Now, consider an alternative menu which, keeping all else equal, lets all types $\theta \geq \sup P$ choose the test $(\pi_P^\omega)_{\omega \in \{L, H\}}$ which is originally only for element P . Note that the resulting menu does not violate (OB-H) as seen in the above paragraph. Moreover, we have (IC) by construction, hence feasible. Since Lemma 12 implies that $x(\theta)$ is a constant on $P \cup Q$ and $\pi_P^L \geq \pi_Q^L$ for all $Q \geq P$, the alternative menu achieves a higher revenue. Therefore, optimality implies that $\pi_P^L = \pi^L(\theta)$ for all $\theta > \inf P$. However, this implies that every type $\theta > \inf P$ chooses a common test, and therefore, we can pool these elements into a single element without changing revenue. Hence, we can redefine P to be an interval including 1, which completes the proof. $\square$

LEMMA 14. One of the following monotone menus maximizes expected revenue among all menus: a menu that consists of a single test, or a menu such that the conditional mean of every element is strictly less than $\bar{\mu}$. In both cases, (OB-H) binds for every element.

PROOF. Let $(\pi^\omega)_{\omega \in \{L, H\}}$ be an optimal menu. Then, it is feasible under a monotone partition $\mathcal{P}$. Lemma 10 implies that we can assume the highest element has a positive measure. Moreover, Lemma 12 implies that there exists an interval $P^* = [\theta_*, 1]$ such that $x(\theta) = \pi^H(\theta) - \pi^L(\theta)$ is a constant on P^* and zero over the remaining region. Therefore, Lemma 6 implies that either $P^* \in \mathcal{P}$ or $\mu_P < \bar{\mu}$ for almost all $P \in \mathcal{P}$ such that $P \subset P^*$.

There are two cases to consider. In the former case, there is a single element in interval P^* in an optimal menu. Suppose (OB-H) for P^* does not bind. If $x_{P^*} > 0$, and therefore $1 \geq \pi_{P^*}^H > \pi_{P^*}^L$, we can increase $\pi_{P^*}^L$ and thereby raise revenue without violating any constraints. If $x_{P^*} = 0$, then we must have $\mu_{P^*} > \bar{\mu}$. Optimality then implies that $(\pi_{P^*}^H, \pi_{P^*}^L) = (1, 1)$. However, since $\mathbb{E}[\theta] < \bar{\mu} < \mu_{P^*}$ by assumption,

there exists a positive measure of types in $[0, 1] \setminus P^*$ who do not choose this test. Assigning the test $(1, 1)$ to a subset of these types while keeping everything else unchanged preserves both incentive compatibility and the obedience constraint, while increasing revenue. This contradicts optimality. Taken together, these arguments imply that (OB-H) for P^* must bind.

Next, suppose $\mu_P < \bar{\mu}$ for almost all $P \in \mathcal{P}$ such that $P \subset P^*$. Let P be the highest element. Then, we can assume $F(P) > 0$ as discussed in the proof of Lemma 12. By Lemma 13, (OB-H) for all elements other than P binds. Now, for any small $\varepsilon > 0$, consider an alternative menu such that, keeping all else fixed, sender types $\theta \in \Delta = [\theta_*, \theta_* + \varepsilon]$ choose the test for the highest element. By Lemma 9, we have $\pi_P^H = 1$, and therefore, $x_P = x_Q$ implies $\pi_P^L > \pi_Q^L$ for all $Q \neq P$. This means that the alternative menu achieves higher revenue. Moreover, for any element $Q \neq P$, either $Q \setminus \Delta = \emptyset$ or $\mu_{Q \setminus \Delta} \geq \mu_Q$ by construction, which implies that (OB-H) for all elements are still satisfied. Since the alternative menu must not be feasible for any $\varepsilon > 0$, (OB-H) for element P must bind. $\square$

The next lemma shows that the above observations imply the optimality of the menu that we describe in Theorem 1.

LEMMA 15. Menu $\mathcal{M}(\theta_*, \theta^*)$ is optimal for some $0 < \theta_* \leq \theta^* < \bar{\mu}$.

PROOF. Take any optimal menu $(\pi^\omega)_{\omega \in \{L, H\}}$ under monotone partition $\mathcal{P}$. By Lemmas 14, 9, and 13, if there is some element such that the conditional mean is not strictly less than $\bar{\mu}$, menu $\mathcal{M}(\theta_*, \theta^*)$ is optimal for $0 < \theta_* = \theta^* < \bar{\mu}$ such that $\mathbb{E}[\theta | \theta \geq \theta^*] = \bar{\mu}$. Here, the strict inequality $\theta_* > 0$ holds by assumption $\xi < \bar{\mu}$. Therefore, assume in the following that $\mu_P < \bar{\mu}$ for all $P \in \mathcal{P}$.

Lemma 12 and (IC) imply that there exists a threshold θ_* such that $x(\theta) = 0$ for all $\theta < \theta_*$ and $x(\theta) = x$ for all $\theta > \theta_*$ for some $x \in [0, 1]$. Moreover, Lemma 14 implies that (OB-H) for every element P must bind. Therefore, for each P that is above θ_* , we get

$$\pi_P^L = \frac{\mu_P(1 - \bar{\mu})}{\bar{\mu} - \mu_P} x. \quad (\text{L-LB})$$

Moreover, since $x(\theta) + \pi^L(\theta) \leq 1$ must hold for all θ , we obtain

$$x \left\{ \frac{\mu_P(1 - \bar{\mu})}{\bar{\mu} - \mu_P} + 1 \right\} \leq 1 \iff x \leq \frac{\bar{\mu} - \mu_P}{\bar{\mu}(1 - \mu_P)}, \quad (\text{X-LB})$$

for each element P above θ_* .

Therefore, for revenue R under the menu $(\pi^\omega)_{\omega \in \{L, H\}}$, we obtain

$$\begin{aligned} R &= \int_0^1 \left\{ \psi(\theta) + \frac{\mu_{P(\theta)}(1 - \bar{\mu})}{\bar{\mu} - \mu_{P(\theta)}} \right\} x(\theta) dF(\theta) \\ &= x \int_{\theta_*}^1 \left\{ \psi(\theta) + \frac{\mu_{P(\theta)}(1 - \bar{\mu})}{\bar{\mu} - \mu_{P(\theta)}} \right\} dF(\theta) \\ &\leq \frac{\bar{\mu} - \mu_{P(1)}}{\bar{\mu}(1 - \mu_{P(1)})} \int_{\theta_*}^1 \left\{ \psi(\theta) + \frac{\mu_{P(\theta)}(1 - \bar{\mu})}{\bar{\mu} - \mu_{P(\theta)}} \right\} dF(\theta) \end{aligned}$$

where we denote by $P(\theta) \in \mathcal{P}$ the element to which θ belongs. The last inequality follows from the right-hand side of (X-LB) being decreasing in μ_P .

Write $P(1) = [\theta^*, 1]$ and $\mu^* = \mu_{P(1)}$. Recall that $\mu_P < \bar{\mu}$ for all $P \in \mathcal{P}$. Moreover, the fraction $\mu/[\bar{\mu} - \mu]$ is convex in $\mu \in [0, \bar{\mu})$. Hence, Jensen's inequality applied to $[\theta_*, \theta^*]$ further implies that revenue R is bounded above by

$$\frac{\bar{\mu} - \mu^*}{\bar{\mu}(1 - \mu^*)} \left[\int_{\theta_*}^{\theta^*} \left\{ \psi(\theta) + \frac{\theta(1 - \bar{\mu})}{\bar{\mu} - \theta} \right\} dF(\theta) + \int_{\theta^*}^1 \left\{ \psi(\theta) + \frac{\mu^*(1 - \bar{\mu})}{\bar{\mu} - \mu^*} \right\} dF(\theta) \right].$$

This is exactly the revenue under menu $\mathcal{M}(\theta_*, \theta^*)$. Finally note that if $\theta_* = 0$, then increasing θ_* results in a strictly higher revenue (when $\theta_* = 0 < \theta^*$, this is implied by $\psi(0) < 0$, and when $\theta_* = \theta^* = 0$, the derivative of the revenue with respect to the single threshold is $[(1 - \bar{\mu})\xi^2 f(0) + (1 - \xi)(\bar{\mu} - \xi)]/[\bar{\mu}(1 - \xi)^2] > 0$); hence $\theta_* > 0$.

This finishes the proof. $\square$

PROOF OF THEOREM 1. The result follows from Lemma 15. $\square$

C.2 Proof of Proposition 1

PROOF OF PROPOSITION 1. When the certifier optimally offers a single test (π^H, π^L) , i.e., the menu has a single threshold $\bar{\theta}$ such that types $\theta \geq \bar{\theta}$ purchase the test,

threshold type $\bar{\theta}$ is indifferent whether to purchase the test, and the certifier's revenue is

$$(1 - F(\bar{\theta})) \left(\bar{\theta} \pi^H + (1 - \bar{\theta}) \pi^L \right) = (1 - F(\bar{\theta})) \left(\bar{\theta} + (1 - \bar{\theta}) \frac{\mu}{1 - \mu} \frac{1 - \bar{\mu}}{\bar{\mu}} \right)$$

where

$$\mu = \mathbb{E}[\theta | \theta \geq \bar{\theta}] = \int_{\bar{\theta}}^1 \theta dF(\theta) / (1 - F(\bar{\theta})).$$

Note that $\partial \mu / \partial \bar{\theta} = f(\bar{\theta})(\mu - \bar{\theta}) / [1 - F(\bar{\theta})]$. Let $\hat{\theta}$ be such that $\mathbb{E}[\theta | \theta \geq \hat{\theta}] = \bar{\mu}$. The derivative of the revenue with respect to $\bar{\theta}$ at $\bar{\theta} = \hat{\theta}$ is

$$f(\hat{\theta}) \left(\frac{1 - \hat{\theta}}{\bar{\mu}} \frac{\bar{\mu} - \hat{\theta}}{1 - \bar{\mu}} - 1 \right).$$

This gives us the condition under which it is optimal to decrease threshold $\bar{\theta}$ from level $\hat{\theta}$. $\square$

C.3 Proof of Proposition 2

PROOF OF PROPOSITION 2. Take any menu that satisfies (IC), (F), (OB). Note that (F) and (OB) imply the aggregate obedience constraint (AOB), as discussed in Section 4. Then, the joint surplus of the certifier and the sender is given by

$$\mathbb{E}[\theta \pi^H(\theta) + (1 - \theta) \pi^L(\theta)] \leq \mathbb{E}[\theta \pi^H(\theta)] / \bar{\mu},$$

where the last inequality follows from (AOB), which holds with equality if the obedience constraint for high action holds with equality. The right-hand side is maximized by setting $\pi^H(\theta) = 1$ for all θ . Moreover, since (AOB) binds if we set

$$\pi^L(\theta) = \frac{\xi}{1 - \xi} \frac{1 - \bar{\mu}}{\bar{\mu}},$$

for all θ , we obtain the first statement. The latter statement is trivial, given that the receiver's surplus is maximized when the receiver observes the true state ω , which is the case under the fully informative test $\pi^H = 1$ and $\pi^L = 0$. $\square$

C.4 Proof of Proposition 3

We take a general Lagrangian approach of [Luenberger \(1969\)](#) to solve the optimization problem in the unobservable case. For completeness, we replicate Theorem 1 in Chapter 8.6 of [Luenberger \(1969\)](#) in the following.

LEMMA 16 ([Luenberger \(1969\)](#), Theorem 1, Section 8.6). Let X be a convex subset of a vector space and Z be a normed vector space. Let $f : X \rightarrow \mathbb{R}$ be convex and let $G : X \rightarrow Z$ be convex. Fix $b \in Z$ and assume there exists $x_0 \in X$ such that $G(x_0) < b$, and that $\inf\{f(x) : x \in X, G(x) \leq b\}$ is finite. Then

$$\inf_{x \in X, G(x) \leq b} f(x) = \max_{z^* \geq 0} \inf_{x \in X} \{f(x) + \langle G(x) - b, z^* \rangle\},$$

and the maximum is attained by some $z_0^* \geq 0$. If the infimum on the left is attained at $x_0 \in X$, then $\langle G(x_0) - b, z_0^* \rangle = 0$ and x_0 attains the minimum on the right.

Before proving Proposition 3, we remark that the aggregate obedience constraint for the high action implies that for the low action.

LEMMA 17. If $\xi \leq \bar{\mu}$, then,

$$\bar{\mu} \leq \frac{\mathbb{E}[\theta \pi^H(\theta)]}{\mathbb{E}[\theta \pi^H(\theta) + (1 - \theta) \pi^L(\theta)]}$$

implies the aggregate obedience condition (AOB).

PROOF. Seeking a contradiction, suppose that we have

$$\begin{aligned} \bar{\mu} \cdot \mathbb{E}[\theta \pi^H(\theta) + (1 - \theta) \pi^L(\theta)] &\leq \mathbb{E}[\theta \pi^H(\theta)], \\ \bar{\mu} \cdot \mathbb{E}[\theta(1 - \pi^H(\theta)) + (1 - \theta)(1 - \pi^L(\theta))] &< \mathbb{E}[\theta(1 - \pi^H(\theta))]. \end{aligned}$$

Summing up these inequalities, we obtain $\bar{\mu} < \mathbb{E}[\theta] = \xi$. $\square$

The next lemma proves that the optimal menu can be found among the menus that also satisfy the following constraint:

$$x(\theta) = \pi^H(\theta) - \pi^L(\theta) \geq 0. \quad (\text{P})$$

Tests that satisfy (P) are more plausible than tests that do not, as real-life tests typically depend on the sender providing evidence of a high state to the certifier (although this is not explicitly modeled): for instance, students demonstrate their abilities by performing well on test questions. If (P) is satisfied, the sender has incentives to provide this evidence, otherwise not.

LEMMA 18. Suppose that the test choice is not observable to the receiver. The maximum revenue is equal to the value of the following problem:

$$\max_{\pi^\omega: [0,1] \rightarrow [0,1]} \left\{ \int_0^1 (\psi(\theta)x(\theta) + \pi^L(\theta)) dF(\theta) \right\} \quad (\text{R})$$

subject to (IC), (AOB) and (P), where $x(\theta) = \pi^H(\theta) - \pi^L(\theta)$.

PROOF. We only need to show that there exists an optimal menu for Lemma 2 that satisfies $\theta_0 = 0$ and (P). Consider any menu π^H and π^L that satisfies (IC), (AOB), and (IR) for some type θ_0 . Then, consider the following menu:

$$(\hat{\pi}^H(\theta), \hat{\pi}^L(\theta)) = \begin{cases} (\pi^L(\theta), \pi^L(\theta)) & \text{if } \theta \leq \theta_0, \\ (\pi^H(\theta), \pi^L(\theta)) & \text{otherwise.} \end{cases}$$

Define $\hat{x}(\theta) = \hat{\pi}^H(\theta) - \hat{\pi}^L(\theta)$ for each θ . Then, we have $\hat{x}(\theta) = \max\{x(\theta), 0\} \geq 0$, and hence this new menu satisfies (IR) with $\theta_0 = 0$.

The menu $\hat{\pi}^\omega$ results in a revenue higher than the original menu: Let R be the original revenue and $\hat{R}$ the revenue under the menu $\hat{\pi}^\omega$. Then

$$\begin{aligned} \hat{R} - R &= - \int_0^{\theta_0} \psi(\theta)x(\theta)f(\theta)d\theta - \int_0^{\theta_0} x(\theta)d\theta \\ &= - \int_0^{\theta_0} x(\theta)[1 + \psi(\theta)f(\theta)]d\theta \\ &= - \int_0^{\theta_0} x(\theta)[\theta f(\theta) + F(\theta)]d\theta \geq 0, \end{aligned}$$

where the last inequality follows from $x(\theta) \leq 0$ below θ_0 and $\theta f(\theta) + F(\theta) \geq 0$. Moreover, since the original menu satisfies (IC) and (AOB), one can see that, together with Lemma 17, the new menu also satisfies these constraints, which follows

by the construction. In summary, there must exist an optimal menu that satisfies (IR) with $\theta_0 = 0$, completing the proof. $\square$

Here, we are ready to prove Proposition 3.

PROOF OF PROPOSITION 3. First, suppose that ψ is strictly concave. We consider a relaxed problem without (IC). Later, we will check that the solution of this relaxed problem satisfies (IC). Set

$$G(x, \pi^L) = \int_0^1 \left(\theta(\bar{\mu} - 1)x(\theta) + (\bar{\mu} - \theta)\pi^L(\theta) \right) dF(\theta).$$

From Lemma 17, condition (AOB) can be written as $G(x, \pi^L) \leq 0$.

Consider choosing $(\pi^L, x) \in B^\Theta$ and maximizing expected revenue (R) subject to the aggregate obedience condition (AOB), where

$$B = \{(\pi^L, x) \in \mathbb{R}^2 \mid 1 \geq \pi^L + x \geq 0, x \geq 0, \pi^L \geq 0\}.$$

We derive a solution using a Lagrangian approach in Luenberger (1969). Construct a Lagrangian with a Lagrange multiplier $\lambda \geq 0$:

$$\mathcal{L} = \int_0^1 \left([\psi(\theta) + \lambda\theta(1 - \bar{\mu})]x(\theta) + [1 + \lambda(\theta - \bar{\mu})]\pi^L(\theta) \right) dF(\theta).$$

Note that the objective function is a linear functional of $(x, \pi^L) \in B^\Theta$ and it is also bounded. Moreover, B^Θ is convex, and we also have $G(x, \pi^L) < 0$ when $x(\theta) = 1$ and $\pi^L(\theta) = 0$ for all θ . Therefore, Lemma 16 applies.

In particular, if there exists a Lagrange multiplier λ such that $(x, \pi^L) \in B^\Theta$ maximizes the Lagrangian $\mathcal{L}$ and satisfies (AOB) with equality, then (x, π^L) is a solution. To simplify notation, we let $g_\lambda(\theta) = \psi(\theta) + \lambda\theta(1 - \bar{\mu})$ and $h_\lambda(\theta) = 1 + \lambda(\theta - \bar{\mu})$.

Fix any $\lambda \geq 0$, which is determined later. Note that a pair (x, π^L) clearly maximizes the Lagrangian $\mathcal{L}$ if it pointwise maximizes $\mathcal{L}$ for each θ . Since $\mathcal{L}$ is a linear functional and B is a simple polytope, one can see that

$$(x(\theta), \pi^L(\theta)) = \begin{cases} (1, 0) & \text{if } g_\lambda(\theta) \geq \max\{0, h_\lambda(\theta)\}, \\ (0, 1) & \text{if } h_\lambda(\theta) \geq \max\{0, g_\lambda(\theta)\}, \\ (0, 0) & \text{if } 0 \geq \max\{g_\lambda(\theta), h_\lambda(\theta)\}, \end{cases} \quad (\text{M})$$

is a pointwise maximizer, where ties are broken arbitrarily.

Here we simplify the expression (M). Note that g_λ is a strictly concave function by assumption. Hence, the set of types θ such that $g_\lambda(\theta) \geq 0$ is an interval that contains 1.²⁵ Moreover, $h_\lambda(\theta)$ is linearly increasing, with

$$g_\lambda(1) = \psi(1) + \lambda(1 - \bar{\mu}) = 1 + \lambda(1 - \bar{\mu}) = h_\lambda(1).$$

Thus, the set of types θ such that $g_\lambda(\theta) \geq \max\{0, h_\lambda(\theta)\}$ is also an interval containing 1. We can analogously see that the set of types θ such that $0 \geq \max\{g_\lambda(\theta), h_\lambda(\theta)\}$ is an interval. Therefore,

$$(x(\theta), \pi^L(\theta)) = \begin{cases} (1, 0) & \text{if } \theta \geq \bar{\theta}(\lambda), \\ (0, 1) & \text{if } \bar{\theta}(\lambda) \geq \theta \geq \underline{\theta}(\lambda), \\ (0, 0) & \text{if } \underline{\theta}(\lambda) \geq \theta, \end{cases} \quad (\text{M2})$$

for some $\bar{\theta}(\lambda)$ and $\underline{\theta}(\lambda)$ with $\bar{\theta}(\lambda) \geq \underline{\theta}(\lambda)$. Formally, we define

$$\bar{\theta}(\lambda) = \min\{\theta \mid g_\lambda(\theta) \geq \max\{0, h_\lambda(\theta)\}\},$$

$$\underline{\theta}(\lambda) = \max\{\theta \mid 0 \geq \max\{g_\lambda(\theta), h_\lambda(\theta)\}\},$$

where we set $\underline{\theta}(\lambda) = 0$ when $\max\{g_\lambda(\theta), h_\lambda(\theta)\} > 0$ for all $\theta \in \Theta$. Note that they are both well-defined as both g_λ and h_λ are continuous. Moreover, strict concavity implies that $\bar{\theta}(\lambda)$ and $\underline{\theta}(\lambda)$ are continuous in λ .

It remains to find a Lagrange multiplier λ such that (AOB) is satisfied with equality, that is, $G(x, \pi^L) = 0$. Using (M2), we obtain

$$G(x, \pi^L) = G^*(\lambda) = -(1 - \bar{\mu}) \int_{\bar{\theta}(\lambda)}^1 \theta dF(\theta) + \int_{\underline{\theta}(\lambda)}^{\bar{\theta}(\lambda)} (\bar{\mu} - \theta) dF(\theta).$$

Note that the right-hand side is continuous in λ since both $\bar{\theta}(\lambda)$ and $\underline{\theta}(\lambda)$ are continuous functions of λ , as discussed in the earlier paragraph.

Here, we find a desired λ using the intermediate value theorem. First, note that we have $g_0(\theta) = \psi(\theta) \leq 1 = h_0(\theta)$, where the inequality is strict for any $\theta \neq 1$.

²⁵ The non-emptiness of the set follows from $g_\lambda(1) \geq \psi(1) = 1 > 0$.

Therefore, we have $\bar{\theta}(0) = 1$ and $\underline{\theta}(0) = 0$, hence

$$G^*(0) = \int_0^1 (\bar{\mu} - \theta) dF(\theta) = \bar{\mu} - \xi > 0, \quad (\text{ST})$$

where the last inequality follows by assumption. Second, note that the linear function h_λ intersects with 0 at point $\theta_\lambda = \bar{\mu} - 1/\lambda$. For sufficiently large λ ,

$$g_\lambda(\theta_\lambda) = \psi(\theta_\lambda) + (\lambda\bar{\mu} - 1)(1 - \bar{\mu}) > 0 = h_\lambda(\theta_\lambda).$$

Since $g_\lambda(1) = h_\lambda(1)$, g_λ is strictly concave, and h_λ is linear, this means that $g_\lambda(\theta) \geq h_\lambda(\theta)$ for all $\theta \geq \theta_\lambda$. Fix any such $\lambda^* > 0$. By definition, $\bar{\theta}(\lambda^*) = \underline{\theta}(\lambda^*) = \theta^*$, hence

$$G^*(\lambda^*) = -(1 - \bar{\mu}) \int_{\theta^*}^1 \theta dF(\theta) < 0. \quad (\text{EN})$$

Therefore, it follows from (ST) and (EN) that there exists λ with $0 < \lambda < \lambda^*$ such that the aggregate obedience condition holds with equality, that is, $G^*(\lambda) = 0$. Note that $\bar{\theta}(\lambda) > \underline{\theta}(\lambda)$ must hold.

In summary, there exists $\lambda \geq 0$ such that the menu π^ω induced by (M2) maximizes revenue (R) subject to (AOB). Moreover, it follows from the definition that $x(\theta)$ is increasing in θ , hence (IC). Therefore, by $\pi^H(\theta) = x(\theta) + \pi^L(\theta)$, we obtain the solution. Finally, Lemma 16 also states that any optimal solution must maximize the Lagrangian with some multiplier λ . Since ψ is strictly concave, we have $g_\lambda(\theta) = h_\lambda(\theta)$, $g_\lambda(\theta) = 0$, or $h_\lambda(\theta) = 0$ only on a measure-zero set of types θ . This means that any pointwise maximizer of the Lagrangian must take the form (M2) almost surely.

It remains to extend the result to the case where ψ is only weakly concave. Take a sequence of strictly concave functions ψ_n converging uniformly to ψ , with $\psi_n(1) = 1$ and $\psi_n(\theta) < 1$ for all $\theta < 1$: E.g., $\psi_n(\theta) = (1 - \varepsilon_n)\psi(\theta) + \varepsilon_n(2\theta - \theta^2)$ with $\varepsilon_n \rightarrow 0$. For each n , consider the auxiliary problem obtained by replacing ψ in the objective (R) by ψ_n , while keeping F and the feasible set B^Θ subject to (AOB) fixed. The preceding Lagrangian argument uses only the stated continuity, strict concavity, and normalization properties of ψ_n , so it applies even though ψ_n need not be the virtual type function associated with F . Thus, this auxiliary problem has

an optimal menu of the form (M2). The objective function, viewed as a function of (ψ, x, π^L) , is continuous. The domain of optimization is compact under weak topology. Therefore, the correspondence of optimal solutions is upper hemicontinuous in ψ by the maximum theorem. Hence, taking a convergent subsequence of the corresponding cutoff pairs $(\underline{\theta}_n, \bar{\theta}_n)_n$, the limiting cutoff menu is optimal for ψ . The aggregate obedience must bind at each n and hence in the limit, and thus $\bar{\theta} > \underline{\theta}$; if $(\underline{\theta}, \bar{\theta}) = (1, 1)$, then the revenue is zero and is not optimal. $\square$

REFERENCES

- Albano, Gian Luigi and Alessandro Lizzeri (2001) “Strategic Certification and Provision of Quality,” *International Economic Review*, 42 (1), 267–283. [5]
- Ali, S Nageeb, Nima Haghpanah, Xiao Lin, and Ron Siegel (2022) “How to Sell Hard Information,” *The Quarterly Journal of Economics*, 137 (1), 619–678. [5]
- Banks, Jeffrey S and Joel Sobel (1987) “Equilibrium selection in signaling games,” *Econometrica: Journal of the Econometric Society*, 647–661. [28]
- Bergemann, Dirk, Alessandro Bonatti, and Alex Smolin (2018) “The Design and Price of Information,” *American Economic Review*, 108 (1), 1–48. [5, 6, 8, 10, 12, 17]
- Bergemann, Dirk, Tibor Heumann, and Stephen Morris (2022) “Screening with Persuasion,” *arXiv preprint arXiv:2212.03360*. [6]
- Bergemann, Dirk and Martin Pesendorfer (2007) “Information Structures in Optimal Auctions,” *Journal of Economic Theory*, 137 (1), 580–609. [6]
- Boleslavsky, Raphael and Aaron Kolb (2025) “The Design and Price of Influence,” *arXiv preprint arXiv:2408.03689*. [7]
- Celik, Gorkem and Roland Strausz (2025) “Informative Certification: Screening vs. Acquisition,” Technical report, CRC TRR 190 Rationality and Competition. [6]
- Chopra, Hershdeep and Jeffrey Ely (2025) “Incentive-Compatible Information Design,” Technical report, Mimeo. [7, 26]

- 1 Corrao, Roberto (2023) "Mediation Markets: The Case of Soft Informa- 1
2 tion," Technical report, Working Paper. [6] 2
- 3 Dvoretzky, Aryeh, Abraham Wald, and Jacob Wolfowitz (1951) "Elimination of 3
4 randomization in certain statistical decision procedures and zero-sum two-person 4
5 games," *The Annals of Mathematical Statistics*, 1–21. [29] 5
- 6 Ely, Jeffrey (2026) "Screening With Tests," Technical report, Mimeo. [7] 6
7 7
- 8 Grossman, Sanford J (1981) "The Informational Role of Warranties and Private 8
9 Disclosure About Product Quality," *The Journal of Law and Economics*, 24 (3), 461– 9
10 483. [5] 10
- 11 Grossman, Sanford J and Oliver D Hart (1980) "Disclosure Laws and Takeover 11
12 Bids," *The Journal of Finance*, 35 (2), 323–334. [5] 12
- 13 Hancart, Nathan (2024) "The Optimal Menu of Tests." [6] 13
14 14
- 15 Haupt, Andreas A, Nicole Immorlica, and Brendan Lucier (2023) "Certification 15
16 Design for a Competitive Market," *arXiv preprint arXiv:2301.13449*. [5] 16
- 17 Hörner, Johannes and Andrzej Skrzypacz (2016) "Selling Information," *Journal of* 17
18 *Political Economy*, 124 (6), 1515–1562. [5] 18
- 19 Ichihashi, Shota and Alex Smolin (2022) "Data Collection by an Informed Seller." 19
20 [5] 20
- 21 Kamenica, Emir and Matthew Gentzkow (2011) "Bayesian Persuasion," *American* 21
22 *Economic Review*, 101 (6), 2590–2615. [5, 16] 22
- 23 Kolotilin, Anton, Timofiy Mylovanov, and Andriy Zapechelnyuk (2022) "Censor- 23
24 ship as optimal persuasion," *Theoretical Economics*, 17 (2), 561–585. [21] 24
25 25
- 26 Lizzeri, Alessandro (1999) "Information Revelation and Certification Intermedi- 26
27 aries," *The RAND Journal of Economics*, 214–231. [5] 27
- 28 Luenberger, David G (1969) *Optimization by Vector Space Methods*: John Wiley & 28
29 Sons. [24, 34, 40, 48, 50] 29
- 30 Mekonnen, Teddy, Zeky Murra-Anton, and Bobak Pakzad-Hurson (2023) "Per- 30
31 suaded Search," *arXiv preprint arXiv:2303.13409*. [5] 31
32 32

- 1 Milgrom, Paul R (1981) “Good News and Bad News: Representation Theorems and 1
2 Applications,” *The Bell Journal of Economics*, 380–391. [5] 2
- 3 Mussa, Michael and Sherwin Rosen (1978) “Monopoly and product quality,” *Journal* 3
4 *of Economic theory*, 18 (2), 301–317. [6] 4
- 5 Myerson, Roger B (1981) “Optimal Auction Design,” *Mathematics of Operations Re-* 5
6 *search*, 6 (1), 58–73. [11] 6
- 7 7
- 8 Park, In-Uck and Robert Evans (2025) “Selling Information for Bilateral Trade,” 8
9 *Available at SSRN 5362376*. [5] 9
- 10 Rayo, Luis (2013) “Monopolistic signal provision,” *The BE Journal of Theoretical* 10
11 *Economics*, 13 (1), 27–58. [6] 11
- 12 Rayo, Luis and Ilya Segal (2010) “Optimal Information Disclosure,” *Journal of Polit-* 12
13 *ical Economy*, 118 (5), 949–987. [5] 13
- 14 14
- 15 Shaked, Moshe and J George Shanthikumar (2007) *Stochastic orders*: Springer. [34] 15
- 16 Stahl, Konrad and Roland Strausz (2017) “Certification and market transparency,” 16
17 *The Review of Economic Studies*, 84 (4), 1842–1868. [5] 17
- 18 18
- 19 Weksler, Ran and Boaz Zik (2023) “Monopolistic Screening in Certification Mar-
- 20 kets,” *Available at SSRN 4675804*. [6, 10] 20
- 21 Yang, Frank and Kai Hao Yang (2025) “Multidimensional Monotonicity and Eco-
- 22 nomic Applications,” *arXiv preprint arXiv:2502.18876*. [41] 22
- 23 Yang, Kai Hao (2022) “Selling Consumer Data for Profit: Optimal Market-
- 24 Segmentation Design and Its Consequences,” *American Economic Review*, 112 (4), 24
25 1364–1393. [5] 25
- 26 Zapechelnyuk, Andriy (2020) “Optimal Quality Certification,” *American Economic* 26
27 *Review: Insights*, 2 (2), 161–176. [5] 27
- 28 28
- 29 29
- 30 30
- 31 31
- 32 32