# How to Sell Hard Information[*]

Nima Haghpanah ⓡ S. Nageeb Ali ⓡ Xiao Lin ⓡ Ron Siegel [†]

October 15, 2020

Abstract

The seller of an asset has the option to buy hard information about the value of the asset from an intermediary. The seller can then disclose the acquired information before selling the asset in a competitive market. We study how the intermediary designs and sells hard information to robustly maximize her revenue across all equilibria. Even though the intermediary could use an accurate test that reveals the asset's value, we show that robust revenue maximization leads to a noisy test with a continuum of possible scores that are distributed exponentially. In addition, the intermediary always charges the seller for disclosing the test score to the market, but not necessarily for running the test. This enables the intermediary to robustly appropriate a significant share of the surplus resulting from the asset sale even though the information generated by the test provides no social value.

Contents

1 Introduction

This paper studies settings in which individuals purchase hard information from an intermediary that they can verifiably disclose to influence the actions of others.[1] Such settings are ubiquitous. For example, entrepreneurs often seek evidence that they can disclose to venture capitalists to obtain more funding, sellers of physical and financial assets routinely pay for evaluations that enable them to obtain better prices, and workers commonly seek certification before applying for positions. The size of the market for hard information, and certification in particular, is in the hundreds of billions.[2]

An important justification for the existence of information intermediaries in these settings is that they generate economic value. The information they provide may facilitate assortative matching or alleviate moral hazard or adverse selection.[3] It may also affect how the surplus is divided between the parties (other than the intermediary). For example, a seller who knows that his asset is valuable may be unable to credibly convey that information to buyers on his own, but objective evidence obtained from a trusted third-party would allow him to do so, enabling him to negotiate a better price for his asset.

But would we expect to find information intermediaries where they do not provide economic value? Perhaps not, since purchasing information from them is voluntary and if they do not provide economic value, they can make a profit only by reducing the surplus of the other agents. We consider such a setting with an information intermediary and show, perhaps surprisingly, that the intermediary is able to appropriate a significant part of the surplus arising from trade by generating noisy information and charging certain fees, even when the equilibrium played by the other parties is chosen adversarially to the intermediary's interests.

In our model, an agent owns an asset that he would like to sell in a competitive market. Both the agent and the market have symmetric information about the asset's market value. Before selling the asset, the agent can purchase additional, hard information from an intermediary about the asset's value that the agent can share with the market to improve the terms of trade. The intermediary chooses a *test*, which stochastically maps the asset's value to a score that can be verifiably disclosed, and a two-part tariff for her services. The tariff comprises a *testing fee* for running the test and a *disclosure fee* for disclosing the resulting score to the market. If the agent pays the testing fee, the test is run and the agent observes the resulting score. He then chooses whether to pay the disclosure fee to obtain hard information that enables him to disclose the score to the market. The market cannot distinguish between the agent

---

[1] For a survey of the literature on the design and sale of information by intermediaries see Bergemann and Bonatti (2019).

[2] See, for example, https://tinyurl.com/y5ce9p4b.

[3] Issues of moral hazard and certification have been studied by Albano and Lizzeri (2001), Marinovic, Skrzypacz, and Varas (2018), Zapechelnyuk (2020), and Saeedi and Shourideh (2020) among others.

not disclosing the score and the test not having been run. Because the market is competitive, the market price for the asset following disclosure or non-disclosure equals the asset's expected value conditioning on all the information available to the market and the equilibrium choices of the agent.

The intermediary's choice of test and fees shapes the disclosure game between the agent and the market: the test determines the degree and kind of private information the agent will have in the trading relationship, and the fees determine his learning and disclosure costs. The intermediary's goal is to extract as much surplus from the agent as possible. Should the intermediary use a noisy test to determine the asset's value or should she use an accurate one? What combination of testing and disclosure fees should she charge? Finally, how much of the trading surplus can she extract?

To address these questions, we first observe that the agent's outside option depends on the market's expectation of the agent's behavior. If the market expects the agent to pay the testing fee with some probability but the agent attempts to opt out of the disclosure game by not paying the testing fee, he cannot prove to the market that he chose to opt out. Instead, when the market sees no score being disclosed, it rationally concludes that the test may have been run but the agent chose not to disclose a low score. The market weights this "bad news" contingency and the "no news" contingency in which the agent decided not to pay the testing fee, where the weights depend on what the market believes the agent does in equilibrium. Thus, unlike in a standard mechanism-design problem, the agent's outside option depends on both the test-fee structure and the equilibrium that is played in the induced game between the agent and the market.

If the intermediary can select the equilibrium of the induced game, then an optimal test-fee structure comprises a fully revealing test, a high testing fee, and no disclosure fee. The intermediary selects an equilibrium in which the agent always pays the testing fee, and when the agent chooses not to disclose the score, the market believes that the test revealed that the asset has its lowest possible value (say $\underline{\theta}$). The agent discloses the test score whenever it reveals that the asset's value is not the lowest possible value. The high testing fee charged by the intermediary extracts all of the agent's expected surplus (minus $\underline{\theta}$) from selling the asset, so his expected payoff is $\underline{\theta}$.[4] This is consistent with a key intuition from standard mechanism design: since the agent's payoff beyond what is needed to satisfy his individual rationality constraint is due to information rents, when the agent starts with no private information the designer can keep the agent's payoff at his individual rationality level, extract the full surplus, and achieve this by charging an upfront fee.

But the game induced by this test-fee structure has another equilibrium, in which the agent

---

[4]The intermediary can also extract all the surplus by using a binary score test, making testing free, and charging a high disclosure fee.

never pays the testing fee so the intermediary's revenue is zero. In this equilibrium, the market treats non-disclosure as "no news," and the resulting market price is the asset's ex ante expected value. Given this, it is optimal for the agent not to pay for the intermediary's services. Thus, choosing this test-fee structure leaves the intermediary vulnerable to obtaining zero revenue. We show in Proposition 1 that this is not an accident: any test-fee structure that has an equilibrium in which the intermediary extracts (approximately) all the surplus also has an equilibrium in which the intermediary's revenue is (approximately) zero.

Motivated by the above discussion, our goal is to identify *robustly optimal test-fee structures*, namely those that guarantee the highest revenue to the intermediary across *all* equilibria of the induced game. This corresponds to the intermediary choosing the test-fee structure that maximizes her revenue assuming that the equilibrium of the induced game is selected adversarially to her interests. Our motivation for studying robustly optimal test-fee structures is twofold. First, the intermediary may be unable to coordinate the behavior of the agent and the market on her most preferred equilibrium. The uncertainty about which equilibrium will be played could motivate her to be cautious and therefore use test-fee structures that guarantee her a high revenue across all equilibria.[5] Second, for any test-fee structure, the sum of the agent and the intermediary's revenue is constant across equilibria and equal to the asset's ex ante expected value, so the intermediary's least preferred equilibrium is the agent's most preferred equilibrium. And the agent and the market may be able to coordinate on the agent's preferred equilibrium. As we will see, the intermediary obtains a substantial share of the surplus when using the robustly optimal test-fee structure. This will show that the ability of the intermediary to secure a profit even though the information she provides has no social value and does not benefit the agent (or the market) ex-ante does not depend on the equilibrium played in the induced game.

Finding the robustly optimal test-fee structure involves optimizing across all test-fee structures and all equilibria of the induced games. Different tests induce different distributions of private information about the asset's value for the agent in the induced games, and different fees change the agent's incentives to obtain and disclose this private information. Thus, the optimization entails comparing the equilibria of disclosure games that vary in both the amount of the agent's private information and his disclosure costs. Despite this richness, we find that robustly optimal test-fee structures take a relatively simple form regardless of the distribution of the asset's value. As we illustrate in Figure 1, robustly optimal tests are in the "step-exponential-step" class: they feature an exponential distribution over a continuum of scores (even if the asset's value is drawn from a finite set), one atom below this continuum, and pos-

---

[5]Our focus on adversarial equilibrium selection is shared by a rapidly growing literature in mechanism and information design, including Bergemann, Brooks, and Morris (2017), Du (2018), Hoshino (2019), Ziegler (2019), Inostroza and Pavan (2020), Dworczak and Pavan (2020), Halac, Kremer, and Winter (2020), Halac, Lipnowski, and Rappoport (2020), and Mathevet, Perego, and Taneva (2020).
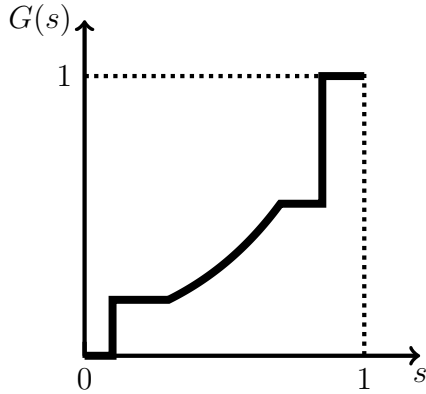
Figure 1: A robustly optimal score distribution if the asset quality θ is drawn from [0, 1], where G is the marginal CDF on scores. The score distribution features atoms on a low and a high score, and an exponential distribution over a continuum of intermediate scores.

sibly one atom above it. The optimal disclosure fee is always positive, but the optimal testing fee may be positive or 0. The resulting payoff to the intermediary is positive and bounded away from the full surplus.

To derive these features of the robustly optimal test-fee structure we first observe that the intermediary does not provide any added value to the agent ex ante. This is because the market draws correct (Bayesian) inferences and is competitive, so for every test-fee structure and any equilibrium, the ex ante expected market price is the ex ante expected value of the asset. Thus, for every test and for all positive fees, the agent strictly prefers an equilibrium in which the market expects him not to have the asset tested and consequently offers him the asset's ex ante expected value. The agent then retains the full surplus and the intermediary's revenue is zero. If the market were to observe whether the asset is tested, the agent could achieve this as an equilibrium outcome by not paying the testing fee. But the market *does not* observe whether the asset is tested. Therefore, the market's expectation of whether the asset has been tested and the agent observed the resulting score must be consistent with the agent's unobserved equilibrium choice of whether to pay the testing fee.

This is how the intermediary obtains a positive payoff robustly: she uses *option value as a carrot* to make it non-credible for the agent not to pay for the asset to be tested. Because the market only learns that the test has been run if the agent pays the disclosure fee and discloses the test score, the intermediary creates option value for the agent by offering a test that generates high test scores with some probability and setting the testing and disclosure fees sufficiently low. If this option value is sufficiently high, the agent cannot credibly refrain from paying the testing fee. The intermediary then obtains at least the testing fee in every equilibrium. Moreover, the market then treats non-disclosure as concealing a low score, which further motivates the agent to pay the disclosure fee and disclose the test score. In effect, the agent is trapped by market expectations that he has paid the testing fee and will disclose if the test score is sufficiently high.

4

But even if the agent pays the testing fee with certainty, multiple equilibria may exist. These equilibria differ in the set of scores that the agent discloses, and therefore in the probability of disclosure and the intermediary's revenue. The exponential test score distribution is robustly optimal because it eliminates potential equilibria in which the agent discloses with low probability. We develop an intuition for this result by showing that the intermediary can be thought of as choosing an optimal "demand curve for testing," subject to the demand curve being feasible and the quantity of testing demanded corresponding to the one in the equilibrium least favorable to the intermediary. We illustrate this approach in Section 2, and provide a general analysis in Sections 3–5.

Section 6 describes two extensions. First, we consider a setting in which testing is costly for the intermediary. We show that if testing costs increase in the Blackwell order, then our main results continue to hold, that is, there exists a robustly optimal test in the step-exponential-step class. Moreover, if the increase is strict, then every robustly optimal test is in this class. Second, we consider an intermediary who can sell the agent multiple pieces of evidence, and gives him the choice of which to disclose. We show that this additional flexibility does not improve the intermediary's revenue guarantee.

The main takeaway of our paper is that even if information is neither socially valuable nor ex ante valuable to a seller, an intermediary selling hard information can profit from the seller's inability to commit. The intermediary robustly maximizes her revenue by using test-fee structures that generate option value, use noise, and include strictly positive disclosure fees. Thus, the presence of profitable intermediaries, as in the large certification industry is not, in itself, evidence that the provision of hard information improves the welfare of market participants.

Our work builds on the rich literature on verifiable disclosure and persuasion games, initiated by Grossman (1981) and Milgrom (1981). Their main insight is that if a privately informed agent could costlessly and verifiably disclose evidence about his type, the unique equilibrium involves full disclosure. The subsequent literature suggests a number of mechanisms that dampen this force, including exogenously costly disclosure (Jovanovic, 1982; Verrecchia, 1983) and lacking evidence with positive probability (Dye, 1985). Matthews and Postlewaite (1985) and Shavell (1994) consider an uninformed agent who decides whether to take a fully revealing test. Matthews and Postlewaite (1985) show that the unique equilibrium involves testing and full disclosure of the test result when disclosure is voluntary but involves no testing when disclosure of the test result is mandatory. Shavell (1994) assumes the agent bears a privately-known cost of testing, and studies how this cost dampens unraveling.

The features in the models above also apply to our setting: the agent faces a cost of obtaining information about her type and a cost of disclosing that information in a verifiable form, and with positive probability, the agent may lack evidence. But we derive these features

endogenously because the intermediary chooses the evidence structure as well as the cost of learning and disclosing the evidence; the probability that the market attributes to the agent having evidence is determined in equilibrium. Treating these features as endogenous objects reveals a tradeoff: all else equal, the intermediary would like the market to unravel (so that the agent discloses with maximal probability), but the instruments from which she earns revenue are exactly those that counter unraveling. This "quantity-price" tradeoff leads to the price-theoretic approach to evidence generation we develop, in which the intermediary both chooses the optimal price and designs the optimal demand curve subject to constraints that correspond to Bayes rule and adversarial equilibrium selection.

A closely related strand of the disclosure literature studies choices made by agents to influence market perceptions. Ben-Porath, Dekel, and Lipman (2018) model how an agent chooses projects when he obtains evidence of project returns with positive probability. Their analysis emphasizes option value from the possibility of disclosure as motivating the agent to choose riskier projects. DeMarzo, Kremer, and Skrzypacz (2019) study how an agent chooses tests and disclosures to influence the market valuation of his asset. While most of their paper concerns evidence being costlessly generated in-house by the agent, they also consider the choices that would be made by a monopolistic intermediary who can charge a testing fee but not a disclosure fee, assuming that equilibria are selected to favor the intermediary. Our analysis shows that when the intermediary chooses to protect herself from the worst equilibrium (or assumes that the agent and market can coordinate on the agent's best equilibrium) and can charge a disclosure fee, she optimally designs her tests differently.[6]

In the papers discussed above, and in our work, the players are symmetrically informed at the outset of their interaction. By contrast, Lizzeri (1999) studies signaling dynamics when the agent is perfectly and privately informed at the outset about the value of the asset and chooses whether to have the asset tested. In his model, testing is voluntary (and observable) and if the agent has the asset tested, disclosure is mandatory. The intermediary's gains come from the prejudicial expectation that the market forms when the privately informed agent chooses to not have the asset tested. Lizzeri shows that the intermediary can extract the full surplus using a nearly uninformative test.[7] Were the agent instead uninformed prior to testing, the agent would retain the full surplus and the intermediary's revenue would be 0.[8] Our work

---

[6]Another strand of the literature less closely related to our work studies organizational settings in which an agent exerts effort to acquire hard information to persuade others; see Aghion and Tirole (1997), Dewatripont and Tirole (1999), and Che and Kartik (2009). Shishkin (2019) considers a sender who commits to an evidence structure and then chooses whether to disclose evidence to persuade a receiver. Both the probability of obtaining evidence and the form that evidence takes are determined endogenously. He finds that the optimal evidence structure is binary certification.

[7]Harbaugh and Rasmusen (2018) study a related model in which the intermediary's objective is to provide as much information as possible to the receiver. They show that the intermediary nevertheless chooses to provide coarse information to encourage the privately informed agent to have the asset tested.

[8]As an application of a novel result on Bayesian updating, Kartik, Lee, and Suen (2020) study a version of

sheds light on a different and complementary economic force: we show that when *both* testing and disclosure are voluntary, an intermediary can gain from using option-value to incentivize an uninformed agent to always have the asset tested. The intermediary then capitalizes on prejudicial inferences that are made if the agent chooses not to disclose the test score.

2   An Example

Consider an agent who will sell an asset in a competitive asset market. The market value of the asset, θ, is either 0 or 1, each with equal probability. Neither the agent nor the asset market know the asset's value, but an intermediary can run a test that generates information about the asset's value. The intermediary chooses the test, T, which stochastically maps the asset's value to an unbiased score s in $[0, 1]$, so $s = E[\theta|s]$. If the agent wants the asset tested, he has to pay the intermediary a testing fee ϕ_t. If the asset is tested, the score is reported to the agent. The agent then chooses whether to pay an additional disclosure fee ϕ_d to disclose the score as hard information to the market; otherwise, no score is disclosed to the market. If no score is disclosed, the market cannot distinguish between (i) the asset not being tested and (ii) the asset being tested but the agent not disclosing the score. The intermediary chooses both the test T and the fees (a test-fee structure), and her choice induces a disclosure game between the agent and the asset market. Some test-fee structures have multiple equilibria, and the intermediary's objective is to choose a test-fee structure that guarantees the highest payoff across equilibria.

We use this example to illustrate several features of our analysis. We show that the intermediary benefits from using noisy tests that pool low and high asset values as intermediate scores. We visually depict how our problem maps to setting a price on an optimally chosen demand curve. And we show why the robustly optimal test uses an exponential score distribution. To illustrate each of these features, it suffices to assume that testing is free and the intermediary charges only a disclosure fee. As we later show, these features also arise when it is optimal to charge a testing fee.

We first describe the intermediary's revenue guarantee if she used a fully revealing test, i.e., a test in which the score is equal to the asset's value θ. Figure 2(a) depicts our analysis of this case using a *(inverse) demand correspondence*, which for any disclosure fee in $[0, 1]$ (on the vertical axis) traces the corresponding probabilities of disclosure consistent with equilibria of the induced game. For any such disclosure fee, there is an equilibrium in which the agent has the asset tested, discloses a score of 1, and conceals a score of 0; if no score is disclosed, the market offers a price of 0 for the asset. The agent has no incentive to deviate because the payoff of $1 - \phi_d$ from disclosing a score of 1 exceeds the payoff from not disclosing, which is in turn

Lizzeri's model in which the agent is privately and imperfectly informed about the asset's value.
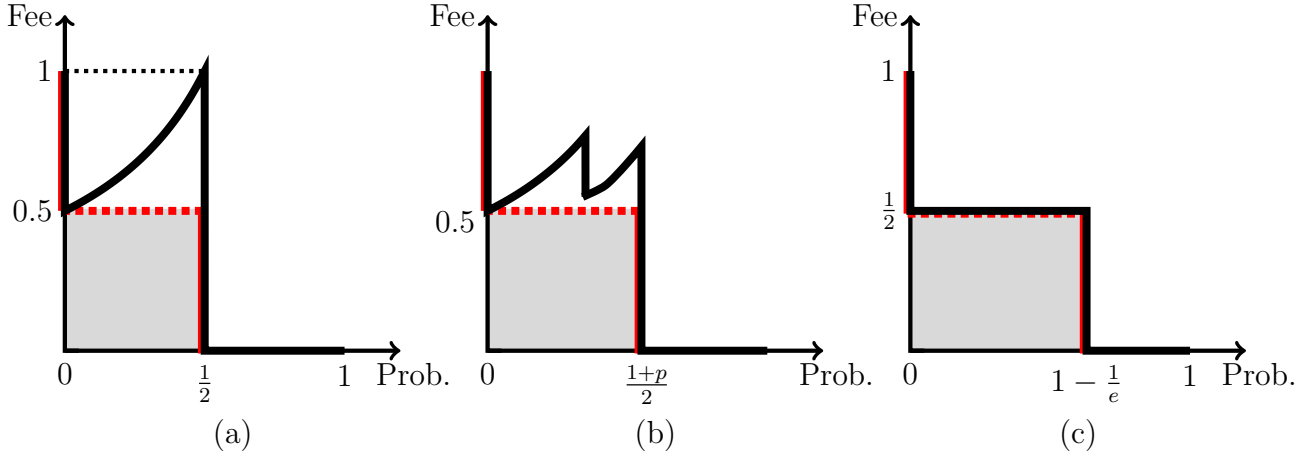
Figure 2: The black curve graphs the demand correspondence, and the red curve represents the robust demand curve for (a) the fully revealing test, (b) the test with three scores, and (c) the robustly optimal test.

equal to that from disclosing a score of 0. This is the unique equilibrium for any disclosure fee in $(0, 1/2)$, and the resulting revenue for the intermediary is half the disclosure fee.[9]

For disclosure fees that weakly exceed $1/2$, other equilibria exist. There is an equilibrium in which the agent has the asset tested but never discloses the score; if no score is disclosed, the market offers a price of $1/2$. Given this market price, the agent prefers not to disclose a score of 1 (or 0) because the disclosure fee is at least $1/2$. Because the agent does not disclose any score, the intermediary's revenue is zero in this equilibrium. There is also a mixed strategy equilibrium in which the agent discloses a score of 1 with an interior probability. For disclosure fees higher than $1/2$, the demand correspondence in Figure 2(a) depicts the three "quantities of disclosure" associated with these three equilibria.

Since we study the intermediary's robust revenue across equilibria, we identify the "robust demand curve" for disclosure, which maps each disclosure fee (price) to the *lowest* probability of disclosure (quantity) across all the equilibria associated with that fee. This is the red curve in Figure 2(a). The robust revenue for a given disclosure fee is the product of the disclosure fee and the lowest equilibrium probability of disclosure, i.e., the area of the rectangle under the robust demand curve at the point associated with that disclosure fee. The maximal robust revenue of $\approx 1/4$ for the fully revealing test is the area of the shaded rectangle in Figure 2(a), obtained from a disclosure fee that is slightly below $1/2$.[10]

Let us now see how the intermediary can robustly improve her revenue by introducing an intermediate score that pools the asset's two possible values, 0 and 1. Consider a test with three possible scores—0, 3/4, and 1—generated by the conditional distribution in Table 1. The demand correspondence and the robust demand curve for this test are shown in Figure 2(b).

---

[9]If the disclosure fee is 0, then for any probability p, there is an equilibrium in which the agent discloses a score of 1 and with probability p discloses a score of 0, since the agent's payoff is then 0 whether or not he discloses.

[10]As is common with adversarial equilibrium selection, the intermediary cannot robustly achieve a payoff of exactly $1/4$ with a disclosure fee of $1/2$—since there is then an equilibrium in which she obtains a payoff of 0.

8

	$s = 0$	$s = 3/4$	$s = 1$
$\theta = 0$	$1 - p$	p	0
$\theta = 1$	0	$3p$	$1 - 3p$

Table 1: Distribution of score s conditional on asset value θ, where p is in $(0, 1/3)$.

For strictly positive disclosure fees less than $1/2$, there is a unique equilibrium: the agent has the asset tested, and discloses his score if it is $3/4$ or 1; if no score is disclosed, the market offers a price of 0. The probability of disclosure in this equilibrium is $(1 + p)/2$. For disclosure fees higher than $1/2$, other equilibria exist, and all of them feature a lower probability of disclosure. Thus, by setting a disclosure fee slightly below $1/2$, the intermediary can obtain a revenue of $(1 + p)/4$, which is more than her revenue from the robustly optimal fully revealing test. By introducing the intermediate score $3/4$, the intermediary increases the quantity of disclosure demanded because the agent pays the disclosure fee not only when the asset's value is 1 but also, with some probability, when its value is 0.

But the intermediary is limited in how much pooling she can introduce: a value of p that is too high introduces a bad equilibrium in which the agent has the asset tested but only discloses a score of 1. If no score is disclosed, the market concludes that the score is either 0 or $3/4$ and offers a price of $3p/(1 + 3p)$. In this equilibrium, the probability of disclosure is $(1 - 3p)/2$. To eliminate this equilibrium, the intermediary sets the value of p so that if the score is $3/4$, the agent prefers to disclose his score rather than obtain the non-disclosure price:

$$\underbrace{\frac{3}{4} - \phi_d}_{\text{Disclose a score of } 3/4} \quad > \quad \underbrace{\frac{3p}{1 + 3p}}_{\text{Nondisclosure price}}$$

With a disclosure fee slightly below $1/2$, the above inequality holds for p less than $1/9$. The robust revenue is attained by setting $p = 1/9$ and charging a disclosure fee slightly below $1/2$. The restriction on p is a "cross-equilibrium" constraint, which ensures that a low-revenue equilibrium is not created. Cross-equilibrium constraints lead to the exponential score distribution in the robustly optimal tests.

To see why, let us return to the demand curve approach. These constraints lead to tests that induce a rectangular demand correspondence like the one in Figure 2(c): intuitively, whenever the demand correspondence or the robust demand curve are not a rectangle, there is some slack that the intermediary can use to change the test in a way that increases the probability of disclosure (by having more pooling) in the equilibrium that has the lowest probability of disclosure, , i.e., without violating cross-equilibrium constraints.

Which tests lead to rectangular demand correspondences? Those whose marginal score distribution includes an exponentially distributed component. To see why, notice that with a rectangular demand correspondence, there is a disclosure fee ϕ_d for which there is an interval
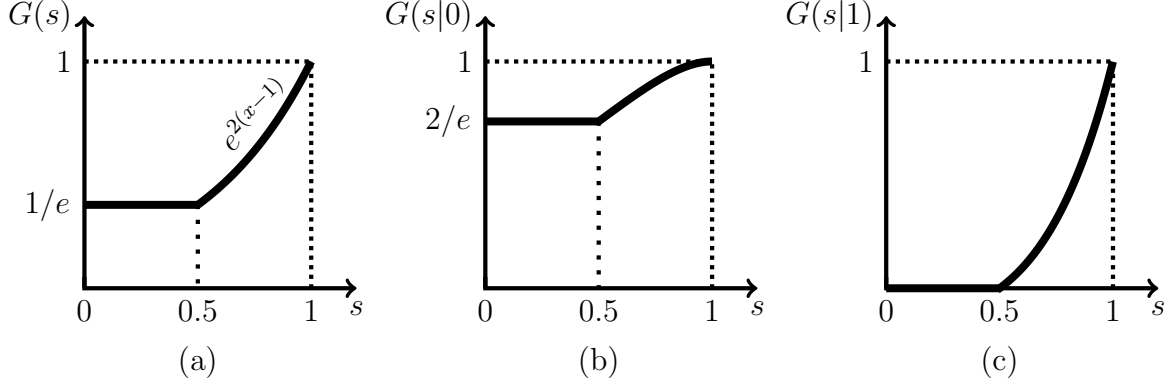
Figure 3: The robustly optimal distribution of scores where (a) $G(\cdot)$ is the marginal probability, (b) $G(\cdot|0)$ is the conditional probability of scores for $\theta = 0$, and (c) $G(\cdot|1)$ is the conditional probability for $\theta = 1$.

of equilibrium disclosure probabilities. In Figure 2(c), this disclosure fee is $\phi_d = 1/2$. This continuum of equilibria implies that there is a continuum of scores s for which

$$s - \phi_d = E[s'|s' \leq s]. \tag{1}$$

The LHS above is the market price from disclosing a score s, and the RHS is that from not disclosing when the set of scores that do not disclose are those weakly below s. Thus, when (1) holds there is an equilibrium with a disclosure threshold of s. Using $G(s)$ as the marginal CDF on scores and $G(s'|s' \leq s)$ as the conditional CDF on scores no higher than s, the RHS can be re-written as $\int_0^s (1 - G(s'|s' \leq s))ds'$. It follows then that Equation 1 is equivalent to

$$\phi_d = \int_0^s G(s'|s' \leq s)ds' = \frac{\int_0^s G(s')ds'}{G(s)} = \left(\frac{d}{ds} \Big(\ln(\int_0^s G(s')ds') \Big) \right)^{-1}. \tag{2}$$

Because the above is true for an interval of scores s, it defines a differential equation whose solution is $G(s) = \alpha e^{\frac{s}{\phi_d}}$ for some constant α, , i.e., the exponential score distribution.[11]

We illustrate this solution in Figure 3, where (a) shows the marginal score distribution in the robustly optimal test. This marginal distribution is generated by a test in which if the value of the asset is 0, with probability $2/e$ the test score is 0 and with complementary probability the test score is distributed exponentially on $[1/2, 1]$, and if the value of the asset is 1, the test score is distributed exponentially on $[1/2, 1]$. The testing fee is 0 (by assumption) and the disclosure

---

[11]The exponential score distribution may bring to mind the work of Roesler and Szentes (2017) and Condorelli and Szentes (2020). In those papers and in our work, a player optimally chooses a demand curve by manipulating an information structure or a distribution of values. In those papers, a buyer optimally chooses a unit-elastic demand curve that leads a seller to charge a low price and be indifferent between that price and higher prices. In a similar vein, Ortner and Chassang (2018) show that a principal can reduce the cost of monitoring by paying a monitor wages that generate a unit-elastic demand curve for bribes. In our paper, adversarial equilibrium selection leads the intermediary to choose a test that generates a rectangular demand curve, which corresponds to an exponential distribution of scores. The driving force is not to generate any indifference but to eliminate equilibria in which the agent does not disclose intermediate scores.

fee is slightly less than $1/2$. This test-fee structure induces a game with a unique equilibrium: the agent has the asset tested, and discloses the score if it exceeds the disclosure fee. If no score is disclosed the market offers a price of 0. The resulting revenue for the intermediary is $\approx \frac{1}{2}\left(1 - \frac{1}{e}\right)$, which is more than $1/4$ but less than the full surplus of $1/2$.

3   Model

3.1   The Setting

A risk-neutral agent sells an asset in a competitive market comprising two risk-neutral buyers who have a common value for the asset. The agent and both buyers are symmetrically informed about the asset's market value, θ, knowing only that it is drawn according to a distribution F with support $\Theta \subseteq [\underline{\theta}, \overline{\theta}]$, where $0 \leq \underline{\theta} < \overline{\theta} < \infty$ are the lowest and highest values in Θ, with an expected value of μ.

Prior to selling the asset, the agent can pay an intermediary to evaluate the asset, and if he does so, he has the option, for a fee, to disclose that evaluation to the market. The intermediary chooses how to evaluate the asset and sets the fees for her services, and we call such a scheme a *test-fee structure*. A test-fee structure comprises:

   a. a *test*, T, where $T : \Theta \to \Delta S$ is a measurable function that stochastically maps the asset's value to a score in some set S,
   b. a *testing fee*, $\phi_t \in \mathbb{R}$, which the agent pays for the asset to be evaluated, and
   c. a *disclosure fee*, $\phi_d \in \mathbb{R}$, which the agent pays to disclose the score to the market.

Because the test will only be used to determine the asset's expected market value, we henceforth assume without loss of generality that each test score is an unbiased estimate of the value. That is, $S \equiv [\underline{\theta}, \overline{\theta}]$ and for every market value θ in Θ and score s in S, $E[\theta | T(\theta) = s] = s$. We denote the set of all tests by $\mathcal{T}$. For each test T, let $G_T(s)$ denote the probability that if the intermediary evaluates the asset using test T, the resulting score is at most s. We refer to G_T as the *marginal score distribution*.

Turning to fees, we denote the pair of testing and disclosure fees by $\phi \equiv (\phi_t, \phi_d)$. The testing fee is the price the intermediary charges to evaluate and generate information about the asset and show it to the seller. The disclosure fee is the additional price to make that information hard or verifiable so that the agent can disclose it to the buyers.

We denote the test-fee structure by (T, ϕ). Each test-fee structure (T, ϕ) induces a game $\mathcal{G}(T, \phi)$ between the agent and the buyers with the following timeline:

   1. Nature chooses the value θ of the asset according to the distribution F.

2. The agent, without observing θ, decides whether to have the asset tested. If so, he pays the testing fee ϕ_t and observes a score s drawn according to $T(\theta) \in \Delta S$.

3. The agent decides whether to disclose the score to the market. If so, he pays the disclosure fee ϕ_d and the buyers observe the score. If he does not disclose, or if he does not take the test, the buyers observe a null message, N.

4. The buyers bid simultaneously for the asset, and the asset is sold to the highest bidder, with ties broken uniformly.

The agent's payoff is the price at which he sells the asset minus any testing and disclosure fees that he pays. The payoff of the buyer that buys the asset at price p is $\theta - p$; the other buyer's payoff is zero.[12]

In the induced game $\mathcal{G}(T, \phi)$, an agent's (behavioral) strategy is (σ_t, σ_d), where σ_t in $[0, 1]$ is the probability with which he has the asset tested and $\sigma_d : S \to [0, 1]$ is the score-contingent probability with which he discloses the score. Let $\hat{S}(T) \equiv \cup_{\theta \in \Theta} Im(T(\theta))$ denote the set of scores in the image of T. Buyer i's (pure) strategy $\sigma_i : \hat{S}(T) \cup \{N\} \to \mathbb{R}$ specifies buyer i's bid following a disclosure of score s or the null message N. A belief system $\rho_i : \hat{S}(T) \cup \{N\} \to \Delta \Theta$ specifies buyer i's posterior belief about the asset's value following a disclosure of score s or the null message N. We denote a strategy profile $\sigma = (\sigma_t, \sigma_d, \sigma_1, \sigma_2)$ and belief system $\rho = (\rho_1, \rho_2)$ by (σ, ρ).

For a game $\mathcal{G}(T, \phi)$, we denote the set of Perfect Bayesian Equilibria (henceforth PBE) by $\Sigma(T, \phi)$.[13] Because the buyers have the same beliefs and compete in a first-price auction, each buyer's bid is equal to the asset's expected value given all the available information.

3.2   Maximal Revenue Guarantees

We study the intermediary's *maximal revenue guarantee*, namely the highest payoff that she can guarantee herself by choosing a test-fee structure, i.e., assuming that the equilibrium in the induced game is chosen adversarially to her interests.

---

[12]We have framed our analysis in terms of this market game for concreteness. But the issues and our analysis apply to other settings in which an agent obtains evidence to persuade others. For example, a principal may decide how much to invest in a project based on her beliefs about the project's success, and the agent may acquire evidence to persuade the principal to invest more. Our model is isomorphic to such a setting if the principal's investment increases linearly in her posterior expectation and the agent's payoff increases linearly in the principal's investment.

[13]A PBE corresponds to each player behaving in a way that is sequentially rational, beliefs μ_i being derived via Bayes Rule whenever possible, and off-path beliefs satisfying "You can't signal what you don't know" (Fudenberg and Tirole, 1991). The important implication for our setting is that both buyers have the same beliefs ($\rho_1(s) = \rho_2(s)$ for any score s) both on and off the equilibrium path.

Definition 1. *The intermediary's **maximal revenue guarantee** is*

$$R_M \equiv \sup_{(T,\phi)\in\mathcal{T}\times\mathbb{R}^2} \inf_{(\sigma,\rho)\in\Sigma(T,\phi)} \sigma_t\left(\phi_t + \phi_d\int_S \sigma_d(s)dG_T\right). \tag{3}$$

Recall that σ_t is the probability with which the agent pays for the intermediary to test the asset and $\sigma_d(s)$ is the probability with which the agent pays to disclose a score of s. The first term in the parenthesis in (3) is the intermediary's revenue from the testing fee, and the second term is the revenue from the disclosure fee.

We compare this maximal revenue guarantee with the *full informational surplus* that the intermediary could (conceivably) extract in equilibrium, which is $R_F \equiv \mu - \underline{\theta}$, where μ is the expected value of the asset. To see why, notice that because the market price of the asset is never lower than $\underline{\theta}$, the agent can guarantee himself that payoff by never taking the test. And since the total surplus in the economy is μ, the intermediary cannot obtain more than R_F in any equilibrium.

Why do we study the maximal revenue guarantee? Although there are test-fee structures that have equilibria in which the intermediary obtains the full informational surplus, such test-fee structures have other equilibria in which the intermediary obtains 0. We alluded to this fact in the introduction and in the example in Section 2. In fact, for any test-fee structure in which the intermediary obtains close to the full informational surplus in some equilibrium, there is another equilibrium in which the intermediary obtains close to 0.

Proposition 1. *There exists $\varepsilon^* > 0$ such that for any $\varepsilon < \varepsilon^*$, if there exists an equilibrium of a test-fee structure (T,ϕ) in which the intermediary's revenue is at least $R_F - \varepsilon$, then there exists another equilibrium of that test-fee structure in which the intermediary's revenue is no more than $\delta(\varepsilon)$, where $\lim_{\varepsilon\to 0}\delta(\varepsilon) = 0$.*

Proposition 1 formalizes the challenge that multiple equilibria present to the intermediary: choosing a test-fee structure because its most favorable equilibrium generates a high revenue leaves the intermediary vulnerable to an equilibrium that generates low or zero revenue. An immediate implication of Proposition 1 is that the maximal revenue guarantee, R_M, is bounded away from the full information surplus, R_F, by at least ε^*. Our main result, Proposition 2, derives a tight bound on R_M that, for every μ, applies across distributions for which the asset's expected value is μ. Because the proof of Proposition 1 uses techniques that we develop later in the paper, we do not develop the intuition here. The proof and a graphical intuition are in the Appendix.

A separate rationale for focusing on the maximal revenue guarantee is that for any test-fee structure, the equilibrium that minimizes the intermediary's revenue also maximizes the agent's payoff. Thus, if the intermediary fears that the agent and the asset market will coordinate on

the agent's preferred equilibrium, she would choose a test-fee structure that attains her maximal revenue guarantee.[14]

Our analysis also shows how to maximize the intermediary's revenue across test-fee structures that admit a *unique* equilibrium. The value of this related problem is at most R_M, but because the robustly optimal test-fee structure that we identify in Proposition 2 has a unique equilibrium, it is also a solution to this related problem.

4 Preliminary Steps

To simplify the problem of solving for the maximal revenue guarantee, we take the following preliminary steps:

1. We frame the analysis of tests purely in terms of marginal score distributions.
2. We show that it suffices to consider only those test-fee structures in which the agent pays the testing fee with probability 1 in every equilibrium.
3. For every such test-fee structure, we characterize the equilibrium with the lowest revenue for the intermediary in terms of a score threshold for disclosure.
4. Because finding the maximal revenue guarantee involves optimizing with strict constraints, we formulate a relaxed problem with weak constraints that has the same solution.

4.1 Shifting from Tests to Marginal Score Distributions

For the intermediary's revenue, all that matters about a test is its marginal score distribution: for all fees, if two tests have the same marginal score distribution, then they have the same equilibria.[15] Thus, henceforth, we refer to (G, ϕ) as a test-fee structure, where G is a CDF on $[\underline{\theta}, \overline{\theta}]$ that corresponds to the marginal score distribution of some test T and ϕ is a pair of testing and disclosure fees. We denote by $\underline{s}_G$ the lowest score in the support of G.

Focusing on the set of marginal score distributions induced by all possible tests is useful because this set is easy to characterize. Recall that a distribution G is a *mean-preserving contraction* of F if its support is in $[\underline{\theta}, \overline{\theta}]$ and $\int_{\underline{\theta}}^{s'} G(s)ds \leq \int_{\underline{\theta}}^{s'} F(s)ds$ for all $s' \in [\underline{\theta}, \overline{\theta}]$, with equality at $s' = \overline{\theta}$. We denote by $\Gamma(F)$ the set of distributions that are mean-preserving

---

[14]Thus, in the game induced by a test-fee structure, the agent's preferred equilibrium is the uniquely Pareto efficient equilibrium from the perspective of players of that game.

[15]Take two test-fee structures (T, ϕ) and (T', ϕ) such that T and T' have the same marginal score distribution. Take an equilibrium (σ, ρ) of $\mathcal{G}(T, \phi)$. Suppose that (σ, ρ) were played in $\mathcal{G}(T', \phi)$ and observe that given ρ and that $G_T = G'_T$, the strategy profile σ remains sequentially rational for the agent and the buyers. Moreover, given σ, ρ continues to satisfy Bayes' Rule and the appropriate consistency condition. Therefore, (σ, ρ) is an equilibrium of $\mathcal{G}(T', \phi)$.

contractions of F. We then have the following classical result.[16]

Lemma 0. *A marginal score distribution G is induced by an (unbiased) test if and only if G is in $\Gamma(F)$.*

We use this formulation to rewrite revenue guarantees: for a test-fee structure (G, ϕ), let $\hat{\Sigma}(G, \phi)$ be the equilibrium set of the induced game between the agent and the market. The *revenue guarantee* of (G, ϕ) is the lowest revenue generated in any equilibrium of this test-fee structure: $R(G, \phi) \equiv \inf_{(\sigma, \rho) \in \hat{\Sigma}(G, \phi)} \sigma_t \left(\phi_t + \phi_d \int_S \sigma_d(s) dG \right)$. The maximal revenue guarantee is thus $R_M = \sup_{(G, \phi) \in \Gamma(F) \times \mathbb{R}^2} R(G, \phi)$.

4.2   Using Option-Value as a Carrot

We show that in every test-fee structure, either the asset is tested with probability 1 in every equilibrium or there exists an equilibrium in which the asset is tested with probability 0 (and the intermediary's revenue is 0).

Lemma 1. *If a test-fee structure (G, ϕ) satisfies*

$$\phi_t < \int_{\mu + \phi_d}^{\bar{\theta}} [s - (\mu + \phi_d)] dG, \tag{P}$$

then the asset is tested with probability 1 in every equilibrium; otherwise, there exists an equilibrium in which the asset is tested with probability 0.

The logic of Lemma 1 is that (P) is a *participation constraint* that must hold for the asset to be tested with positive probability in every equilibrium. To see why, consider a test-fee structure (G, ϕ) and an equilibrium in which the asset is tested with probability 0. Because the market expects non-disclosure with probability 1, the price of the asset conditional on non-disclosure is μ. If the agent deviates and has the asset tested, then he optimally pays ϕ_d and discloses the score whenever it is higher than $\mu + \phi_d$. This is strictly profitable if

$$\mu < -\phi_t + \underbrace{\int_{\underline{\theta}}^{\bar{\theta}} \max\{\mu, s - \phi_d\} dG}_{\text{Option Value}},$$

which, by re-arranging, yields (P). Thus, if (P) holds, the asset is tested with positive probability in every equilibrium. The proof of Lemma 1 shows if (P) holds, the asset is in fact tested with probability 1 in every equilibrium.

Using this result, we show that it suffices to restrict attention to non-negative fees: for any test-fee structure with a negative testing or disclosure fee, the intermediary can improve her revenue guarantee by using a test-fee structure that has non-negative fees.

Lemma 2. *For any test-fee structure (G, ϕ) with $\phi_d < 0$ or $\phi_t < 0$, there exists a test-fee structure (G', ϕ') with non-negative fees such that $R(G, \phi)$ is strictly less than $R(G', \phi')$.*

4.3 Adversarial Disclosure Thresholds

For any test-fee structure that satisfies (P), an adversarial equilibrium is one that minimizes the disclosure probability, but in which the asset is tested with probability 1. We show that such equilibria are threshold equilibria in which the agent does not disclose his score if he obtains a score at the threshold.

With Lemmas 1 and 2 in hand, we restrict attention to test-fee structures that satisfy (P) and have non-negative disclosure fees. Given a test-fee structure (G, ϕ), consider thresholds τ that weakly exceed $\underline{s}_G$, where $\underline{s}_G$ is the lowest score in the support of G. We say that such a threshold τ is an *equilibrium threshold* for a test-fee structure (G, ϕ) if

$$\tau - \phi_d = E_G[s|s \le \tau]. \tag{4}$$

If τ satisfies (4), then there is an equilibrium in which the asset is tested with probability 1 and the agent discloses the score if and only if it strictly exceeds τ. To see why, suppose that the agent behaves in this way and (4) holds. Then the LHS of (4) is the difference between the market price of the asset when a score of τ is disclosed and the disclosure fee, and the RHS is the market price for the asset when no score is disclosed. Thus, the agent is indifferent between disclosing and not disclosing a score of τ. Because the market price following disclosure increases in the score but the market price following non-disclosure is constant in the score, the agent strictly prefers not to disclose scores lower than τ and to disclose scores higher than τ.

A test-fee structure may have multiple equilibrium thresholds. It may also have other equilibria, in which if the agent obtains a threshold score, he chooses to disclose his score with strictly positive probability rather than probability 0. The mixed strategy equilibria for the fully revealing and the 3-score tests in Section 2 are examples of such equilibria. We prove in Lemma 3 below that from the standpoint of adversarial equilibrium selection, it suffices to focus on equilibria in which the agent discloses his score only if *strictly* exceeds the threshold.

We proceed as follows. For each test-fee structure, we show that a highest equilibrium threshold exists and provide a characterization of the highest equilibrium threshold that we later use to find the robustly optimal test-fee structure. Finally, we show that this highest equilibrium threshold corresponds to an adversarial equilibrium in which the agent discloses his

score if and only if it strictly exceeds the highest equilibrium threshold.

Lemma 3. *If a test-fee structure (G, ϕ) satisfies* (P) *and $\phi_d \geq 0$, then the following are true:*

(a) *A highest equilibrium threshold τ exists.*

(b) *A score threshold $\tau \geq \underline{s}_G$ is the highest equilibrium threshold if and only if*

$$
\begin{aligned}
\tau - \phi_d &= E_G[s|s \leq \tau], \\
\tau' - \phi_d &> E_G[s|s \leq \tau'] \qquad \forall \tau' > \tau.
\end{aligned}
\tag{HE}
$$

(c) *There exists an adversarial equilibrium in which the agent discloses score s if and only if $s > \tau$ where τ is the highest equilibrium threshold.*

Figure 4 illustrates our characterization of the highest equilibrium threshold.
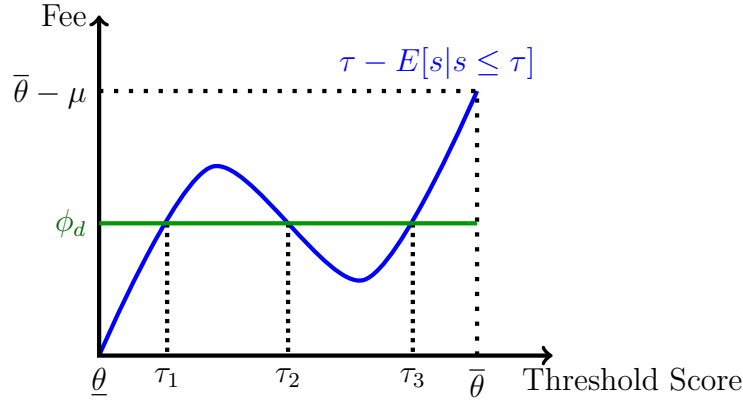
Figure 4: The worst equilibrium at a disclosure fee is that with the highest equilibrium threshold. For example, at a disclosure fee of ϕ_d, the highest equilibrium threshold is τ_3.

4.4 A Relaxation with an Identical Value

We have seen that if a test-fee structure satisfies (P), then the agent takes the test with probability 1 and the adversarial equilibrium is characterized by the highest equilibrium threshold. The revenue guarantee in the robustly optimal test-fee structure is therefore

$$
R_M = \sup_{(G,\phi,\tau)\in\Gamma(F)\times\mathbb{R}^3} \phi_t + \phi_d(1 - G(\tau)) \qquad \text{s.t. (P) and (HE).}
$$

The intermediary's maximization problem has constraints with strict inequalities, but we show that the problem's value R_M is unchanged if those inequalities are made weak. These are the *weak participation constraint*

$$
\phi_t \leq \int_{\mu+\phi_d}^{\bar{\theta}} [s - (\mu + \phi_d)]dG
\tag{w-P}
$$

17

and the *weak-highest equilibrium constraint* (which defines the weak-highest equilibrium threshold),

$$\tau - \phi_d = E_G[s|s \leq \tau]$$
$$\tau' - \phi_d \geq E_G[s|s \leq \tau'] \qquad \forall \tau' > \tau. \tag{w-HE}$$

We write the relaxed problem as

$$\max_{(G,\phi,\tau)\in\Gamma(F)\times\mathbb{R}^3} \phi_t + \phi_d(1 - G(\tau)) \qquad \text{s.t. (w-P) and (w-HE)}.$$

In addition to having a solution and the same value as the original problem, the relaxed problem has several attractive features. First, for any test-fee structure that solves the relaxed problem, there is a "nearby" test-fee structure whose revenue guarantee is close to the maximal revenue guarantee R_M. Second, for any convergent sequence of test-fee structures that achieves the maximal revenue guarantee, the limiting test-fee structure is a solution to the relaxed problem. Thus, solutions to the relaxed problem identify necessary and sufficient features of test-fee structures whose revenue guarantees approximate R_M. We call these solutions *robustly optimal test-fee structures*.

Let us formalize this discussion. A sequence of test-fee structures $\{(G^n, \phi^n)\}_{n=1,2,\ldots}$ converges to a test-fee structure (G, ϕ) if G^n converges weakly to G and ϕ^n converges to ϕ. For a test-fee structure (G, ϕ) and threshold τ, we denote the associated equilibrium revenue for the intermediary by $\hat{R}(G, \phi, \tau)$. We use this notation to link the relaxed and original problems.

Lemma 4.

(a) *An optimal solution (G, ϕ, τ) to the relaxed problem exists and $R_M = \hat{R}(G, \phi, \tau)$.*

(b) *Consider any optimal solution (G, ϕ, τ) to the relaxed problem. Then there exists a sequence of test-fee structures and thresholds $\{(G^n, \phi^n, \tau^n)\}_{n=1,2,\ldots}$ such that (i) for each n, (G^n, ϕ^n, τ^n) satisfy (P) and (HE), (ii) (G^n, ϕ^n) converges to (G, ϕ), and (iii) $R_M = \lim_{n\to\infty} \hat{R}(G^n, \phi^n, \tau^n)$.*

(c) *Consider any sequence of test-fee structures and thresholds $\{(G^n, \phi^n, \tau^n)\}_{n=1,2,\ldots}$ such that (i) for each n, (G^n, ϕ^n, τ^n) satisfy (P) and (HE), (ii) (G^n, ϕ^n) converges to a test-fee structure (G, ϕ), and (iii) $R_M = \lim_{n\to\infty} \hat{R}(G^n, \phi^n, \tau^n)$. Then there exists τ such that (G, ϕ, τ) satisfy (w-P) and (w-HE), and $R_M = \hat{R}(G, \phi, \tau)$.*

5 Robustly Optimal Test-fee Structures

This section contains our main result. We show that robustly optimal test-fee structures use tests that have a "step-exponential-step" form; i.e., the distributions over scores are exponential over an interval, have up to two mass points, one above and one below the interval, and have zero density everywhere else. We also show that the optimal disclosure fees are strictly positive, and we derive a tight bound on the intermediary's maximal revenue guarantee.
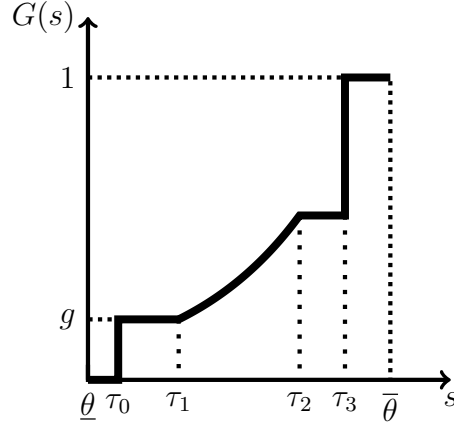
Figure 5: A step-exponential-step distribution G.

As illustrated in Figure 5, a test-fee structure (G, ϕ) is in the *step-exponential-step* class if there exists $g \in [0, 1]$ and thresholds $\tau_0 < \tau_1 < \tau_2 \le \tau_3$ such that

$$
G(s) = \begin{cases} g & \text{if } s \in [\tau_0, \tau_1) \\ g e^{(s-\tau_1)/(\tau_1-\tau_0)} & \text{if } s \in [\tau_1, \tau_2) \\ 1 & \text{if } s \ge \tau_3, \end{cases} \tag{5}
$$

G assigns probability 0 to $[0, \tau_0) \cup (\tau_2, \tau_3)$, and the fees are

$$
\phi_d = \tau_1 - \tau_0, \tag{6}
$$
$$
\phi_t = (1 - g e^{(\tau_2 - \tau_1)/\phi_d})(\tau_3 - (\mu + \phi_d)). \tag{7}
$$

Proposition 2. *For any distribution F of the asset's value, the following hold:*

(a) *There exists a test-fee structure in the step-exponential-step class that is robustly optimal.*

(b) *Every robustly optimal test-fee structure has a strictly positive disclosure fee.*

(c) *If (G, ϕ) is a robustly optimal test-fee structure, then the testing fee ϕ_t is strictly positive if and only if scores in $(\mu + \phi_d, \overline{\theta}]$ arise with positive probability.*

(d) *The maximal revenue guarantee is at most $(\overline{\theta} - \mu)(1 - e^{(\underline{\theta}-\mu)/(\overline{\theta}-\mu)})$, and this bound is attained when the support of F is binary, i.e., $\{\underline{\theta}, \overline{\theta}\}$.*
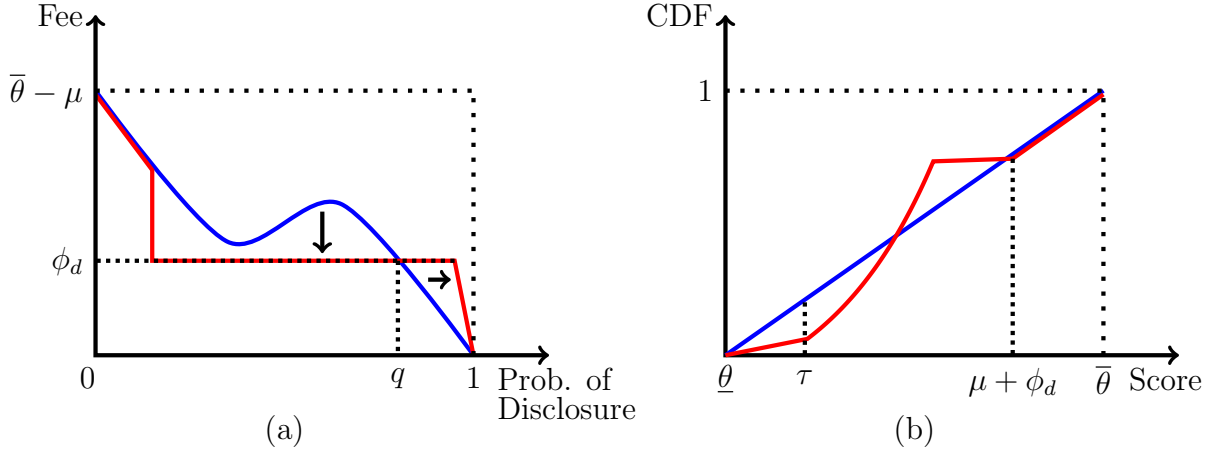
19

Figure 6: How an intermediary gains from "flattening" the robust demand curve. (a) depicts a demand curve in blue, and by flattening the demand curve the intermediary can induce a higher probability of disclosure at disclosure fee ϕ_d. (b) shows that this can be done without changing the score distribution above $\mu + \phi_d$, which guarantees that (P) continues to be satisfied without changing the testing fee.

Proposition 2 partially characterizes the robustly optimal test-fee structures, and provides a tight upper bound on the maximal revenue guarantee. There is always a robust solution in the step-exponential-step class, disclosure fees are strictly positive in any robust solution, and testing fees are strictly positive only if the robustly optimal score distribution takes a particular form. We show in the proof of Proposition 2(a) that if the intermediary uses the robustly optimal test in the step-exponential-step class, charges the testing fee in (7) and a disclosure fee slightly below those in (6), then that test-fee structure has a unique equilibrium. In that equilibrium, the asset is tested with probability 1 and the agent discloses his score whenever it exceeds τ_1. Our previous result, Proposition 1, already established that the maximal revenue guarantee is bounded away from full surplus, but the bound provided was not tight; (d) specifies a tight bound.

Proposition 2 clarifies the interaction between the testing fee, the disclosure fee, and the test. Lemma 10 in the Appendix shows that if the intermediary could not charge a disclosure fee, she would choose a fully revealing test and a testing fee that the agent pays with probability 1. In the robustly optimal solution, the intermediary uses noisy tests because she can charge a disclosure fee. If she could not charge a testing fee, there would be a solution in the step-exponential-step class that would not have a step at the top (above the exponential part of the curve). Her ability to charge a testing fee in addition to a disclosure fee leads to a solution in the step-exponential-step class that may have a step at the top.

The intuition for Proposition 2(a) builds on the robust demand curve approach described in Section 2. Consider a test-fee structure with disclosure fee ϕ_d in which the probability of disclosure in an adversarial equilibrium is q. Suppose, as is shown in Figure 6(a), that the demand correspondence is not flat to the left of the point (q, ϕ_d). Then we can modify the test-fee structure and improve the intermediary's robust revenue as follows: flatten the demand

20

correspondence to the left of (q, ϕ_d) and push it to the right of (q, ϕ_d). Doing this increases the disclosure probability in the adversarial equilibrium for any disclosure fee slightly lower than ϕ_d. Moreover, as seen in Figure 6(b), this modification can be done without changing the score distribution on scores above $\mu + \phi_d$, so the option value of a test, given by the RHS of (P), remains unchanged. Therefore, this modification does not generate a new adversarial equilibrium in which the agent doesn't pay the testing fee. Thus, the intermediary can increase the probability of disclosure by decreasing the disclosure fee arbitrarily slightly and without changing the testing fee, which improves her revenue guarantee.[17]

A similar logic shows that every robustly optimal test-fee structure has a score distribution that is exponential over a non-degenerate interval (Lemma 12 in the Appendix). Outside of that interval, there is some flexibility because scores below τ_1 are those that the agent strictly prefers not to disclose, and scores above $\mu + \phi_d$, which exceeds τ_2, are those for which (HE) is slack. In the step-exponential-step distribution scores below τ_1 are pooled as a single score and scores above $\mu + \phi_d$ are pooled as a single score. This creates a score distribution that is a mean-preserving contraction of the distribution F of the asset value. But there may be other robustly optimal distributions that are the same on the interval $[\tau_1, \tau_2]$ and are also mean-preserving contractions of F.

Turning to disclosure fees, we prove Proposition 2(b) by showing that for the intermediary to charge a testing fee, the test must induce a high option value by having high scores in its support. But then the gain of adding a small disclosure fee outweighs the resulting reduction in the option value and the associated reduction in the testing fee.

Proposition 2(c) follows from the participation constraint (P). If scores in $(\mu + \phi_d, \overline{\theta}]$ arise with positive probability., then even if the market expects the asset to not be tested, the agent strictly prefers to have the asset tested at a testing fee of 0 and a disclosure fee of ϕ_d. Increasing the testing fee slightly while keeping the disclosure fee unchanged increases the intermediary's revenue in the adversarial equilibrium and does not generate a new equilibrium in which the asset is not tested. We use this logic in Proposition 4 below to provide conditions on primitives that guarantee a strictly positive testing fee.

We prove Proposition 2(d) by considering a relaxed problem in which the score distribution need not be a mean-preserving contraction of distribution F, while still having the same expectation, i.e., $E_G[s] = E_F[\theta]$. We solve this relaxed problem completely, and its value then provides an upper bound on the maximal revenue guarantee. This bound is tight when the support of F is binary because the mean-preserving contraction condition is then equivalent to $E_G[s] = E_F[\theta]$. The following result describes the complete solution for this case.

---

[17]A reader may wonder how this is possible while the total surplus is constant and (w-P) (defined on p. 17) binds. This is because (w-P) is a non-standard participation constraint. The modification to the test decreases the agent's payoff in the adversarial equilibrium without affecting his indifference between paying the testing fee and not paying it if the market expects him not to pay it, which is what (w-P) represents.

Proposition 3. *Suppose that the support of distribution F is $\{\underline{\theta}, \overline{\theta}\}$. The unique robustly optimal test-fee structure involves:*

(a) *No testing fee but a strictly positive disclosure fee: $\phi_t^* = 0$ and $\phi_d^* = \overline{\theta} - \mu$.*

(b) *The following marginal score distribution that has an atom at $\underline{\theta}$, a gap above it, and then an exponential form (with no atom at the top):*

$$G^*(s) = \begin{cases} e^{\frac{\underline{\theta} - \mu}{\overline{\theta} - \mu}} & \text{if } s \in [\underline{\theta}, \underline{\theta} + \overline{\theta} - \mu) \\ e^{\frac{s - \overline{\theta}}{\overline{\theta} - \mu}} & \text{if } s \in [\underline{\theta} + \overline{\theta} - \mu, \overline{\theta}]. \end{cases} \tag{8}$$

We know from Proposition 2 that the robustly optimal disclosure fee is always strictly positive; Proposition 3 shows that the robustly optimal testing fee may be 0. We now provide a sufficient condition on the distribution F for the optimal testing fee to be strictly positive.

Proposition 4. *If the distribution F of the asset's value is log-concave, then any robustly optimal test-fee structure includes a strictly positive testing fee.*

Several commonly studied distributions such as the uniform distribution over $[\underline{\theta}, \overline{\theta}]$ and the truncated Normal and Pareto distributions are log-concave (Bagnoli and Bergstrom, 2005). Log-concavity requires the tail of the distribution to be less heavy than the tail of the exponential distribution.

6 Extensions

This section shows that our results also hold when testing is costly for the intermediary and when the intermediary can offer the agent multiple pieces of evidence that he can disclose.

6.1 Costly Tests

Our analysis assumed that testing is costless, but in reality testing is often costly. Suppose that the cost to the intermediary of running a test with marginal score distribution G is $c(G)$. We assume that c is lower semi-continuous in the weak* topology and monotone in the Blackwell order, that is, garbling a test weakly reduces its cost; in other words, whenever G' is a mean-preserving contraction of G, $c(G') \leq c(G)$.[18] This condition is standard when information acquisition is costly, and corresponds to a less informative test being less costly to generate.[19] A special case is when the intermediary starts with a finite set of initial tests and can garble those tests to obtain additional tests.

---

[18]Lower semi-continuity guarantees that solutions to the relevant maximization problems exist.

[19]Recent examples of analyses that assume monotonicity in the Blackwell order are De Oliveira, Denti, Mihm, and Ozbek (2017) and Pomatto, Strack, and Tamuz (2019).

The following is a corollary of our existing results.

Proposition 5. *If costs are lower semi-continuous and monotone in the Blackwell order, there exists a robustly optimal test-fee structure in the step-exponential-step class, and every robustly optimal test-fee structure uses a strictly positive disclosure fee.*

Proposition 5 is a corollary of Lemma 12 in the Appendix, which is one of the steps in the proof of Proposition 2. Lemma 12 shows that for any test-fee structure (G, ϕ), there exists a test-fee structure in the step-exponential-step class that uses a mean-preserving contraction of G and has a weakly higher revenue guarantee. Because such a test-fee structure is a garbling of G, it has a weakly lower cost. Therefore, there must exist a robustly optimal test-fee structure in this class.[20] The argument used to prove Proposition 2 proves that the optimal disclosure fee is strictly positive.

6.2   Multiple Pieces of Evidence

Our analysis assumed that the intermediary provided the agent with a single piece of evidence that he could verifiably disclose. One could envision the intermediary providing the agent with multiple pieces of evidence and a choice of which to disclose. We show that this additional generality does not change the robustly optimal test-fee structures.

To see this, suppose that the intermediary designs, along with the test-fee structure, an arbitrary evidence structure, which specifies a message space $\mathcal{M}$ and the set of messages available for each score, described by $M : S \rightrightarrows \mathcal{M}$. We call this an *evidence-test-fee structure*. Our baseline model corresponds to the special case in which $M(s) = \{s\}$ for every score s. The agent first decides whether to have the asset tested and pay the testing fee ϕ_t. If he pays the testing fee, then he observes the test score s and then chooses whether to disclose each message m in $M(s)$ to the buyers. To disclose any of these messages, the agent pays the disclosure fee ϕ_d to the intermediary. If the agent does not have the asset tested or does not disclose any message, then the buyers observe the null message N. The following result shows that using evidence-test-fees structures does not improve the maximal revenue guarantee.

Proposition 6. *For every adversarial equilibrium in the game induced by an evidence-test-fee structure, there exists a test-fee structure that has an adversarial equilibrium with the same revenue for the intermediary.*

If the intermediary chose the equilibrium in addition to the test-fee structure, Proposition 6 would follow from the logic of the revelation principle. But adversarial equilibrium selection

---

[20]If costs are strictly monotone in the Blackwell order, then every robustly optimal test-fee structure is in the step-exponential-step class.

introduces a new consideration: because in the game induced by an evidence-test-fee structure the agent has more actions than in the game induced by a test-fee structure, it also has more deviations, and these deviations may exclude strategy profiles that would otherwise be equilibria. Thus, it is conceivable that under adversarial equilibrium selection some evidence-test-fee structure generates more revenue for the intermediary than any test-fee structure. We show this is not true by finding a way to prune the evidence-test-fee structure to obtain a test-fee structure in a way that does not introduce any equilibria with lower revenue to the resulting game.

7   Conclusion

When assets are traded, it is common for sellers to disclose to buyers information obtained from a third party about the value of the assets. One rationale for the existence of such information intermediaries is that their presence generates economic value by, for example, mitigating moral hazard or facilitating assortative matching. A less obvious rationale for the presence of such intermediaries, even if they provide no economic value, is that once sellers of assets have the option to obtain hard information from an intermediary, potential buyers may have an unfavorable view of assets whose sellers do not disclose favorable information.

Our paper investigates the scope of this second rationale. In a setting in which information has no social value, we study how an intermediary designs and prices evidence for a disclosure game between an asset owner and the asset market. We show that even if the equilibria of the disclosure game are chosen to minimize the intermediary's revenue, she can still guarantee herself a high revenue across equilibria. We study how she accomplishes this.

First, she uses option value as a carrot. Because the agent prefers not to have the asset tested and the market to anticipate this, the intermediary chooses a test and fees so that the agent cannot credibly refrain from having the asset tested. We show that this corresponds to the intermediary creating option value by including high score realizations in the test and charging sufficiently low testing and disclosure fees. The market then correctly expect the agent to have the asset tested, and treats non-disclosure with prejudice. In this way, the intermediary exploits the agent's commitment problem to guarantee herself revenue.

Second, the intermediary uses positive disclosure fees and noisy tests. Disclosure fees and noisy tests are closely related; were the intermediary constrained to charging only testing fees, she could optimally use tests that are fully revealing. To develop an intuition for the optimal combination of test and fees we propose a demand curve approach in which every test corresponds to a robust demand curve, and the intermediary can be thought of as choosing an optimal price on an optimal robust demand curve. The optimal disclosure fee maximizes the area of a rectangle under the robust demand curve, so the optimal robust demand curve has

a rectangular component, which leads to an exponential distribution of scores. Finally, while we have focused on two-part tariffs, which are often used in practice, it would be interesting to study a broader range of pricing structures, including disclosure fees that depend on the realized test score and fees paid by prospective buyers of the asset.

References

Aghion, Philippe and Jean Tirole. 1997. "Formal and Real Authority in Organizations." *Journal of Political Economy* 105 (1):1–29.

Albano, Gian Luigi and Alessandro Lizzeri. 2001. "Strategic Certification and Provision of Quality." *International Economic Review* 42 (1):267–283.

Bagnoli, Mark and Ted Bergstrom. 2005. "Log-Concave Probability and Its Applications." *Economic Theory* 26 (2):445–469.

Ben-Porath, Elchanan, Eddie Dekel, and Barton L. Lipman. 2018. "Disclosure and Choice." *The Review of Economic Studies* 85 (3):1471–1501.

Bergemann, Dirk and Alessandro Bonatti. 2019. "Markets for Information: An Introduction." *Annual Review of Economics* 11:85–107.

Bergemann, Dirk, Benjamin Brooks, and Stephen Morris. 2017. "First-Price Auctions With General Information Structures: Implications for Bidding and Revenue." *Econometrica* 85 (1):107–143.

Blackwell, David A and Meyer A Girshick. 1979. *Theory of Games and Statistical Decisions.* Courier Corporation.

Che, Yeon-Koo and Navin Kartik. 2009. "Opinions as Incentives." *Journal of Political Economy* 117 (5):815–860.

Condorelli, Daniele and Balázs Szentes. 2020. "Information Design in the Holdup Problem." *Journal of Political Economy* 128 (2):681–709.

De Oliveira, Henrique, Tommaso Denti, Maximilian Mihm, and Kemal Ozbek. 2017. "Rationally inattentive preferences and hidden information costs." *Theoretical Economics* 12 (2):621–654.

DeMarzo, Peter M., Ilan Kremer, and Andrzej Skrzypacz. 2019. "Test Design and Minimum Standards." *American Economic Review* 109 (6):2173–2207.

Dewatripont, Mathias and Jean Tirole. 1999. "Advocates." *Journal of Political Economy* 107 (1):1–39.

Du, Songzi. 2018. "Robust mechanisms under common valuation." *Econometrica* 86 (5):1569–1588.

Dworczak, Piotr and Alessandro Pavan. 2020. "Preparing for the Worst but Hoping for the Best: Robust (Bayesian) Persuasion." Working Paper.

Dye, Ronald A. 1985. "Strategic Accounting Choice and the Effects of Alternative Financial Reporting requirements." *Journal of Accounting Research* :544–574.

Fudenberg, Drew and Jean Tirole. 1991. *Game Theory*. MIT Press.

Gentzkow, Matthew and Emir Kamenica. 2016. "A Rothschild-Stiglitz Approach to Bayesian Persuasion." *American Economic Review* 106 (5):597–601.

Grossman, Sanford J. 1981. "The Informational Role of Warranties and Private Disclosure About Product Quality." *The Journal of Law and Economics* 24 (3):461–483.

Halac, Marina, Ilan Kremer, and Eyal Winter. 2020. "Raising Capital from Heterogeneous Investors." *American Economic Review* 110 (3):889–921.

Halac, Marina, Elliot Lipnowski, and Daniel Rappoport. 2020. "Rank Uncertainty in Organizations." Working Paper.

Harbaugh, Rick and Eric Rasmusen. 2018. "Coarse Grades: Informing the Public by Withholding Information." *American Economic Journal: Microeconomics* 10 (1):210–35.

Hoshino, Tetsuya. 2019. "Multi-Agent Persuasion: Leveraging Strategic Uncertainty." Working Paper.

Inostroza, Nicolas and Alessandro Pavan. 2020. "Persuasion in Global Games with Application to Stress Testing." Working Paper.

Jovanovic, Boyan. 1982. "Truthful Disclosure of Information." *The Bell Journal of Economics* :36–44.

Kartik, Navin, Frances Lee, and Wing Suen. 2020. "Information Validates the Prior: A Theorem on Bayesian Updating and Applications." *American Economic Review: Insights* (Forthcoming).

Kleiner, Andreas, Benny Moldovanu, and Philipp Strack. 2020. "Extreme Points and Majorization: Economic Applications." Working Paper.

Lizzeri, Alessandro. 1999. "Information Revelation and Certification Intermediaries." *The RAND Journal of Economics* :214–231.

Marinovic, Iván, Andrzej Skrzypacz, and Felipe Varas. 2018. "Dynamic Certification and Reputation for Quality." *American Economic Journal: Microeconomics* 10 (2):58–82.

Mathevet, Laurent, Jacopo Perego, and Ina Taneva. 2020. "On Information Design in Games." *Journal of Political Economy* 128 (4):1370–1404.

Matthews, Steven and Andrew Postlewaite. 1985. "Quality Testing and Disclosure." *The RAND Journal of Economics* :328–340.

Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications." *The Bell Journal of Economics* :380–391.

Ortner, Juan and Sylvain Chassang. 2018. "Making Corruption Harder: Asymmetric Information, Collusion, and Crime." *Journal of Political Economy* 126 (5):2108–2133.

Pomatto, Luciano, Philipp Strack, and Omer Tamuz. 2019. "The cost of information." Working Paper.

Roesler, Anne-Katrin and Balázs Szentes. 2017. "Buyer-Optimal Learning and Monopoly Pricing." *American Economic Review* 107 (7):2072–80.

Rothschild, Michael and Joseph E Stiglitz. 1970. "Increasing risk: I. A definition." *Journal of Economic Theory* 2 (3):225–243.

Saeedi, Maryam and Ali Shourideh. 2020. "Optimal Rating Design." Working Paper.

Shavell, Steven. 1994. "Acquisition and Disclosure of Information Prior to Sale." *The RAND Journal of Economics* :20–36.

Shishkin, Denis. 2019. "Evidence Acquisition and Voluntary Disclosure." Working Paper.

Verrecchia, Robert E. 1983. "Discretionary Disclosure." *Journal of Accounting and Economics* 5:179–194.

Zapechelnyuk, Andriy. 2020. "Optimal Quality Certification." *American Economic Review: Insights* 2 (2):161–76.

Ziegler, Gabriel. 2019. "Adversarial Bilateral Information Design." Working Paper.

A   Appendix

A.1   Proofs for Section 3

A.1.1   Proof of Proposition 1 on p. 13

Proof. Consider a test-fee structure (G, ϕ). If (w-P) is violated, then Lemma 1 shows that there is an equilibrium with zero revenue, and the proposition follows. Suppose that (w-P) holds. Using integration by parts, we can rewrite (w-P) as $\phi_t \leq \int_{\mu + \phi_d}^{\bar{\theta}} [1 - G(s)] ds$. This expression implies that any testing fee that satisfies (w-P) is at most the area above the score distribution G from $\mu + \phi_d$ to $\bar{\theta}$, shaded dark in Figure 7. The revenue from disclosure is at most $\phi_d \Pr[s \geq \underline{\theta} + \phi_d]$, shaded light in Figure 7. This is because in any equilibrium, a score strictly less than $\underline{\theta} + \phi_d$ strictly prefers to conceal. So the total revenue is at most the shaded area above G. Since G is a mean-preserving contraction of the prior distribution,

$$\mu = \int_{\underline{\theta}}^{\bar{\theta}} s \, dG(s) = \int_{\underline{\theta}}^{\bar{\theta}} [1 - G(s)] ds + \bar{\theta}$$

and therefore the area above G is equal to $R_F = \mu - \bar{\theta}$.
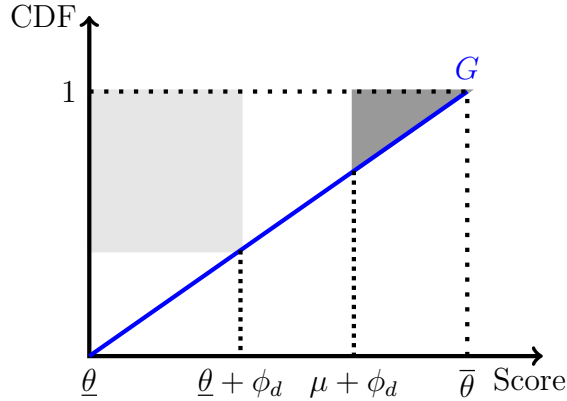
Figure 7: The revenue from testing fee fee is shaded dark. The revenue from disclosure fee is at most the area shaded light.

Now suppose that there exists an equilibrium in which revenue is $R_F - \varepsilon$. Since revenue is at most the shaded area above G and the total area above G is R_F, then the un-shaded area above G must be at most ε. In particular, the area above G from $\underline{\theta} + \phi_d$ to $\mu + \phi_d$ is at most ε. Since G is monotone,

$$\phi_t \leq \int_{\mu + \phi_d}^{\bar{\theta}} [1 - G(s)] ds \leq (\bar{\theta} - (\mu + \phi_d))(1 - G(\mu + \phi_d))$$

$$\leq \left(\frac{\bar{\theta} - (\mu + \phi_d)}{\mu - \underline{\theta}} \right) \int_{\underline{\theta} + \phi_d}^{\mu + \phi_d} [1 - G(s)] ds \leq \left(\frac{\bar{\theta} - \mu}{\mu - \underline{\theta}} \right) \varepsilon$$

28

Therefore, as ε goes to zero, the revenue from testing goes to zero as well. Thus to complete the proof we only need to show that the revenue from disclosure also goes to zero.

For ε small enough so that $\mu + \phi_d \geq \underline{\theta} + \phi_d + \sqrt{\varepsilon}$, we have

$$\varepsilon \geq \int_{\underline{\theta}+\phi_d}^{\mu+\phi_d} [1 - G(s)]ds \geq \int_{\underline{\theta}+\phi_d}^{\underline{\theta}+\phi_d+\sqrt{\varepsilon}} [1 - G(s)]ds \geq \sqrt{\varepsilon}(1 - G(\underline{\theta} + \phi_d + \sqrt{\varepsilon})),$$

where the third inequality follows since G is monotone. That is, the probability that the score is more than $\tau := \underline{\theta} + \phi_d + \sqrt{\varepsilon}$ is at most $\sqrt{\varepsilon}$. Thus if there exists an equilibrium threshold above τ, the disclosure probability in that equilibrium is at most $\sqrt{\varepsilon}$. To show that there exists an equilibrium threshold above τ, we apply Lemma 3 by showing that $E[s|s \leq \tau] > \tau - \phi_d$.

The expectation of G can be written as

$$\mu = G(\tau)E[s|s \leq \tau] + (1 - G(\tau))E[s|s > \tau] \leq G(\tau)E[s|s \leq \tau] + (1 - G(\tau))\overline{\theta}.$$

Rearranging terms yields

$$\overline{\theta} - E[s|s \leq \tau] \leq \frac{\overline{\theta} - \mu}{G(\tau)} \leq \frac{\overline{\theta} - \mu}{1 - \sqrt{\varepsilon}}.$$

Therefore since $\tau = \underline{\theta} + \phi_d + \sqrt{\varepsilon}$ and $\underline{\theta} < \mu$, for small enough ε we have[21]

$$\tau - \phi_d = \underline{\theta} + \sqrt{\varepsilon} < \frac{\mu - \overline{\theta}\sqrt{\varepsilon}}{1 - \sqrt{\varepsilon}} = \overline{\theta} - \frac{\overline{\theta} - \mu}{1 - \sqrt{\varepsilon}} \leq E[s|s \leq \tau],$$

and therefore by Lemma 3 there must exist an equilibrium threshold higher than τ.

To complete the proof, recall that $\phi_d < \overline{\theta} - \mu$, so the revenue from disclosure fee is no more than $\sqrt{\varepsilon}(\overline{\theta} - \mu)$. Thus the total revenue is bounded by $\delta(\varepsilon) \equiv (\frac{\overline{\theta}-\mu}{\mu-\underline{\theta}})\varepsilon + \sqrt{\varepsilon}(\overline{\theta} - \mu)$. $\qquad\square$

A.2 Proofs for Section 4

A.2.1 Proof of Lemma 1 on p. 15.

Proof. Consider a test-fee structure (G, ϕ) that satisfies (P). Assume towards a contradiction that there exists an equilibrium in which the agent has the asset tested with probability strictly less than 1. In this equilibrium, let p_N be the price that the agent obtains when he does not disclose his score. In any equilibrium, a agent discloses the score if $s - \phi_d > p_N$ and does not disclose his score if $s - \phi_d < p_N$. Therefore, the expected value conditional on non-disclosure is no higher than the prior expected value: $p_N \leq \mu$.

Because the agent has the asset tested with probability strictly less than 1, his equilibrium payoff equals p_N. Consider the following deviation: having the asset tested with probability 1,

---

[21]Precisely, for $\varepsilon \leq \varepsilon^* \equiv (\frac{\mu-\underline{\theta}}{1+\overline{\theta}})^2$.

and disclosing if and only if the score s satisfies $s > p_N + \phi_d$. The payoff from this deviation is $-\phi_t + \int_{p_N+\phi_d}^{\bar{\theta}} (s - \phi_d) dG + \int_{\underline{\theta}}^{p_N+\phi_d} p_N dG$ where the first term is the testing fee, the second is the payoff from disclosing a high score, and the third is the payoff from concealing a low score. Taking the difference between this deviation payoff and the equilibrium payoff of p_N yields

$$
\begin{aligned}
-\phi_t + \int_{p_N+\phi_d}^{\bar{\theta}} (s - \phi_d) dG + \int_{\underline{\theta}}^{p_N+\phi_d} p_N dG - p_N &= -\phi_t + \int_{p_N+\phi_d}^{\bar{\theta}} (s - p_N - \phi_d) dG \\
&\geq -\phi_t + \int_{\mu+\phi_d}^{\bar{\theta}} (s - \mu - \phi_d) dG \\
&> 0.
\end{aligned}
$$

where the equality follows from algebra, the first inequality follows from $p_N \leq \mu$, and the second inequality follows from (P). Therefore, the deviation is strictly profitable, and this strategy profile is not an equilibrium.

Now consider a test-fee structure (G, ϕ) that violates (P), and where

$$
\phi_t \geq \int_{\mu+\phi_d}^{\bar{\theta}} [s - (\mu + \phi_d)] dG. \tag{9}
$$

We show that now there exists an equilibrium where the agent has the asset tested with probability 0. Consider a strategy profile in which the agent's strategy involves having the asset tested with probability 0, and therefore, he has no opportunity to disclose on the path of play. Off-path, the agent discloses when obtaining a score $s \geq \mu + \phi_d$.

In this strategy profile, the market price from non-disclosure is μ. We claim that this strategy profile is an equilibrium. First, the agent's behavior is sequentially rational off-path. Second, by deviating to take the test, the agent receives an expected payoff lower than μ:

$$
\int_{\underline{\theta}}^{\mu+\phi_d} \mu\, dG + \int_{\mu+\phi_d}^{\bar{\theta}} [s - \phi_d] dG - \phi_t = \mu + \int_{\mu+\phi_d}^{\bar{\theta}} [s - (\mu + \phi_d)] dG - \phi_t \leq \mu
$$

where the equality is algebra, and the inequality follows from (9). So the agent has no profitable deviation, and the strategy profile is an equilibrium. $\qquad\square$

A.2.2   Proof of Lemma 2 on p. 16

Proof. Consider a test-fee structure (G, ϕ) where at least one of the fees is negative.

If $\phi_t - \int_{\mu+\phi_d}^{\bar{\theta}} [s - (\mu + \phi_d)] dG < 0$, from Lemma 1, the asset is tested with probability 1 in all equilibria. Suppose $\phi_d < 0$, then in the unique equilibrium all scores disclose, so $R(G, \phi) = \phi_t + \phi_d$. Let $\phi_d' = 0$ and $\phi_t' = \phi_t + (1 - G(\mu + \phi_d))\phi_d$. Then $\phi_t' - \int_{\mu+\phi_d'}^{\bar{\theta}} [s - (\mu +$

$\phi'_d)]dG \leq \phi_t - \int_{\mu+\phi_d}^{\bar{\theta}}[s - (\mu + \phi_d)]dG < 0$, so we can find $\phi''_t > \phi'_t$ and $\phi''_d = \phi'_d = 0$ such that $\phi''_t - \int_{\mu+\phi''_d}^{\bar{\theta}}[s - (\mu + \phi''_d)]dG < 0$. Then Lemma 1 implies that under (G, ϕ'') the asset is tested with probability 1 and $R(G, \phi'') = \phi''_t > \phi'_t = \phi_t + (1 - G(\mu + \phi_d))\phi_d \geq \phi_t + \phi_d = R(G, \phi)$. Now if $\phi''_t \geq 0$, we have constructed a test-fee structure with non-negative fees that generates a strictly higher revenue. If $\phi''_t < 0$, since $\phi''_d = 0$, it implies $R(G, \phi'') \leq 0$. But then we can easily construct a test-fee strcuture (G''', ϕ''') that generates a strictly positive revenue: let $G''' = F$ and $\phi'''_d = 0$. By assumption F is not degenerated (i.e., $\underline{\theta} < \bar{\theta}$), so $\int_\mu^{\bar{\theta}}(s - \mu)dF > 0$. Let $\phi'''_t \in (0, \int_\mu^{\bar{\theta}}(s - \mu)dF)$, and from Lemma 1 in all equilibria the asset is tested with probability 1 under (G''', ϕ'''), and $R(G''', \phi''') = \phi'''_t > 0$. Now suppose $\phi_t < 0$ and $\phi_d \geq 0$. Let $\phi''''_t = 0$ and $\phi''''_d = \phi_d$, then (G, ϕ'''') satisfies (P) unless $\int_{\mu+\phi_d}^{\bar{\theta}}[s - (\mu + \phi_d)]dG = 0$. So if $\int_{\mu+\phi_d}^{\bar{\theta}}[s - (\mu + \phi_d)]dG > 0$, a higher testing fee is charged while the revenue from disclosure fee is unchanged, because the disclosure decision doesn't depend on the testing fee (this can also be seen from ϕ_t does not appear in (HE)). If $\int_{\mu+\phi_d}^{\bar{\theta}}[s - (\mu + \phi_d)]dG = 0$, then $G(\mu + \phi_d) = 1$, which implies $\tau = \mu + \phi_d$ is the threshold satisfying (HE), and $R(G, \phi) = \phi_t + 0 < 0$. But then the (G''', ϕ''') constructed above gives a strictly higher revenue with non-negative fees.

If $\phi_t - \int_{\mu+\phi_d}^{\bar{\theta}}[s - (\mu + \phi_d)]dG \geq 0$, from Lemma 1, $R(G, \phi_d) \leq 0$. Then the (G''', ϕ''') constructed above gives a strictly higher revenue with non-negative fees.

$\square$

A.2.3   Proof of Lemma 3 on p. 17

The following lemma is used in our proof.

Lemma 5. *Suppose f is an increasing function defined on $[a, b] \subset \mathbb{R}$, and g is a continuous function defined on $[a, b]$. If $f(a) > g(a)$ and $f(b) < g(b)$, there exists $x^* \in (a, b)$ such that $f(x^*) = g(x^*)$. Moreover, $\bar{x} = \max\{x | f(x) = g(x)\}$ exists and $\bar{x} = \sup\{x \in [a, b] | f(x) \geq g(x)\}$.*

Proof. Define $S = \{x \in [a, b] | f(x) \geq g(x)\}$. Becase S is non-empty and bounded above by b, there exists a unique supremum $\bar{x} = \sup S$. Since f is increasing, $f(a) > g(a)$, and g is continuous, it follows that for sufficiently small $\varepsilon > 0$, $f(a+\varepsilon) > f(a) > g(a+\varepsilon)$, which implies $\bar{x} > a$. Consider an increasing sequence $x_n \in S$ such that $x_n \to \bar{x}$. Since f is an increasing function, $f(\bar{x}) \geq \limsup_{n\to\infty} f(x_n)$. Recall that $f(x_n) \geq g(x_n)$, so taking limsup for both sides yields

$$f(\bar{x}) \geq \limsup_{n\to\infty} f(x_n) \geq \limsup_{n\to\infty} g(x_n) = g(\bar{x})$$

where the equality holds because g is continuous. So $f(\bar{x}) \geq g(\bar{x})$, which also implies $\bar{x} \neq b$.

Now we prove that $f(\bar{x}) = g(\bar{x})$. Suppose towards a contradiction that $f(\bar{x}) > g(\bar{x})$. Because g is continuous, there exists $\varepsilon > 0$ such that $f(\bar{x}) > g(\bar{x} + \varepsilon)$. Since f is increasing, it then

31

follows that

$$f(\bar{x} + \varepsilon) \geq f(\bar{x}) > g(\bar{x} + \varepsilon)$$

which means $\bar{x} + \varepsilon \in S$, contradicting $\bar{x} = \sup S$. Therefore, it follows that $\bar{x} = \max\{x | f(x) = g(x)\}$. □

Proof of Lemma 3(a). A threshold τ is an equilibrium if and only if

$$E_G[s|s \leq \tau] = \tau - \phi_d.$$

Define $a(\tau) = E_G[s|s \leq \tau]$ and $b(\tau) = \tau - \phi_d$. Since $a(\tau)$ is bounded by μ and $b(\tau) \to \infty$ as $\tau \to \infty$, there exists $\bar{\tau}$ large enough such that $a(\tau) < b(\tau)$ for all $\tau \geq \bar{\tau}$. So if an equilibrium exists, the threshold τ must be in $[\underline{s}_G, \bar{\tau}]$.

We have $a(\underline{s}_G) = \underline{s}_G \geq b(\underline{s}_G)$ and $a(\bar{\tau}) \leq b(\bar{\tau})$, a is increasing and b is continuous. From Lemma 5, there exists $\tau \in [\underline{s}_G, \bar{\tau}]$ such that $a(\tau) = b(\tau)$, and $\max\{\tau | a(\tau) = b(\tau)\}$ exists. So the set of equilibrium thresholds is non-empty and a highest equilibrium threshold exists. □

Proof of Lemma 3(b). The "if" part follows directly from definition, so we prove the "only if" part. Suppose that τ is the highest threshold of (G, ϕ), so $\tau - \phi_d = E[s|s \leq \tau]$ and for all $\tau' > \tau$, $\tau' - \phi_d \neq E[s|s \leq \tau']$. We show that indeed we must have $\tau' - \phi_d > E[s|s \leq \tau']$.

Suppose for contradiction that there exists $\tau' > \tau$ such that $\tau' - \phi_d < E[s|s \leq \tau']$. Since $a(\tau') > b(\tau')$, $a(\bar{\tau}) \leq b(\bar{\tau})$, a is increasing and b is continuous, from Lemma 5, there exists $\tau^* \in (\tau', \bar{\tau}] \subset (\tau, \bar{\tau}]$ such that $a(\tau^*) = b(\tau^*)$. It implies threshold $\tau^* > \tau$ is an equilibrium, which leads to a contradiction. Therefore, for all $\tau' > \tau$, we must have $\tau' - \phi_d > E[s|s \leq \tau']$. □

Proof of Lemma 3(c). We show that the agent discloses and only discloses scores greater than the highest threshold τ is an adversarial equilibrium. Suppose there exists another equilibrium that gives the intermediary a lower revenue. In this equilibrium, let $\tilde{\tau} = p_N + \phi_d$, where p_N is the price of no disclosure. Since all scores strictly greater than $\tilde{\tau}$ must disclose and all scores strictly lower than $\tilde{\tau}$ must conceal in equilibrium, the equilibrium price satisfies

$$E[s < \tilde{\tau}] \leq p_N \leq E[s|s \leq \tilde{\tau}].$$

The intermediary's revenue from this equilibrium is $\phi_t + \phi_d[1 - G(\tilde{\tau}) + \lambda(G(\tilde{\tau}) - \sup_{s < \tau} G(s))]$, for some $\lambda \in [0, 1]$, which denotes the probability that the agent discloses when $s = \tau$. Since this equilibrium gives the intermediary a lower revenue,

$$1 - G(\tilde{\tau}) + \lambda(G(\tilde{\tau}) - \sup_{s < \tau} G(s)) < 1 - G(\tau),$$

which implies $\tilde{\tau} > \tau$. But then from part (b) the characterization of τ and the fact that $\tilde{\tau} > \tau$,

$$\tilde{\tau} - \phi_d > E[s|s \le \tilde{\tau}] \ge p_N$$

which contradicts $\tilde{\tau} = p_N + \phi_d$. $\qquad\square$

A.2.4   Proof of Lemma 4 on p. 18

We prove Lemma 4 by first proving part (c), then (a), and then (b).

We use the following lemma to prove Lemma 4(c).

Lemma 6. *If $(G^n, \tau^n) \to (G, \tau)$, and $\lim_{n\to\infty} G^n(\tau^n)$ exists, then $\lim_{n\to\infty} G^n(\tau^n) \le G(\tau)$.*

Proof. It suffices to show that for any $c > G(\tau)$, there exists N such that for $n \ge N$, $G^n(\tau^n) < c$. Because G is right-continuous, it follows that for any $c > G(\tau)$, there exists $\bar{\varepsilon} > 0$ such that

$$c > G(\tau + 2\bar{\varepsilon}) + \bar{\varepsilon} \tag{10}$$

Notice that $\tau^n \to \tau$, so there exists N_1 such that for $n > N_1$, $\tau^n \le \tau + \bar{\varepsilon}$, which implies

$$G^n(\tau^n) \le G^n(\tau + \bar{\varepsilon}) \tag{11}$$

Because weak convergence is metrized by the Levy metric,

$$L(G, F) = \inf\{\varepsilon > 0 | F(x - \varepsilon) - \varepsilon \le G(x) \le F(x + \varepsilon) + \varepsilon \text{ for all } x\}.$$

Thus, if G^n converges weakly to G, there exists N_2 such that for all $n > N_2$, $L(G, G^n) < \bar{\varepsilon}$, and hence

$$G(\tau + 2\bar{\varepsilon}) + \bar{\varepsilon} \ge G^n(\tau + \bar{\varepsilon}) \tag{12}$$

Combining (10), (11) and (12), it follows that for every $n \ge \max\{N_1, N_2\}$,

$$c > G(\tau + 2\bar{\varepsilon}) + \bar{\varepsilon} \ge G^n(\tau + \bar{\varepsilon}) \ge G^n(\tau^n).$$

$\qquad\square$

Proof of Lemma 4(c). Let us begin by verifying (w-P). We have

$$\phi_t = \lim_{n\to\infty} \phi_t^n \le \lim_{n\to\infty} \int_{\mu+\phi_d^n}^{\bar{\theta}} [s - (\mu + \phi_d^n)]dG^n(s) = \int_{\mu+\phi_d}^{\bar{\theta}} [s - (\mu + \phi_d)]dG(s)$$

where the equalities follow from taking limits and the inequality follows since (G^n, ϕ^n, τ^n) satisfies (P). So (G, ϕ) satisfies (w-P).

Now we turn to (HE). Since $\tau_n - \phi_d^n = E[s|s \le \tau^n] \le \mu$ and $\phi_d^n \to \phi_d$, we have τ^n being bounded above. On the other hand $\tau^n \ge \underline{s}_G \ge \underline{\theta}$ so τ^n is bounded below. From Bolzano-Weierstrass theorem, there exists a subsequence n_k and τ such that $\lim_{k\to\infty} \tau_{n_k} = \tau$.

For any $\tau' > \tau$, we show that $\phi_d G(\tau') \le \int_{\underline{\theta}}^{\tau'} G(s)ds$. If $\phi_d G(\tau') > \int_{\underline{\theta}}^{\tau'} G(s)ds$, there exists small enough ε such that $\phi_d G(\tau' + \varepsilon) > \int_{\underline{\theta}}^{\tau'+\varepsilon} G(s)ds$ and G is continuous at $\tau' + \varepsilon$. So $\lim_{n\to\infty} \phi_d^n G^n(\tau' + \varepsilon) = \phi_d G(\tau' + \varepsilon)$. But then we have a contradiction:

$$\phi_d G(\tau' + \varepsilon) > \int_{\underline{\theta}}^{\tau'+\varepsilon} G(s)ds = \lim_{k\to\infty} \int_{\underline{\theta}}^{\tau'+\varepsilon} G_{n_k}(s)ds \ge \lim_{k\to\infty} \phi_d^{n_k} G_{k_n}(\tau' + \varepsilon) = \phi_d G(\tau' + \varepsilon)$$

where the second inequality holds because $\tau' > \tau$ implies for large enough k, $\tau' + \varepsilon > \tau' > \tau_{n_k}$ and thus $\phi_d^{n_k} G_{n_k}(\tau' + \varepsilon) \le \int_{\underline{\theta}}^{\tau'+\varepsilon} G_{n_k}(s)ds$.

Next we show that $\phi_d G(\tau) = \int_{\underline{\theta}}^{\tau} G(s)ds$. Since G is right continuous and for all $\tau' > \tau$, $\phi_d G(\tau') \le \int_{\underline{\theta}}^{\tau'} G(s)ds$, we must have $\phi_d G(\tau) \le \int_{\underline{\theta}}^{\tau} G(s)ds$. Since $\phi_d^{n_k} G_{n_k}(\tau_{n_k}) = \int_{\underline{\theta}}^{\tau_{n_k}} G_{n_k}(s)ds$ for all k, $\lim_{k\to\infty} \phi_d^{n_k} G_{n_k}(\tau_{n_k})$ exists and equals $\lim_{k\to\infty} \int_{\underline{\theta}}^{\tau_{n_k}} G_{n_k}(s)ds = \int_{\underline{\theta}}^{\tau} G(s)ds$. Therefore

$$\phi_d G(\tau) \ge \lim_{k\to\infty} \phi_d^{n_k} G_{n_k}(\tau_{n_k}) = \int_{\underline{\theta}}^{\tau} G(s)ds$$

where the inequality holds from Lemma 6. Notice that the argument above also implies $\phi_d G(\tau) = \lim_{k\to\infty} \phi_d^{n_k} G_{n_k}(\tau_{n_k})$.

Now we have shown $(G, (\phi_d, \phi_t))$ and τ satisfies (w-P) and (w-HE), and next we consider the revenues comparison:

$$V = \lim_{n\to\infty} \phi_t^n + \phi_d^n(1 - G^n(\tau^n)) = \phi_t + \lim_{k\to\infty} \phi_d^{n_k}(1 - G_{n_k}(\tau_{n_k})) = \phi_t + \phi_d(1 - G(\tau)).$$

So $(G, (\phi_d, \phi_t))$ and τ indeed generate a revenue as the limit of the revenues generated by $(G^n, (\phi_d^n, \phi_t^n))$ and τ^n. □

We use the following lemma to prove Lemma 4(a) and Lemma 4(b).

Lemma 7. *For every test-fee structure and threshold (G, ϕ, τ) that satisfy (w-P) and (w-HE), there exists a sequence of test-fee structures and thresholds $\{(G^n, \phi^n, \tau^n)\}_{n=1,2,...}$ such that (i) for each n, (G^n, ϕ^n, τ^n) satisfy (P) and (HE), (ii) (G^n, ϕ^n) converges to (G, ϕ), and (iii) $\hat{R}(G, \phi, \tau) \le \lim_{n\to\infty} \hat{R}(G^n, \phi^n, \tau^n)$.*

Proof. Consider a test-fee structure (ϕ_t, ϕ_d, G) and τ satisfying (w-P) and (w-HE). Therefore,

$$\phi_t \le \int_{\max\{\mu+\phi_d, \bar{\theta}\}}^{\bar{\theta}} (s - \mu - \phi_d)dG(s) \text{ and}$$
$$\tau - \phi_d = E[s|s \le \tau]$$

34

$$\tau' - \phi_d \geq E[s|s \leq \tau'] \quad \text{for all } \tau' > \tau.$$

Let $\phi_d^n = \phi_d - \frac{1}{n}$ and $\phi_t^n = \phi_t - \frac{1}{n}$, then

$$\phi_t^n < \phi_t \leq \int_{\mu+\phi_d}^{\bar{\theta}} (s - \mu - \phi_d) dG(s) \leq \int_{\mu+\phi_d^n}^{\bar{\theta}} (s - \mu - \phi_d^n) dG(s),$$

so (P) is satisfied.

From Lemma 3, we know that under (ϕ_t^n, ϕ_d^n, G), the highest equilibrium threshold $\hat{\tau}_n$ exists. Moreover, since

$$\tau - \phi_d^n > \tau - \phi_d = E[s|s \leq \tau]$$
$$\tau' - \phi_d^n > \tau' - \phi_d \geq E[s|s \leq \tau'] \quad \text{for all } \tau' > \tau.$$

It implies the highest threshold $\hat{\tau}_n < \tau$. So the revenue under (ϕ_t^n, ϕ_d^n, G) is

$$\phi_t^n + \phi_d^n(1 - G(\hat{\tau}_n)) \geq \phi_t^n + \phi_d^n(1 - G(\tau)) \geq \phi_t + \phi_d(1 - G(\tau)) - \frac{2}{n}$$

where the last term goes to $\hat{R}(G, \phi, \tau)$ as $n \to \infty$. $\qquad \square$

Proof of Lemma 4(a). Consider any sequence $(G^n, (\phi_t^n, \phi_d^n))$ and τ^n satisfying (P) and (HE) and generating value $V_n \to R_M$.

The disclosure fee being non-negative ($\phi_d^n \geq 0$) implies $\phi_t^n \leq \int_\mu^{\bar{\theta}}(s - \mu) dG^n(s) \leq \bar{\theta} - \mu$. For any $\phi_d^n > \bar{\theta} - \mu$, (P) and (HE) implies $\phi_t^n < 0$ and $G(\tau^n) = 1$, and thus $V_n < 0$. So without loss of generality, we can consider $\phi_d^n \leq \bar{\theta} - \mu$. Moreover, for any $\phi_t^n < -1$, by replacing it with $\phi_t^n = -1$, (P) and (HE) are still satisfied, and V^n increases. So without loss of generality, we can restrict attention to $\phi_t^n \geq -1$.

Now we have a sequence $(G^n, (\phi_t^n, \phi_d^n))$ such that $\phi_t^{n_j} \in [-1, \bar{\theta}-\mu]$, $\phi_d^{n_j} \in [0, \bar{\theta}-\mu]$, $G^n \in \Gamma(F)$. Proposition 1 of Kleiner, Moldovanu, and Strack (2020) proves that $\Gamma(F)$ is compact, which implies that there exists a converging subsequence $(G^{n_k}, (\phi_t^{n_k}, \phi_d^{n_k})) \to (G, (\phi_t, \phi_d))$. From Lemma 4(c), we can find τ such that $(G, (\phi_t, \phi_d))$ and τ satisfies the constraints (w-P) and (w-HE) in the relaxed problem, and generates a revenue R_M. Moreover, from Lemma 7, any test-fee structure and threshold satisfying (w-P) and (w-HE) must at most generate a revenue of R_M, so $(G, (\phi_t, \phi_d))$ and τ is an optimal solution to the relaxed problem. $\qquad \square$

Proof of Lemma 4(b). By Lemma 7, there exists a sequence of test-fee structures and thresholds $\{(G^n, \phi^n, \tau^n)\}_{n=1,2,...}$ such that (i) for each n, (G^n, ϕ^n, τ^n) satisfy (P) and (HE), (ii) (G^n, ϕ^n) converges to (G, ϕ), and (iii) $\lim_{n\to\infty} \hat{R}(G^n, \phi^n, \tau^n) \geq \hat{R}(G, \phi, \tau) = R_M$. By definition of R_M, we also have $\lim_{n\to\infty} \hat{R}(G^n, \phi^n, \tau^n) \leq R_M$. Therefore, $\lim_{n\to\infty} \hat{R}(G^n, \phi^n, \tau^n) = R_M$. $\qquad \square$

A.3   Proofs for Section 5

This section contains the proofs of Propositions 2–4. The three statements of Proposition 2 build on each other in reverse order. We therefore prove Proposition 2 in reverse order, beginning with (d) and ending with (a). Proposition 3 is a corollary of one of the steps (Lemma 9) used in Proposition 2(d).

A.3.1   Proof of Proposition 2(d) on p. 19

We prove Proposition 2(d) by solving a relaxed problem where the mean-preserving contraction constraints are relaxed to requiring the score distribution to have same expectation as the prior mean:

$$\max_{(G,\phi,\tau)\in\Delta[\underline{\theta},\overline{\theta}]\times\mathbb{R}^3} \hat{R}(G,\phi,\tau) \qquad \text{s.t. (w-P), (w-HE), and } E_G[s] = E_F[s] \qquad \text{(RE)}$$

Recall that $\hat{R}(G,\phi,\tau) \equiv \phi_t + \phi_d(1 - G(\tau))$ is the revenue corresponding to a test-fee structure (G,ϕ) and a threshold τ. This is a relaxed problem because $G \in \Gamma(F)$ implies $G \in \Delta[\underline{\theta},\overline{\theta}]$ and $E_G[s] = E_F[s]$.

We use the following lemma to solve the relaxed problem. The lemma shows that in any optimal testing fee structure, the only possible score above $\mu + \phi_d$ is $\overline{\theta}$. In other words, the score distribution G is flat from $\mu + \phi_d$ to $\overline{\theta}$, with possibly a discrete jump at $\overline{\theta}$ so that $G(\overline{\theta}) = 1$.

Lemma 8. *If a test-fee structure (G,ϕ) with a weak-highest equilibrium threshold τ is an optimal solution to the relaxed problem* (RE), *then $G(s) = G(\mu + \phi_d)$ for all $s \in [\mu + \phi_d, \overline{\theta})$.*

Proof. Using integration by parts, the testing fee of an optimal test-fee structure is

$$\phi_t = \int_{\mu+\phi_d}^{\overline{\theta}} [s - (\mu + \phi_d)]dG(s) = -\int_{\mu+\phi_d}^{\overline{\theta}} G(s)ds + (\overline{\theta} - (\mu + \phi_d))$$

Thus, revenue is

$$\hat{R}(G,\phi,\tau) = -\phi_d G(\tau) - \int_{\mu+\phi_d}^{\overline{\theta}} G(s)ds + (\overline{\theta} - \mu).$$

Since τ is a weak-highest equilibrium threshold, $\phi_d G(\tau) = \int_{\underline{\theta}}^{\tau} G(s)ds$. Substituting this equality into revenue, we have

$$\hat{R}(G,\phi,\tau) = -\int_{\underline{\theta}}^{\tau} G(s)ds - \int_{\mu+\phi_d}^{\overline{\theta}} G(s)ds + (\overline{\theta} - \mu) = \int_{\tau}^{\mu+\phi_d} G(s)ds, \qquad (13)$$

where the second equality followed from the constraint that the integral of G is $1 - \mu$.

Now consider an optimal test-fee structure (G, ϕ) with a weak-highest equilibrium threshold τ, and suppose for contradiction that $G(s) > G(\mu + \phi_d)$ for some $s \in [\mu + \phi_d, \overline{\theta})$. Construct a distribution G' as follows. Let $G'(s) = \alpha G(s)$ for some α and all $s \le \mu + \phi_d$, and $G'(s) = G'(\mu + \phi_d)$ for all $s \in [\mu + \phi_d, \overline{\theta})$. By the assumption that $G(s) > G(\mu + \phi_d)$ for some $s \in [\mu + \phi_d, \overline{\theta})$, there exists $\alpha > 1$ such that the areas under G and G' are equal. Define $\phi'_d = \phi_d$, and ϕ'_t so that the upper bound on the testing fee (w-P) holds with equality for distribution G'.

We now show that τ is a weak-highest equilibrium threshold of the test-fee structure (G', ϕ'), and gives higher revenue than (G, ϕ). Indeed, for any τ' such that $\tau \le \tau' \le \mu + \phi_d$, since G' is equal to G multiplied by α, we have

$$\frac{\int_{\underline{\theta}}^{\tau'} G(s)ds}{G(\tau')} = \frac{\int_{\underline{\theta}}^{\tau} G(s)ds}{G(\tau)}.$$

Thus, from integration by parts, $\tau' - E_{G'}[s|s \le \tau'] = \tau' - E_G[s|s \le \tau'] \ge \phi_d$, with equality at $\tau' = \tau$. In addition, for all $\tau' \ge \mu + \phi_d$, since G' is flat we have

$$\frac{\int_{\underline{\theta}}^{\tau'} G'(s)ds}{G'(\tau')} \ge \frac{\int_{\underline{\theta}}^{\mu+\phi_d} G(s)ds}{G(\mu + \phi_d)} \ge \phi_d.$$

It remains to show that the revenue of (G', ϕ') is higher than that of (G, ϕ). This fact follows directly from (13), since $G' > G$ below $\mu + \phi_d$.[22] $\square$

Equipped with Lemma 8, we solve the relaxed problem (RE).

Lemma 9. *The value of* (RE) *is* $(\overline{\theta} - \mu)(1 - e^{\frac{\underline{\theta} - \mu}{\overline{\theta} - \mu}})$, *and it has a unique solution* $\phi_t^* = 0$, $\phi_d^* = \overline{\theta} - \mu$, *and*

$$G^*(s) = \begin{cases} e^{\frac{\underline{\theta}-\mu}{\overline{\theta}-\mu}} & \text{if } s \in [\underline{\theta}, \underline{\theta} + \overline{\theta} - \mu) \\ e^{\frac{s-\overline{\theta}}{\overline{\theta}-\mu}} & \text{if } s \in [\underline{\theta} + \overline{\theta} - \mu, \overline{\theta}]. \end{cases}$$

Proof. We first argue that the test-fee structure (G^*, ϕ^*) achieves the revenue bound $(\overline{\theta} - \mu)(1 - e^{\frac{\underline{\theta}-\mu}{\overline{\theta}-\mu}})$. We will later show that $(\overline{\theta} - \mu)(1 - e^{\frac{\underline{\theta}-\mu}{\overline{\theta}-\mu}})$ is an upper bound on revenue and therefore (G^*, ϕ^*) is optimal. We start by verifying that (G^*, ϕ^*) is a feasible solution.

Weak participation constraint is satisfied Since $\mu + \phi_d^* = \overline{\theta}$, the weak participation constraint (w-P) holds with equality, i.e.,

$$\int_{\mu+\phi_d^*}^{\overline{\theta}} [s - (\mu + \phi_d^*)]dG^*(s) = 0 = \phi_t^*.$$

---

[22]If $\tau = \mu + \phi_d$, $\hat{R}(G, \phi, \tau) = 0$. As we have shown in the proof of Lemma 2, the optimal revenue is strictly positive. So in an optimal test-fee structure (G, ϕ) with threshold τ, $\tau < \mu + \phi_d$.

The constraint $E_G[s] = E_F[\theta]$ is satisfied Directly from the definition of G^*, for any $\tau' \geq \underline{\theta} + \overline{\theta} - \mu$ we have

$$\int_{\underline{\theta}}^{\tau'} G^*(s)ds = (\overline{\theta} - \mu)e^{\frac{\underline{\theta} - \mu}{\overline{\theta} - \mu}} + (\overline{\theta} - \mu)e^{\frac{s - \overline{\theta}}{\overline{\theta} - \mu}}\Big|_{\underline{\theta} + \overline{\theta} - \mu}^{\tau'} = (\overline{\theta} - \mu)e^{\frac{\tau' - \overline{\theta}}{\overline{\theta} - \mu}}.$$

Thus, in particular, $\int_{\underline{\theta}}^{\overline{\theta}} G^*(s) = \overline{\theta} - \mu$ and hence, by integration by parts,

$$E_{G^*}[s] = \int_{\underline{\theta}}^{\overline{\theta}} s \, dG^*(s) = \overline{\theta} - \int_{\underline{\theta}}^{\overline{\theta}} G^*(s)ds = \mu.$$

Therefore, $E_G[s] = E_F[\theta]$.

Weak-highest equilibrium constraint is satisfied We show that $\tau = \underline{\theta} + \overline{\theta} - \mu$ is a weak-highest equilibrium threshold, i.e., it satisfies (w-HE). For any $\tau' \geq \underline{\theta} + \overline{\theta} - \mu$,

$$E[s|s \leq \tau'] = \frac{\int_{\underline{\theta}}^{\tau'} s \, dG^*(s)}{G^*(\tau')} = \frac{\tau'G(\tau') - \int_{\underline{\theta}}^{\tau'} G^*(s)ds}{G^*(\tau')} = \tau' - \frac{(\overline{\theta} - \mu)e^{\frac{\tau' - \overline{\theta}}{\overline{\theta} - \mu}}}{e^{\frac{\tau' - \overline{\theta}}{\overline{\theta} - \mu}}} = \tau' - \phi_d^*.$$

Therefore, $\tau = \underline{\theta} + \overline{\theta} - \mu$ satisfies (w-HE) and is a weak-highest equilibrium threshold.
The lemma below shows that $E[s] = \mu$ and $\tau = \underline{\theta} + \overline{\theta} - \mu$ is a weak-highest equilibrium threshold.

Thus, the revenue of the test fee structure (G^*, ϕ^*) with weak-highest equilibrium threshold $\tau = \underline{\theta} + \overline{\theta} - \mu$ is

$$\phi_t^* + \phi_d^*(1 - G(\underline{\theta} + \overline{\theta} - \mu)) = 0 + (\overline{\theta} - \mu)(1 - e^{\frac{\underline{\theta} - \mu}{\overline{\theta} - \mu}}).$$

Upper bound on revenue It remains to show that the revenue of any test-fee structure is at most $(\overline{\theta} - \mu)(1 - e^{\frac{\underline{\theta} - \mu}{\overline{\theta} - \mu}})$.

For any τ' such that $\tau \leq \tau' \leq \mu + \phi_d$, the total area under G is

$$\overline{\theta} - \mu = \int_{\underline{\theta}}^{\tau'} G(s)ds + \int_{\tau'}^{\overline{\theta}} G(s)ds$$

$$\leq \phi_d e^{\frac{1}{\phi_d}(\tau' - \tau)}G(\tau) + \int_{\tau'}^{\overline{\theta}} G(s)ds$$

$$\leq \phi_d e^{\frac{1}{\phi_d}(\tau' - \tau)}G(\tau) + (\overline{\theta} - \tau')G(\mu + \phi_d),$$

38

where the first inequality followed from Lemma 11, and the second inequality followed since $G(s) \leq G(\mu + \phi_d)$ for all $s \leq \mu + \phi_d$, and $G(s) = G(\mu + \phi_d)$ for all $s \geq \mu + \phi_d$ by Lemma 8. Let $\tau' = \phi_d + \frac{\mu - \bar{\theta}(1 - G(\mu + \phi_d))}{G(\mu + \phi_d)}$. Notice that $\tau' \in [\tau, \mu + \phi_d]$ because from the definition of τ, $\tau = \phi_d + E[s|s \leq \tau] \leq \phi_d + E[s|s \leq \mu + \phi_d] = \tau'$, and $\mu < \bar{\theta}$ implies $\frac{\mu - \bar{\theta}(1 - G(\mu + \phi_d))}{G(\mu + \phi_d)} < \mu$. Therefore, by the above discussion we have

$$
\begin{aligned}
G(\tau) &\geq \frac{1}{\phi_d} e^{-\frac{1}{\phi_d}(\tau' - \tau)} ((\bar{\theta} - \mu) - (\bar{\theta} - \tau')G(\mu + \phi_d)) \\
&= \frac{1}{\phi_d} e^{-\frac{1}{\phi_d}(\tau' - \tau)} ((\bar{\theta} - \mu) - (\frac{\bar{\theta} - \mu}{G(\mu + \phi_d)} - \phi_d)G(\mu + \phi_d)) \\
&= e^{-\frac{1}{\phi_d}(\tau' - \tau)} G(\mu + \phi_d) \\
&\geq e^{-\frac{1}{\phi_d}(\bar{\theta} - \underline{\theta} - \frac{\bar{\theta} - \mu}{G(\mu + \phi_d)})} G(\mu + \phi_d),
\end{aligned}
\tag{14}
$$

where the last inequality followed since $\underline{\theta} + \phi_d \leq \tau$.

We now use (14) and Lemma 8 to bound revenue. Since by Lemma 8, the distribution is flat above $\mu + \phi_d$, the testing fee is

$$
\phi_t = \int_{\mu + \phi_d}^{\bar{\theta}} [s - (\mu + \phi_d)]dG(s) = (1 - G(\mu + \phi_d))(\bar{\theta} - (\mu + \phi_d))
$$

Thus revenue is

$$
\hat{R}(G, \phi, \tau) \leq \phi_d \left[1 - e^{-\frac{1}{\phi_d}(\bar{\theta} - \underline{\theta} - \frac{\bar{\theta} - \mu}{G(\mu + \phi_d)})} G(\mu + \phi_d) \right] + (1 - G(\mu + \phi_d))(\bar{\theta} - (\mu + \phi_d)).
$$

The above expression is non-decreasing in ϕ_d. Since $\phi_d \leq \bar{\theta} - \mu$, an upper bound on revenue is obtained by substituting $\phi_d = \bar{\theta} - \mu$ into the above expression, which yields

$$
(\bar{\theta} - \mu)(1 - e^{-\frac{\bar{\theta} - \underline{\theta}}{\bar{\theta} - \mu} + \frac{1}{G(\mu + \phi_d)}} G(\mu + \phi_d)).
$$

This expression is non-decreasing in $G(\mu + \phi_d)$. Since $G(\mu + \phi_d) \leq 1$, an upper bound on revenue is obtained by substituting $G(\mu + \phi_d) = 1$ into the above expression, which yields $(\bar{\theta} - \mu)(1 - e^{\frac{\underline{\theta} - \mu}{\bar{\theta} - \mu}})$, completing the proof. □

Proof of Proposition 2(d). The solution to the relaxed problem in Lemma 9 is $(\bar{\theta} - \mu)(1 - e^{\frac{\underline{\theta} - \mu}{\bar{\theta} - \mu}})$. Therefore $(\bar{\theta} - \mu)(1 - e^{\frac{\underline{\theta} - \mu}{\bar{\theta} - \mu}})$ is an upper bound on the revenue of any test-fee structure. Moreover, if the support of the prior is binary (i.e. $\{\underline{\theta}, \bar{\theta}\}$), $G \in \Gamma(F)$ is equivalent to $G \in \Delta[\underline{\theta}, \bar{\theta}]$ and $E_G[s] = E_F[s]$, so the bound is attained in this case. □

A.3.2   Proof of Proposition 2(c) on p. 19

Proof of Proposition 2(c). In a robustly optimal test-fee structure, the constraint (w-P) must bind, otherwise the intermediary can strictly increases the revenue by increasing ϕ_t, so $\phi_t = \int_{\mu+\phi_d}^{\bar{\theta}}[s-(\mu+\phi_d)]dG$. Notice that $\int_{\mu+\phi_d}^{\bar{\theta}}[s-(\mu+\phi_d)]dG > 0$ iff $G(\mu+\phi_d) < 1$, which then implies the result.

$\square$

A.3.3   Proof of Proposition 2(b) on p. 19

We prove Proposition 2(b) in two steps. We first identify test-fee structures that are robustly optimal among ones with zero disclosure fees. We show that it is in fact robustly optimal to use a fully revealing test. We will then show that such a test-fee structure can be improved upon using a small disclosure fee.

Lemma 10. *The test-fee structure (F,ϕ) with $\phi_t = \int_{\mu}^{\bar{\theta}}[s-\mu]dF(s)$ and $\phi_d = 0$ is robustly optimal among all test-fee structures with zero disclosure fee.*

Proof. Consider a test-fee structure with zero disclosure fee. The revenue is

$$\int_{\mu}^{\bar{\theta}}(s-\mu)dG(s) = (\bar{\theta}-\mu) - \int_{\mu}^{\bar{\theta}}G(s)ds.$$

Thus revenue is maximized by minimizing $\int_{\mu}^{\bar{\theta}}G(s)ds$.

The MPC constraints require that $\int_{\tau}^{\bar{\theta}}G(s)ds \geq \int_{\tau}^{\bar{\theta}}F(s)ds$ for all $\tau \in [\underline{\theta},\bar{\theta}]$. Specifically at $\tau = \mu$, we have

$$\int_{\mu}^{\bar{\theta}}G(s)ds \geq \int_{\mu}^{\bar{\theta}}F(s)ds.$$

Thus the optimal revenue is at most

$$(\bar{\theta}-\mu) - \int_{\mu}^{\bar{\theta}}F(s)ds$$

which is obtained by setting $G = F$. $\square$

Equipped with Lemma 10, we prove Proposition 2(b).

Proof of Proposition 2(b). Consider the revenue from charging a positive disclosure fee ϕ'_d, and reducing the testing fee to ϕ'_t so that the upper bound on testing fee (w-P) binds, ensuring participation with probability one. This change involves a loss and a gain in revenue. The loss is from lower testing fee, which is the difference between $\bar{\theta} - \mu - \int_{\mu}^{\bar{\theta}}G(s)ds$ and $\bar{\theta} - (\mu+\phi'_d) - \int_{\mu+\phi'_d}^{\bar{\theta}}G(s)ds$. By algebra, the loss is at most $\phi'_d(1 - G(\mu))$.

The gain from charging a positive disclosure fee is $\phi_d'(1 - G(\tau))$, where τ is a weak-highest equilibrium threshold. Notice that since $\underline{\theta}$ is in the support of the distribution, we have $G(\tau') > 0$ for all $\tau' > \underline{\theta}$. Thus, by integration by parts we have

$$\tau' - E[s|s \leq \tau'] = \frac{\int_{\underline{\theta}}^{\tau'} G(s)ds}{G(\tau')} > 0.$$

for all $\tau' > \underline{\theta}$. Therefore, for any $\tau' > \underline{\theta}$, there exists a small enough disclosure fee such that any weak-highest equilibrium threshold is below τ'. In particular, for a small enough disclosure fee, any weak-highest equilibrium threshold τ is below μ and satisfies $G(\tau) \leq G(\mu)$. As a result, for small enough disclosure fee, the gain is at least the loss. Further, if $G(s) > G(\mu)$ for some $s \in [\mu, \overline{\theta})$, then there exists ϕ_d such that the loss is strictly less than $\phi_d'(1 - G(\mu))$. This is the case if G is non-binary. As a result, for a non-binary G, there exists ϕ', $\phi_d' > 0$, such that the revenue of (G, ϕ') is strictly larger than the revenue of (G, ϕ).

Suppose first that the prior F is non-binary. By Lemma 10, the test-fee structure (F, ϕ) where $\phi_d = 0$ is optimal among all test-fee structures with zero disclosure fee. The argument above shows that there exists ϕ' with positive disclosure fee and such that the revenue of (F, ϕ') is strictly higher than that of (F, ϕ), and hence of any test-fee structure with zero disclosure fee.

Now suppose that the prior F is binary. With binary support, the mean-preserving contraction constraints become $E_G[s] = E_F[\theta]$. Therefore, by Lemma 9, the optimal revenue is $(\overline{\theta} - \mu)(1 - e^{\frac{\underline{\theta} - \mu}{\overline{\theta} - \mu}})$. By Lemma 10, the optimal revenue among test-fee structures with zero disclosure fee is obtained by full revelation. The revenue is the testing fee

$$\overline{\theta} - \mu - \int_{\mu}^{\overline{\theta}} F(s)ds = (\overline{\theta} - \mu) - \frac{(\overline{\theta} - \mu)^2}{\overline{\theta} - \underline{\theta}} = \frac{(\overline{\theta} - \mu)(\mu - \underline{\theta})}{\overline{\theta} - \underline{\theta}},$$

which is strictly less than the optimal revenue $(\overline{\theta} - \mu)(1 - e^{\frac{\underline{\theta} - \mu}{\overline{\theta} - \mu}})$ for all $\mu \in (\underline{\theta}, \overline{\theta})$. Since $\overline{\theta}$ is in the support of the distribution, it ensures that $\mu \in (\underline{\theta}, \overline{\theta})$. □

A.3.4   Proof of Proposition 2(a) on p. 19

We use two lemmas in the proof of Proposition 2(a). The first lemma bounds the rate at which the integral of a score distribution can grow given the (w-HE) constraint.

Lemma 11. *Suppose that $\phi_d > 0$. Let τ be a weak-highest equilibrium threshold. Then for any τ_a and τ_b where $\tau \leq \tau_a \leq \tau_b$,*

$$\int_{\underline{\theta}}^{\tau_b} G(s)ds \leq e^{\frac{1}{\phi_d}(\tau_b - \tau_a)} \int_{\underline{\theta}}^{\tau_a} G(s)ds,$$

with equality if and only if (w-HE) *holds with equality for all thresholds in* $[\tau_a, \tau_b]$.

Proof. Using integration by parts, the weak-highest equilibrium threshold constraint (w-HE) can be written as

$$\phi_d \leq \tau' - E[s|s \leq \tau'] = \frac{\int_{\underline{\theta}}^{\tau'} G(s)ds}{G(\tau')},$$

for all $\tau' \geq \tau$, with equality at $\tau' = \tau$. The right hand side is the inverse of the derivative of $\ln(\int_{\underline{\theta}}^{\tau'} G(s)ds)$ with respect to τ'. Thus,

$$\frac{d}{d\tau'}(\ln(\int_{\underline{\theta}}^{\tau'} G(s)ds)) \leq \frac{1}{\phi_d}. \tag{15}$$

Integrating from τ_a to τ_b,

$$\ln(\int_{\underline{\theta}}^{\tau_b} G(s)ds) - \ln(\int_{\underline{\theta}}^{\tau_a} G(s)ds) \leq \frac{1}{\phi_d}(\tau_b - \tau_a).$$

Raising both sides to the power of e,

$$\frac{\int_{\underline{\theta}}^{\tau_b} G(s)ds}{\int_{\underline{\theta}}^{\tau_a} G(s)ds} \leq e^{\frac{1}{\phi_d}(\tau_b - \tau_a)},$$

with equality if and only if (w-HE) holds with equality for all thresholds in $[\tau_a, \tau_b]$. □

The following lemma is the main step in the proof of Proposition 2(a). It shows that for any test-fee structure, there exists a test-fee structure that is in the step-exponential-step class and has a weakly higher revenue guarantee. The lemma further establishes that *any* robustly optimal test must be exponential over an interval.

Lemma 12. *For any test-fee structure* (G, ϕ) *with* $\phi_d > 0$ *and a corresponding weak-highest threshold* τ_1, *there exists a mean-preserving contraction* G' *of* G, *a fee structure* ϕ', *and a threshold* τ_1' *such that the test-fee structure* (G', ϕ') *is in the step-exponential-step class,* τ_1' *is a weak-highest threshold and the robust revenue of* (G', ϕ', τ_1') *is at least the robust revenue of* (G, ϕ, τ_1). *Further, if* (G, ϕ) *is robustly optimal and has weak-highest equilibrium threshold* τ_1, *then* $\phi_d > 0$ *and there exists a threshold* $\tau_2 \in [\tau_1, \mu + \phi_d]$, *such that* G *is exponential from* τ_1 *to* τ_2, *i.e.,*

$$G(s) = G(\tau_1)e^{\frac{1}{\phi_d}(s - \tau_1)}, \forall s \in [\tau_1, \tau_2],$$

42

and is flat from τ_2 to $\mu + \phi_d$, i.e.,

$$G(s) = G(\mu + \phi_d), \forall s \in [\tau_2, \mu + \phi_d].$$

Proof. Given a test-fee structure (G, ϕ) with a weak-highest threshold τ_1, clearly by definition of τ_1, we have $\tau_1 \leq \mu + \phi_d$. Moreover, if $\tau_1 = \mu + \phi_d$, the intermediary receives zero revenue, which is never optimal. So we consider $\tau_1 < \mu + \phi_d$.

We first consider the case $G(\mu + \phi_d) = G(\tau_1)$, where τ_1 is a weak-highest equilibrium threshold under (G, ϕ). Define $\tau_a = E_G[s|s \leq \tau_1]$, and $\tau_b = E_G[s|s > \mu + \phi_d]$. From the definition of τ_1, we have $\tau_a = E_G[s|s \leq \tau_1] = \tau_1 - \phi_d$. The mean constraint requires $G(\tau_1)\tau_a + (1 - G(\tau_1))\tau_b = \mu$, which implies $G(\tau_1) = \frac{\tau_b - \mu}{\tau_b - \tau_a}$.

Given the test G, and to satisfy (w-P), the highest testing fee the intermediary can charge is $\phi_t = \int_{\mu+\phi_d}^{1} [s - (\mu + \phi_d)]dG(s)$. The disclosure fee being charged under equilibrium threshold τ_1 is $\phi_d(1 - G(\tau_1))$. So the intermediary's robust revenue bound under test G is

$$
\begin{aligned}
\int_{\mu+\phi_d}^{\overline{\theta}} [s - (\mu + \phi_d)]dG(s) + \phi_d(1 - G(\tau_1)) &= \overline{\theta} - \mu - \phi_d - \int_{\mu+\phi_d}^{\overline{\theta}} G(s)ds + \phi_d - \phi_d G(\tau_1) \\
&= \int_{\underline{\theta}}^{\mu+\phi_d} G(s)ds - \int_{\underline{\theta}}^{\tau_1} G(s)ds \\
&= \int_{\tau_1}^{\mu+\phi_d} G(s)ds \\
&= G(\tau_1)(\mu + \phi_d - \tau_1) \\
&= \frac{(\tau_b - \mu)(\mu - \tau_a)}{\tau_b - \tau_a}.
\end{aligned}
$$

Now we construct the following G', which is in the class of step-exponential-step with the last step being degenerated:

$$
G'(s) = \begin{cases}
0 & \text{if } s \in [\underline{\theta}, \tau_a) \\
e^{\frac{\tau_a - \mu}{\tau_b - \mu}} & \text{if } s \in [\tau_a, \tau_a + \tau_b - \mu) \\
e^{\frac{s - \tau_b}{\tau_b - \mu}} & \text{if } s \in [\tau_a + \tau_b - \mu, \tau_b] \\
1 & \text{if } s \in [\tau_b, \overline{\theta}].
\end{cases}
$$

It can be verified that G' is a mean-preserving contraction of G, and $\tau_1' = \tau_a + \tau_b - \mu$ is a weak-highest equilibrium threshold under test-fee structure (G', ϕ_d', ϕ_t') with $\phi_d' = \tau_b - \mu$ and $\phi_t' = 0$. In this equilibrium, the intermediary's revenue is $(\tau_b - \mu)(1 - e^{\frac{\tau_a - \mu}{\tau_b - \mu}}) > \frac{(\tau_b - \mu)(\mu - \tau_a)}{\tau_b - \tau_a}$.

Now consider the other case $G(\mu + \phi_d) > G(\tau_1)$. We construct a class of distributions G_α parametrized by $\alpha \in [0, 1]$.

$$
G_\alpha(s) = \begin{cases} \alpha G(s) & \text{if } s \leq \tau_1 \\ \min\{\alpha G(\tau_1) e^{\frac{1}{\phi_d}(s-\tau_1)}, G(\mu+\phi_d)\} & \text{if } \tau_1 < s < \mu + \phi_d \\ G(s) & \text{if } s \geq \mu + \phi_d. \end{cases} \tag{16}
$$

Notice that G_α is well-defined since $\phi_d > 0$, and is monotone non-decreasing and is between 0 and 1. Therefore, G_α is a distribution.

We first show that there exists $\alpha \leq 1$ such that the areas under G_α and G are equal. We do so using a continuity argument. Let $\tau_2(\alpha)$ be the lowest score $\tau' > \tau_1$ such that $G_\alpha(\tau') = G_\alpha(\mu + \phi_d)$. Consider $\alpha = 1$. We have

$$
\int_{\underline{\theta}}^{\tau_2(\alpha)} G_1(s)ds = e^{\frac{1}{\phi_d}(\tau_2(\alpha)-\tau_1)} \int_{\underline{\theta}}^{\tau_1} G_1(s)ds
$$

$$
= e^{\frac{1}{\phi_d}(\tau_2(\alpha)-\tau_1)} \int_{\underline{\theta}}^{\tau_1} G(s)ds
$$

$$
\geq \int_{\underline{\theta}}^{\tau_2(\alpha)} G(s)ds,
$$

where the inequality followed from Lemma 11. As a result, since G_1 is weakly higher than G for all scores above τ_2, we have $\int_{\underline{\theta}}^{\overline{\theta}} G_1(s)ds \geq \int_{\underline{\theta}}^{\overline{\theta}} G(s)ds$. Further,

$$
\int_{\underline{\theta}}^{\mu+\phi_d} G_0(s)ds = 0 \leq \int_{\underline{\theta}}^{\mu+\phi_d} G(s)ds.
$$

As a result, since G_0 and G are equal above $\mu + \phi_d$, we have $\int_{\underline{\theta}}^{\overline{\theta}} G_0(s)ds \leq \int_{\underline{\theta}}^{\overline{\theta}} G(s)ds$. The area under G_α increases continuously in α. Therefore, there exists $\alpha \leq 1$ such that the areas under G_α and G are equal. For the rest of the proof fix such an α.

We now show that G_α is a mean-preserving contraction of G. Since $\alpha \leq 1$, the area under G_α up to any threshold $\tau' \leq \tau$ is weakly less than that for G. For $\tau' \geq \tau_2(\alpha)$, since the total area under G and G_α are equal and G_α is weakly higher than G above $\mu + \phi_d$, we must have $\int_{\underline{\theta}}^{\tau'} G_\alpha(s)ds \leq \int_{\underline{\theta}}^{\tau'} G(s)ds$. Finally, for any τ' such that $\tau \leq \tau' \leq \tau_2(\alpha)$ we have

$$
\int_{\underline{\theta}}^{\tau'} G_\alpha(s)ds = e^{-\frac{1}{\phi_d}(\tau_2-\tau')} \int_{\underline{\theta}}^{\tau_2} G_\alpha(s)ds \leq e^{-\frac{1}{\phi_d}(\tau_2-\tau')} \int_{\underline{\theta}}^{\tau_2} G(s)ds \leq \int_{\underline{\theta}}^{\tau'} G(s)ds,
$$

where the inequality followed from Lemma 11. Thus, G_α is a mean-preserving contraction of G.

Test-fee structure (G_α, ϕ) has weakly higher revenue than (G, ϕ). Since the two distributions are equal above $\mu + \phi_d$, the (w-P) constraint is satisfied for (G_α, ϕ). Further, τ is a weak-highest

equilibrium threshold for (G_α, ϕ). Since $\alpha \leq 1$, the disclosure probability in (G_α, ϕ) is weakly higher than (G, ϕ). In fact, if $\alpha < 1$, the robust revenue of (G_α, ϕ) is more than that of (G, ϕ).

To see the first statement of the lemma, consider a score distribution G' that is equal to G_α except that it pools the scores below τ_1 and also pools the scores above τ_2. Formally,

$$G'(s) = \begin{cases} G_\alpha(\tau_1) & \text{if } s \in [\tau_0, \tau_1], \\ G_\alpha(s) & \text{if } s \in [\tau_1, \tau_2], \\ 1 & \text{if } s \in [\tau_3, \overline{\theta}], \end{cases}$$

where τ_0 and τ_3 are such that the areas under G' and G_α are equal. Notice that the test-fee structure (G', ϕ) is in the step-exponential-step class, with a non-degenerated exponential part because $\tau_2 > \tau_1$. Further distribution G' is a mean-preserving contraction of G_α and therefore G. Finally, the robust revenue of the test-fee structure (G', ϕ) with weak-highest equilibrium threshold τ_1 is equal to that of (G_α, ϕ), and therefore at least that of (G, ϕ). This establishes the first statement of the lemma.

To see the second statement, notice that if (G, ϕ) is optimal, then by Proposition 2(b), $\phi_d > 0$. Further, it must be that $\alpha = 1$, as otherwise the revenue of (G_α, ϕ) is strictly higher. Therefore, $G = G_1$, which means that G is exponential from τ_1 to τ_2, and flat from τ_2 to $\mu + \phi_d$, as claimed. $\qquad\square$

Proof of Proposition 2(a). Consider any robustly optimal test-fee structure (G, ϕ). By Proposition 2(b), $\phi_d > 0$. By Lemma 12, the robust revenue of (G, ϕ) is at most the robust revenue of some step-exponential-step test-fee structure (G', ϕ) where G' is a mean-preserving contraction of G. Therefore, (G', ϕ) is robustly optimal.

Next we show that in fact there is a unique equilibrium under any robustly optimal step-exponential-step test-fee structure with a slightly lower disclosure fee. Let (G, ϕ) be a robustly optimal step-exponential-step test-fee structure:

$$G(s) = \begin{cases} g & \text{if } s \in [\tau_0, \tau_1) \\ ge^{(s-\tau_1)/(\tau_1 - \tau_0)} & \text{if } s \in [\tau_1, \tau_2), \\ 1 & \text{if } s \geq \tau_3, \end{cases}$$

the disclosure fee is $\phi_d = \tau_1 - \tau_0$, and the testing fee is $\phi_t = (1 - ge^{(\tau_2 - \tau_1)/(\tau_1 - \tau_0)})(\tau_3 - (\mu + \phi_d))$. Notice that for any $\tau \in [\tau_1, \tau_2]$,

$$E_G[s|s \leq \tau] = \tau - \frac{\int_\theta^\tau G(s)ds}{G(\tau)} = \tau - \frac{g(\tau_1 - \tau_0)e^{\frac{\tau - \tau_1}{\tau_1 - \tau_0}}}{ge^{\frac{\tau - \tau_1}{\tau_1 - \tau_0}}} = \tau - \tau_1 - \tau_0 = \tau - \phi_d.$$

For $\tau > \tau_2$, $E_G[s|s \leq \tau] \leq \tau - \phi_d$. For $\tau \in [\tau_0, \tau_1)$, $E_G[s|s \leq \tau] = \tau_0 < \tau - \phi_d$.

Consider a test-fee structure (G, ϕ') with $\phi_t' = \phi_t$ and $\phi_d' = \phi_d - \varepsilon$ for $\varepsilon < \tau_1 - \tau_0$, then there exists a unique threshold $\tau' = \tau_1 - \varepsilon$ satisfying (HE). Moreover, since (G, ϕ) satisfies (w-P), (G, ϕ') satisfies (P) because $\int_{\mu+\phi_d}^{\bar{\theta}}[s - (\mu + \phi_d)]dG < \int_{\mu+\phi_d-\varepsilon}^{\bar{\theta}}[s - (\mu + \phi_d - \varepsilon)]dG$. Therefore, in the unique equilibrium under (G, ϕ'), the asset is tested with probability 1 and the agent discloses all the scores above τ_1. □

A.3.5 Proof of Proposition 3 on p. 22

Proof. With binary support, the mean-preserving contraction constraints become $E_G[s] = E_F[\theta]$. Therefore, the proposition follows from Lemma 9. □

A.3.6 Proof of Proposition 4 on p. 22

We first simplify the step-exponential-step distribution identified in Section 5 in the case where the testing fee is zero.

Lemma 13. *Consider an optimal step-exponential-step test-fee structure defined in (5), (6), and (7). If $\phi_t^* = 0$, then $\tau_2 = \mu + \phi_d^*$ and $G^*(\tau_2) = 1$. Further, if the prior distribution is log-concave, then the mean-preserving constraints are slack, $\int_{\underline{\theta}}^{\tau'} G^*(s)ds < \int_{\underline{\theta}}^{\tau'} F(s)ds$, for all τ' from τ_1 to $\bar{\theta}$.*

Proof. For the optimal testing fee to be zero, the area above G^* from $\mu + \phi_d^*$ to $\bar{\theta}$ must be zero, and hence $G^*(\mu + \phi_d) = 1$. From Lemma 12, $G^*(\tau_2) = 1$.

The constraint that the area under the score distribution G^* is $\bar{\theta} - \mu$ can be written as

$$\bar{\theta} - \mu = \int_{\underline{\theta}}^{\tau_1} G^*(s)ds + \int_{\tau_1}^{\tau_2} ge^{\frac{s-\tau_1}{\tau_1-\tau_0}}ds + (\bar{\theta} - \tau_2)$$
$$= g(\tau_1 - \tau_0) + (\tau_1 - \tau_0)ge^{\frac{s-\tau_1}{\tau_1-\tau_0}}|_{\tau_1}^{\tau_2} + (\bar{\theta} - \tau_2)$$
$$= (\tau_1 - \tau_0) + \bar{\theta} - \tau_2.$$

We must therefore have $\tau_2 = \mu + (\tau_1 - \tau_0) = \mu + \phi_d^*$.

We now show that the MPC constraints must be slack on interval $[\tau_1, 1]$. Notice that the constraints are slack on $[\tau_2, \bar{\theta}]$ since G is equal to 1 but $F(\tau')$ is strictly less than one for all $\tau' < 1$. It thus remains to show that the constraints are slack on some interval $[\tau_1', \tau_2]$. Notice that log-concavity implies continuity in the interior, which will be used in the later arguments.

We first claim that $F(\tau_1) \geq G(\tau_1) = e^{-\frac{\mu-\tau_0}{\tau_1-\tau_0}}$. Suppose $F(\tau_1) < G(\tau_1)$, then $F(x) < G(x)$ for any $x \in [\tau_0, \tau_1)$, which implies the constraint must be slack on the interval because $\int_{\underline{\theta}}^{x} F(s)ds = \int_{\underline{\theta}}^{\tau_1} F(s)ds - \int_{x}^{\tau_1} F(s)ds > \int_{\underline{\theta}}^{\tau_1} G(s)ds - \int_{x}^{\tau_1} G(s)ds = \int_{\underline{\theta}}^{x} G(s)ds$. From the right-continuity of F, the constraint is also slack at τ_1. If this is the case, we can construct a new distribution

46

parameterized by $\tau_1' = \tau_1 - \varepsilon$ and $\tau_0' = \tau_0 - \varepsilon$, the distribution on $[\tau_1, \bar{\theta}]$ doesn't change so all the constraints on $[\tau_1, \bar{\theta}]$ are still satisfied. Also the original constraints on $[\underline{\theta}, \tau_1]$ are slack, they are still satisfied for small ε. By charging the same disclosure fee $\phi_d = \tau_1 - \tau_0$, this new distribution can induce a higher disclosure probability, which contradicts to G being optimal.

If $F(\tau_1) = G(\tau_1)$, the same argument goes through if the constructed distribution doesn't violate the mean-preserving constraints. If the constraint is slack at τ_1, then all the constraints to the left are slack, so the constructed distribution is still a profitable deviation. If the constraint binds at τ_1, the right derivative of F at τ_1 must be greater than the right derivative of G. Also the left derivative of F must be greater than G from log-concavity. So the local change of the distribution doesn't violate any constraint either, which leads to a profitable deviation.

So we must have $F(\tau_1) > G(\tau_1)$, and the constraint at τ_1 is slack at τ_1 due to the continuity of F. Moreover, from the log-concavity of F, and $F(\tau_1) > G(\tau_1)$, $F(\tau_2) < G(\tau_2)$, F crosses the exponential part of G from above once and only once. To see this, notice that $\log(F)$ is concave and $\log(G)$ is linear on $[\tau_1, \tau_2]$, and a concave function can only cross a linear function from above once. Let x^* denote the intersection point, we have $F(x) > G(x)$ for $x \in [\tau_1, x^*)$ and $F(x) < G(x)$ for $x \in (x^*, \tau_2]$. Notice that the constraint is slack at τ_2, which means $\int_{\underline{\theta}}^{\tau_2} G(s)ds \le \int_{\underline{\theta}}^{\tau_2} F(s)ds$, then for any $x \in [x^*, \tau_2)$, the constraint is also slack because $\int_{\underline{\theta}}^{x} G(s)ds = \int_{\underline{\theta}}^{\tau_2} G(s)ds - \int_{x}^{\tau_2} G(s)ds < \int_{\underline{\theta}}^{\tau_2} GFds - \int_{x}^{\tau_2} F(s)ds = \int_{\underline{\theta}}^{x} F(s)ds$. Similarly the constraint is slack at τ_1 implies the constraint is also slack at any $x \in (\tau_1, x^*]$.

Therefore, all MPC constraints are slack on $[\tau_1, \bar{\theta})$. $\qquad\square$

Given Lemma 13 we now prove Proposition 4.

Proof of Proposition 4. Consider the class of step-exponential-step distributions defined in (5) where $\tau_3 = \bar{\theta}$. That is,

$$G(s) = \begin{cases} g & \text{if } s = \tau_0 \\ ge^{(s-\tau_1)/(\tau_1-\tau_0)} & \text{if } s \in [\tau_1, \tau_2] \\ 1 & \text{if } s = \bar{\theta}. \end{cases}$$

A degree of freedom is removed given the constraint that the area under G is $\bar{\theta} - \mu$. Notice that the distribution with zero testing fee identified in Lemma 13 is a special case where $\tau_2 = \mu + (\tau_1 - \tau_0)$. Consider such a distribution. We show that reducing τ_1 by a small amount increases the revenue, and does not violate any of the constraints.

The intermediary's revenue from charging disclosure fee $\phi_d = \tau_1 - \tau_0$ and the highest possible testing fee $\phi_t = \int_{\mu+\phi_d}^{\bar{\theta}} (s - \mu - c)dG(s)$ is

$$\phi_t + \phi_d(1 - G(\tau_1)) = \int_{\mu+x_1-x_0}^{\bar{\theta}} [s - \mu - (\tau_1 - \tau_0)]dG(s) + (\tau_1 - \tau_0)(1 - G(\tau_1))$$

47

$$=\bar{\theta} - \mu - (\tau_1 - \tau_0) - \int_{\mu+\tau_1-\tau_0}^{\bar{\theta}} G(s)ds + \tau_1 - \tau_0 - \int_{\underline{\theta}}^{\tau_1} G(s)ds$$

$$= \int_{\tau_1}^{\mu+\tau_1-\tau_0} G(s)ds$$

$$= ge^{\frac{\tau_2-\tau_1}{\tau_1-\tau_0}}(2(\tau_1 - \tau_0) + \mu - \tau_2) - (\tau_1 - \tau_0)g.$$

Consider the change in revenue from decreasing τ_1, evaluated at $\tau_2 = \mu + (\tau_1 - \tau_0)$ and $g = e^{-\frac{\mu-\tau_0}{\tau_1-\tau_0}}$,

$$\frac{d}{d\tau_1}(\phi_t + \phi_d(1 - G(\tau_1))) = (\tau_1 - \tau_0)\frac{d}{d\tau_1}\left(\frac{\tau_2 - \tau_1}{\tau_1 - \tau_0}\right) + 2 - \frac{d\tau_2}{d\tau_1} - g$$

$$= \left(\frac{d\tau_2}{d\tau_1} - 1\right) - \frac{\tau_2 - \tau_1}{\tau_1 - \tau_0} + 2 - \frac{d\tau_2}{d\tau_1} - g$$

$$= 1 - g - \frac{\tau_2 - \tau_1}{\tau_1 - \tau_0}$$

$$= 1 + \frac{\tau_0 - \mu}{\tau_1 - \tau_0} - e^{\frac{\tau_0-\mu}{\tau_1-\tau_0}}$$

$$< 0.$$

The inequality follows from $e^x > 1 + x$ for $x > 0$. Thus decreasing τ_1 strictly increases revenue. $\qquad\square$

A.4 Proofs for Section 6

A.4.1 Proof of Proposition 6 on p. 23

Proof. Consider an evidence-test-fee structure, denoted by fees (ϕ_t, ϕ_d), an unbiased test $T : \Theta \to \Delta S$, and an evidence structure $M : S \rightrightarrows \mathcal{M}$ such that for each s, $M(s)$ is a Borel space. A strategy profile (σ, p) consists of the agent's strategy $\sigma = (\sigma_T, \sigma_D)$, where $\sigma_T \in [0, 1]$ and σ_D maps $s \in S$ to $\Delta(M(s) \cup \{N\})$, and the market price $p : \mathcal{M} \to [\underline{\theta}, \bar{\theta}]$. Let (σ, p) be an adversarial equilibrium. First we consider the case in which the agent has the asset tested with probability 1, that is $\sigma_T = 1$.

Consider the disclosure stage. Let $G_{(\sigma,p)}$ be the induced distribution of prices, i.e., $G_{(\sigma,p)}(x) = Pr[p(\sigma_D(s)) \le x]$ for any x, taking into account both the randomization over the score and the agent's strategy. Let $\tau \equiv p(N) + \phi_d$. We show that the following holds, mirroring our characterization of highest equilibria (HE):

$$\tau - \phi_d \le E_{G_{(\sigma,p)}}[x|x \le \tau], \tag{17}$$

$$\tau' - \phi_d > E_{G_{(\sigma,p)}}[x|x \le \tau'], \forall \tau' > \tau. \tag{18}$$

Let us argue why (17) holds. Since $\tau - \phi_d = p(N)$, it suffices to show that $p(N)$ is weakly less than $E_{G_{(\sigma,p)}}[x|x \le \tau]$. Observe that $p(\sigma_D(s))$ is at least $p(N)$ with probability 1: if $p(\sigma_D(s))$ were strictly less than $p(N)$, the agent could profitably deviate to sending message N and obtain a strictly higher price. But then this implies that $p(N) \le E[p(\sigma_D(s))|p(\sigma_D(s)) \le \tau] = E_{G_{(\sigma,p)}}[x|x \le \tau]$.

To see that (18) holds, suppose for contradiction that $\tau'' - \phi_d \le E_{G_{(\sigma,p)}}[p|p \le \tau'']$ for some $\tau'' > \tau$. By Lemma 5 there exists $\tau' > \tau$ such that $\tau' - \phi_d = E_{G_{(\sigma,p)}}[p|p \le \tau']$. Consider the strategy profile (σ', p') defined as follows. The agent's strategy σ' is the same as σ except that the agent conceals a score s if $p(\sigma(s)) \le \tau'$. For $m \ne N$, $p'(m) = p(m)$, and for $m = N$, the price is $p'(m) = E_{G_{(\sigma,p)}}[x|x \le \tau'] \ge E_{G_{(\sigma,p)}}[x|x \le \tau] \ge p(m)$. Notice that since any message m that is disclosed in (σ', p') is also disclosed in (σ, p), the prices p' are defined on path via the Bayes rule.

To see that (σ', p') is an equilibrium, consider any score s such that $p(\sigma(s)) > \tau$ with positive probability. Since s can send a message $m \ne N$ such that $p(m) > \tau$, it optimally randomizes over messages other than N that lead to the same (and maximum) price, which, abusing notation, we denote $p(\sigma(s))$. Since $p(m) = p'(m)$ for all $m \ne N$, $\sigma(s)$ is optimal among all strategies that send N with probability 0 given prices p'. Therefore, for such a score, it is optimal to follow $\sigma(s)$ if $p'(\sigma(s)) > \tau' = p'(N) + \phi_d$, and to conceal if $p'(\sigma(s)) \le \tau'$, as prescribed by σ'. Now consider a score s such that $p(\sigma(s)) \le \tau$ with probability 1. For such a score, it is optimal given prices p to conceal, i.e., for any message $m \ne N$ that s can send, $p(m) - \phi_d \le p(N)$. Since $p'(m) = p(m)$ and $p'(N) \ge p(N)$, it is also optimal to conceal given prices p', as prescribed by σ'.

Now consider the testing stage. If

$$\phi_t \ge \int_{\mu+\phi_d}^{\bar{\theta}} [x - (\mu + \phi_d)]dG_{(\sigma,p)}. \tag{19}$$

then there exists an equilibrium in the evidence-test-fee structure where the agent has the asset tested with probability 0. The argument parallels that of Lemma 1. Consider a strategy profile (σ', p') such that that $\sigma'_T = 0$, off-path the agent follows $\sigma(s)$ if $p'(\sigma(s)) > \mu + \phi_d$ ($p'(\sigma(s))$ is well-defined as argued above) and otherwise conceals, and the prices are $p'(N) = \mu$, and $p(m) = p'(m)$ for all $m \ne N$. Since the set of disclosed messages in (σ', p') is a subset of that in (σ, p), an argument similar to above shows that the agent's disclosure strategy is sequentially rational. Also, by deviating to take the test, the agent receives an expected payoff lower than μ,

$$\int_{\underline{\theta}}^{\mu+\phi_d} \mu dG_{(\sigma,p)} + \int_{\mu+\phi_d}^{\bar{\theta}} [x - \phi_d]dG_{(\sigma,p)} - \phi_t = \mu + \int_{\mu+\phi_d}^{\bar{\theta}} [x - (\mu + \phi_d)]dG_{(\sigma,p)} - \phi_t \le \mu$$

49

where the equality is algebra, and the inequality follows from (19). Therefore, adversarial revenue is at most zero, which is obtained by any test-fee structure with zero fees. So suppose that (19) is violated.

Now consider a test-fee structure $(G_{(\sigma,p)}, \phi)$. By Lemma 3, τ is a weak-highest equilibrium threshold of the test-fee structure. Also since (19) is violated, by Lemma 1, the test is taken with probability 1 in all equilibria. Therefore, adversarial revenue of the test-fee structure is equal to the adversarial revenue of the evidence-test-fee environment.

Now we consider the case $\sigma_T \in [0, 1)$. Notice that in this case

$$
\phi_t \geq \int_{p(N)+\phi_d}^{\overline{\theta}} [x - (p(N) + \phi_d)] dG_{(\sigma,p)} \geq \int_{\mu+\phi_d}^{\overline{\theta}} [x - (\mu + \phi_d)] dG_{(\sigma,p)}
$$

so (19) holds. Form the same argument above we know that the adversarial revenue is at most zero, which can be obtained in a test-fee structure with any test and zero fees. □