Delegating Experiments

Cole Wittbrodt and Nathan Yoder*

June 30, 2026

Abstract

A principal wants information to help her decide whether to approve a project. She delegates costly experimentation to an agent, who wants her to approve the project and has private information about whether it is good or bad. The principal can influence experimentation by restricting the experiments that the agent can undertake and by committing to approval rules, but transfers are unavailable. For example, the FDA may select a set of clinical trials that are acceptable for testing a new drug and the results needed for approval, but cannot pay drug companies. We show that the principal optimally restricts the set of experiments, but need not commit to an approval rule. Higher types choose tests with higher true positive rates in the optimal menu. Private information distorts the optimal menu by making the false negative rate inefficiently high: too many good projects are rejected.

*Wittbrodt: Columbia University, Graduate School of Arts and Sciences, Department of Economics; E-mail: cole.wittbrodt@columbia.edu. Yoder: University of Georgia, Department of Economics; E-mail: nathan.yoder@uga.edu. We thank Yeon-Koo Che, Grace Chuan, Laura Doval, Navin Kartik, Elliot Lipnowski, Qingmin Liu, Jacopo Perego, Ian Schmutte, Akhil Vohra, and seminar participants at Columbia for helpful comments. Yoder acknowledges summer support from the University of Georgia through a Terry-Sanford Research Award. GPT, Claude, and Gemini provided excellent research assistance.

1 Introduction

Individuals and organizations often delegate the acquisition of information to help them make decisions. Drug regulators delegate the design of clinical trials to pharmaceutical companies; central banks delegate some components of stress testing to the banks that they regulate; and judicial systems delegate the investigation of defendants to police and prosecutors in criminal trials. Frequently, the agents that they delegate to have private information about the things they are asked to investigate, and care about the decisions that result from their investigations. For instance, pharmaceutical firms may be better informed about the quality of their prototype drugs than the Food and Drug Administration (FDA), and want those drugs to be approved; retail banks may be better informed than the Federal Reserve about their resilience to financial shocks, and want to pass stress tests.

In practice, regulators deploy a variety of instruments to constrain the experiments that agents conduct. The FDA restricts the set of acceptable clinical trial protocols by mandating pre-specified endpoints, statistical analysis plans, and significance levels—and can halt trials if it deems that their designs are scientifically deficient.[1] It can also provide assurance that a given trial design will result in approval if successful.[2] Similarly, in bank stress testing, the Federal Reserve restricts the scenarios that banks can use and requires that they be at least as severe as the regulator's own.[3]

This paper studies how a regulator should design these instruments when the agents she delegates to have private information about the state of the world. We consider a principal who can limit the agent's discretion over experiments and commit to an approval rule — that is, a rule specifying the principal's decision as a function of experimental outcomes.

Concretely, we work in a model with two states of the world: a *low* state and a *high* state. A principal will decide whether to approve or reject a project, preferring to approve when the state is high. An agent prefers that the principal approves the project. The agent receives a private, noisy signal about the state. In order to persuade the principal to approve the project, the agent can conduct an experiment. Experiments produce hard information[4] and are costly, with costs independent of the prior.[5] Experiments that fully re-

[1] See U.S. Food and Drug Administration (2024b) and U.S. Food and Drug Administration (2024a).

[2] In the form of a Special Protocol Assessment; see U.S. Food and Drug Administration (2018).

[3] See Board of Governors of the Federal Reserve System (2015).

[4] That is, the results of experimentation are verifiable and credible to any interested party.

[5] For some of our results, we additionally assume a constant marginal cost in the sense of Pomatto et al. (2023).

veal the state are prohibitively costly, and so are never chosen by the agent. The principal can restrict the agent to experiments from a *menu*, and can commit to approval decisions associated with each experiment in the menu. However, the agent cannot be compelled to experiment; he always has the option to not communicate.[6]

One might expect that the rules governing approval are a crucial lever for the regulator. For instance, the FDA could promise to hold firms that conduct thorough drug tests to a *lower* standard of approval to provide incentives. However, our first main result establishes that this commitment power is unnecessary: the approval rule used in the principal's optimal mechanism is *ex-post* optimal (Theorem 1). Hence, the principal can implement her optimal mechanism simply by specifying a menu of experiments that the agent can choose from (e.g. clinical trial guidelines and Federal Rules of Evidence in legal proceedings).

We characterize the principal's optimal menu of experiments and show that it is unique (Theorem 2). Each experiment in the optimal menu is *binary*: it yields a positive or negative outcome, with the project approved if and only if the result is positive. We show that a single scalar — *Youden's index*, the difference between the true and false positive rates — is the aspect of the experiment in which the agent's payoff satisfies the single crossing property in his type. As a result, any implementable menu must be monotone in Youden's index (Proposition 2). The optimal menu satisfies an additional condition: Proposition 3 shows that the true positive rate is increasing in the agent's type, so that higher types are approved more frequently when the state is high. Intuitively, the principal must leave information rents to the agent, and since principal and agent are aligned when the state is high, the most cost-effective way to do so is by increasing the true positive rate. Thus, higher types conduct more powerful tests, though not necessarily at a lower significance level.

These conclusions follow from a characterization of implementable menus that is strikingly parallel to standard contracting results, despite the absence of transfers:[7] As Proposition 1 shows, envelope and monotonicity conditions ensure global incentive compatibility. This is in sharp contrast to one-dimensional delegation problems, where such a characterization does not obtain.[8] Intuitively, types are distinguished by their belief that the state is high. The principal can therefore screen them by creating a tradeoff between expected payoffs conditional on the high and low states, much as screening in a standard contracting problem exploits a tradeoff between quantities and transfers.

[6]That is, he can choose to "veto" a particular delegation set (Kartik et al., 2021).

[7]e.g., Myerson (1981), Mussa and Rosen (1978), Maskin and Riley (1984).

[8]E.g., Alonso and Matouschek (2008).

We next turn to the distortions created by private information. As in standard contracting problems, there is *no distortion at the top*: the highest type's experiment is efficient (Proposition 5). Likewise, there is *distortion everywhere else*: no other type conducts an efficient experiment (Theorem 3), and the principal may find it optimal to inefficiently exclude some types entirely.

Since transfers are absent, this distortion is not relative to a single efficient allocation, but rather an entire Pareto frontier of efficient experiments. We thus describe the way that the optimal menu distorts experimentation by comparing each experiment it induces to the efficient experiments that are Pareto-improving under symmetric information (Theorem 3). Relative to this benchmark, distortion does not necessarily take the form of a less informative experiment: false positive rates may be higher or lower. However, the false negative rate is inefficiently high for all non-excluded types other than the highest. Hence, the principal approves good projects too infrequently.

Our results have implications for existing practice. The FDA gives sponsors flexibility in proposing clinical trial designs because they have product-specific information — e.g., about prior evidence, dose response, and feasibility (U.S. Food and Drug Administration, 2009, 2023). But the FDA also restricts those proposals to ensure that a successful trial provides adequate evidence for approval (U.S. Food and Drug Administration, 2018, 2024a). In the language of our model, this flexibility lets the sponsor choose an experiment that is individually rational, while these restrictions ensure that approval after a successful outcome is ex-post optimal.

However, the public record does not describe these restrictions as devices for screening privately informed sponsors. Our results show how they can serve this role as well: when sponsors are privately informed about drug quality, regulators can raise patient welfare by designing the set of acceptable protocols to elicit that information while ensuring that approval is still optimal ex-post (Theorem 1). In particular, Proposition 2 identifies the key feature of the experiment that allows for such screening (Youden's index). An optimal menu induces more optimistic trial sponsors to conduct more costly trials with higher power (Proposition 3, Corollary 1); the restrictions in the menu raise welfare by requiring such trials to have greater statistical significance (equivalently, lower false positive rate) than the bare minimum required to justify approval.

Related Literature

Our paper belongs to a large body of work on delegation following Holmström (1977). Unlike many papers in this literature, the choice being delegated in our model has much higher dimension than the space of agent types. As we show, this gives the principal a greater ability to utilize the agent's private information: Alonso and Matouschek (2008) show that when delegating a one-dimensional choice, the principal can only choose a subset of types to force into a corner solution, while leaving the rest effectively unconstrained. But in our model, the principal can fully screen the agent's private information even without access to transfers. This feature is reminiscent of Koessler and Martimort (2012), who describe optimal delegation in settings where delegation sets are subsets of $\mathbb{R}^2$. They find that the *spread* between the decisions can allow the designer to screen the agent's private information. Analogous to spread in Koessler and Martimort (2012), we show that in our model, the designer can use the difference in the agent's expected utility conditional on the state being high and the agent's expected utility conditional on the state being low to screen the agent's private information.

Especially relevant to our paper is a recent literature on the delegation of *dynamic* experimentation. Guo (2016) studies a one-armed bandit model where an agent has private information about the risky arm's payoff, and the principal can limit the agent's freedom to allocate resources between arms (in the form of a history-dependent policy). Closer to our paper is McClellan (2022), who also considers a setting where an agent experiments to influence the approval decision of a principal. The key differences relative to our paper are that (a) the principal incentivizes the agent only by committing to an approval rule, rather than explicitly limiting the experiments that the agent can conduct, and (b) because experimentation is dynamic á la Wald (1947), the principal faces additional incentive constraints.[9]

Within the literature on information design (e.g., Kamenica and Gentzkow (2011)), other authors (e.g., Hedlund (2017); Kosenko (2023)) have explored the consequences of private information for the sender. Our paper adds to this literature by allowing another party (the principal) to constrain the sender's experimentation decision. That is, in addition to a *signaling* problem with commitment faced by a sender with private information, we consider a receiver's problem of *screening* that sender.

Contemporaneous work by Hancart (2025) considers a closely related setting, which

[9]The Wald (1947) setup is closely related to flexible information acquisition with likelihood ratio costs; see Morris and Strack (2019). But because the principal only observes the ex-post realization of the agent's experiment in McClellan (2022), she cannot fully restrict the design of the agent's experiment (or infer his private information from it) the way she can in our model.

An earlier version (McClellan, 2017) considers the case of *two-sided commitment* where the agent can commit to an experimentation policy, and so the principal only faces a static incentive constraint. Our results also characterize the optimal mechanism in this setting: As Morris and Strack (2019) show, the cost of attaining a distribution of posteriors with Wald (1947) experimentation is just its LLR cost. Thus, a menu of static-threshold stopping mechanisms in McClellan's (2017) setting (as he shows is optimal with two-sided commitment) is equivalent to a menu of binary Blackwell experiments in ours. In addition, we show that the principal *need not commit* to an approval rule in such settings.

also features a principal delegating a menu of experiments to an agent with private information about the state of the world, and then taking a binary action after viewing the experiment's result. Two differences between our papers stand out as most important. First, we consider the question of delegating to an agent whose private information is a *signal* about the state, whereas Hancart (2025) considers the question of delegating to an agent whose private information is *itself* the payoff-relevant state. This means that the results of an experiment play a different role in our paper: instead of providing information about an agent's type, they provide information that *cannot be inferred from* an agent's type (and in the optimal menu, provide no information about the agent's type that cannot be inferred from his choice of experiment). Second, experimentation is flexible and costly in our paper, whereas in Hancart (2025) it is costless and limited by a technological constraint.

We also follow a large recent literature on contracting for flexible information acquisition. Rappoport and Somma (2017), Whitmeyer and Zhang (2022), and Sharma et al. (2024) each consider models where the principal does not possess private information, and so the contracting problem is one of moral hazard. They focus on the impact of risk aversion and limited liability constraints on the optimal payment scheme, as well as whether that scheme can be implemented by contracting on the experiment's result and/or the realized state, rather than on the experimental protocol itself. Closer to this paper are Yoder (2022) and Wang (2023), who (like us) consider settings where the agent has private information, but (unlike us) where that private information concerns the agent's cost of experimentation. Relative to the bulk of this literature, our paper has three novel features: (i) transfers are unavailable to the principal; (ii) the principal's decision is payoff-relevant to the agent; and (iii) the agent has private information about the state of the world itself.

This article is also related to work on the design of approval rules in statistical decision theory. Tetenov (2016) studies how a regulator can design a statistical test when agents have private information about the state. In his setting, the principal chooses a single experiment which all agents conduct, and commits to approval when the experiment produces a positive result. In contrast, our model highlights the usefulness of a menu of experiments to screen the agent. Jagadeesan and Viviano (2024) study the design of publication rules (i.e., should unsurprising results be published?) when agents decide how to experiment. In their setting, researchers do not possess ex-ante[10] private information, but can manipulate their findings. Adusumilli and Vemulapati (2026) propose allowing tests that satisfy an ex-ante welfare threshold, and show how an experimenter that acquires in-

[10] That is, they are not more informed about the state prior to choosing an experiment.

formation dynamically will respond to that threshold.

The paper proceeds as follows. Section 2 introduces the model. Section 3 studies implementability: we show that the principal need not commit to an approval rule (Theorem 1), reduce her problem to the choice of a menu of binary experiments, and characterize the implementable menus through envelope and monotonicity conditions in which Youden's index plays the role of a sufficient statistic (Propositions 1 and 2). Section 4 characterizes the optimal menu: we establish existence and uniqueness (Theorem 2), show that the true positive rate is increasing in the agent's type (Proposition 3), and then use a characterization of the Pareto frontier (Section 4.2) to describe how private information distorts experimentation—no distortion at the top, but an inefficiently high false negative rate for every other included type (Proposition 5 and Theorem 3). Section 5 concludes.

2 Model

There is a binary state of the world $\omega \in \{0, 1\}$. An *agent* (he) with private but imperfect information about ω can publicly conduct a costly experiment about the state. A *principal* (she) can choose the set of experiments available to the agent. Before the agent observes his private information, both he and the principal place prior probability β_0 on the event $\omega = 1$.

The Principal After observing the experiment conducted by the agent and its result, the principal must decide whether to approve a project (such as a new drug application) whose value to her depends on the state of the world. She wants to approve the project in one state (without loss, $\omega = 1$) but not the other. Approving the project gives the principal a payoff of $w_1 > 0$ when $\omega = 1$ and $w_0 < 0$ when $\omega = 0$, while disapproval yields a payoff of zero.[11] Hence, when her belief is $P(\omega = 1) = \beta$, she is better off approving the project if and only if

$$\beta w_1 + (1 - \beta) w_0 \geq 0 \Leftrightarrow \beta \geq \frac{-w_0}{w_1 - w_0} =: b.$$

We call b the principal's *threshold belief*, and assume that $\beta_0 < b < 1$: approval is optimal if and only if she receives information that is sufficiently suggestive of state 1.

[11] Since the principal cannot make transfers to the agent, letting her disapproval payoff be constant at zero is without loss.

The Agent The agent privately observes a signal about ω that causes him to update his belief to $\theta \in \Theta := \{\theta_0, \dots, \theta_N\}$, where for each $1 \leq n \leq N$, $\theta_n > \theta_{n-1}$. We refer to the interim belief θ as the agent's *type*, and describe its distribution with the probability mass function σ; by Bayes' rule, $\sum_{n=0}^{N} \sigma(\theta_n)\theta_n = \beta_0$. We assume that $\theta_N < b$, so that no type would be able to convince the principal to approve the project with only his private signal.

The agent is motivated to conduct costly experiments by his desire to persuade the principal to approve the project. He has *transparent motives*, in the sense that his payoffs are state-independent: He receives a payoff of $u > 0$ if the project is approved, and zero otherwise.

Experimentation The agent conducts a Blackwell experiment of the form $(S, \pi : \Omega \to \Delta(S))$, where S is a set of signal realizations.[12] We typically abuse notation and refer to this experiment as π when the set S is clear from context. Let Π denote the set of Blackwell experiments.[13] When the result of an experiment is observed by an individual with belief α, π induces a distribution over posterior beliefs $\langle \pi|\alpha \rangle \in \Delta([0,1])$ through Bayes' rule, where $E_{\langle \pi|\alpha \rangle}\beta = \alpha$. We say that two experiments π and π' are *Blackwell equivalent*, denoted $\pi \sim_B \pi'$, if for any $\alpha \in [0,1]$, $\langle \pi|\alpha \rangle = \langle \pi'|\alpha \rangle$.[14] We denote the Blackwell ordering by $\succsim_B$ throughout.

Experimentation is costly for the agent. We assume that, when the agent conducts the experiment π, he pays the *log-likelihood ratio cost* (Pomatto et al., 2023) $C(\pi)$, where

$$C(\pi) \equiv \sum_{\omega \in \{0,1\}} \int_S \log\left(\frac{d\pi(\cdot|\omega)}{d\pi(\cdot|1-\omega)}(s)\right) d\pi(s|\omega) = E_{\langle \pi|\alpha \rangle}[G(\beta|\alpha)] \text{ for any } \alpha \in (0,1),$$

where $G(\beta|\alpha) \equiv \frac{\beta}{\alpha}\log\left(\frac{\beta}{1-\beta}\right) + \frac{1-\beta}{1-\alpha}\log\left(\frac{1-\beta}{\beta}\right)$ is the *posterior-specific cost* for β given an initial belief α.

The log-likelihood ratio (LLR) cost function has three desirable properties in the context of our model. First, LLR costs are *prior-independent*. For a fixed experiment $\pi \in \Pi$, the cost of conducting the experiment $C(\pi)$ does not depend on the agent's type. Experiments should be interpreted as physical processes in our model (e.g. clinical trials and stress tests), which is intuitively consistent with prior-independent information costs (Bloedel and Zhong, 2021). Second, LLR costs are posterior separable.[15] Posterior

[12]Naturally, S must be a Polish set.

[13]Formally, let $\Pi = \{(S, \pi) : S \in \mathcal{S}\}$ be the set of all Blackwell experiments where $\mathcal{S}$ is a sufficiently rich set containing sets of signal realizations. It suffices, for instance, to let $\mathcal{S}$ be any uncountable Polish space.

[14]Equivalently, $\pi \succsim_B \pi'$ and $\pi' \succsim_B \pi$ where $\succsim_B$ is the Blackwell informativeness ordering. When showing two experiments are Blackwell equivalent, it suffices to show that $\langle \pi|\alpha \rangle = \langle \pi'|\alpha \rangle$ for some non-degenerate prior $\alpha \in (0,1)$.

[15]But not uniformly posterior separable.

separability yields tractability, and allows us to use the concavification approach in Kamenica and Gentzkow (2011) to characterize solutions. Third, among prior-independent cost functions, LLR costs are uniquely characterized by a *constant marginal cost condition* (Pomatto et al., 2023): $C(\pi + \pi') = C(\pi) + C(\pi')$ for all experiments independent π, π'. This last feature (and the LLR functional form) are not needed for most of our main results, but allow us to characterize the distortion from private information in our setting (Theorem 3).[16]

Delegation Prior to the agent's experimentation decision, the principal can limit the agent's discretion. In line with much of the delegation literature, we write the principal's choice over delegation sets as their choice over *stochastic direct mechanisms* $\chi : \Theta \to \Delta(\Pi \times \mathcal{A})$, where $\mathcal{A}$ is the set of decision rules $a : \{(\pi, s) | (S, \pi) \in \Pi, s \in S\} \to [0, 1]$ that map experiment/signal realization pairs into approval probabilities.[17] That is, the principal commits to the mapping from types into experiments and commits to a rule specifying the signal realizations after which the project is approved. Denote the set of all direct mechanisms by $\mathcal{X}$.

While we allow the principal to restrict the agent's discretion by committing to any stochastic direct mechanism, we will show in Section 3 that the optimal arrangement simplifies to offering a *menu* of experiments $\{\pi_\theta\}_{\theta \in \Theta}$, and approving only when it is optimal to do so after observing their results. That is, the principal uses approval rules a_θ that are *ex-post optimal* for π_θ, in the sense that for all $s \in \text{supp}\, \pi_\theta(\cdot|1) \cup \text{supp}\, \pi_\theta(\cdot|0)$,

$$a_\theta(\pi_\theta, s) = \begin{cases} 1, & \text{if } w_1 \pi_\theta(s|1)\theta + w_0 \pi_\theta(s|0)(1-\theta) \geq 0; \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Although the principal can restrict the set of experiments available to the agent, the agent can always refuse to experiment at all by conducting an *uninformative experiment*, which we represent as $(\{\underline{s}\}, \pi_0)$ with $\pi_0(\underline{s}|1) = \pi_0(\underline{s}|0) = 1$, in which case (without loss) the project is not approved. We also assume that in the absence of asymmetric information, there are gains from trade between the principal and the agent: there exists $(\pi, a) \in \Pi \times \mathcal{A}$ such that $\int_S (a(\pi, s)u - C(\pi)) d\pi(s|\theta) > 0$ and $\theta \int_S a(\pi, s) w_1 d\pi(s|1) +$

[16] For Proposition 3 and Theorem 2, we only require the posterior cost function to be convex, i.e., $C(\pi) = E_{\langle \pi | \alpha \rangle}[H(\beta|\alpha)]$ for some function H that is convex in its first argument.

[17] Naturally, we appeal to the revelation principle in restricting attention to direct mechanisms. In an earlier version of this paper, we did not allow for the principal to commit to an approval rule. The principal approved only when it was ex-post optimal to do so. In such a setting, one cannot appeal to the revelation principle. Theorem 1 shows that the assumption that the principal can commit to approval rules does not bind for the optimal mechanism.

$(1-\theta)\int_S a(\pi,s)w_0 d\pi(s|0) > 0$ and a is ex-post optimal for π.

3 Implementability

3.1 The Principal's Problem

Appealing to the revelation principle, a mechanism χ is feasible if each type finds it optimal to report their type truthfully, and each type's interim participation constraint is satisfied. That is, mechanisms are feasible if and only if they are individually rational and incentive compatible. Formally, a mechanism is individually rational if for all $\theta \in \Theta$,

$$\int_{\Pi\times\mathcal{A}}\int_S \left(a(\pi,s)u - C(\pi)\right) d\pi(s|\theta)d\chi(\theta) \geq 0. \tag{IR}$$

The inner integral represents the agent's expected utility from a particular experiment/decision rule (π, a), with (an abuse of notation) the distribution over signal realizations s depending on the agent's prior belief θ. The outer integral represents the agent's expectation over experiment/decision rule pairs. A mechanism is incentive compatible if for all $\theta, \theta' \in \Theta$,

$$\int_{\Pi\times\mathcal{A}}\int_S \left(a(\pi,s)u - C(\pi)\right) d\pi(s|\theta)d\chi(\theta) \geq \int_{\Pi\times\mathcal{A}}\int_S \left(a(\pi,s)u - C(\pi)\right) d\pi(s|\theta)d\chi(\theta'). \tag{IC}$$

By the revelation principle, the principal chooses a direct mechanism $\chi \in \mathcal{X}$ to maximize her payoffs subject to individual rationality and incentive compatibility. Formally, she solves

$$\max_{\chi\in\mathcal{X}} \sum_{\theta\in\Theta} \sigma(\theta) \int_{\Pi\times\mathcal{A}} \left(\theta \int_S a(\pi,s)w_1 d\pi(s|1) + (1-\theta)\int_S a(\pi,s)w_0 d\pi(s|0)\right) d\chi(\theta) \; s.t. \text{ (IR), (IC)}. \tag{OPT}$$

Since (OPT) is a linear program in (π, a) (all payoffs are linear in the allocation rule), without loss of generality, we can restrict attention to extreme points of the set of feasible mechanisms. Namely, we can restrict our attention to deterministic mechanisms $\chi : \Theta \to \Pi \times \mathcal{A}$.

While the principal can commit to approving the project even after experimental results that would make approval suboptimal, our first main result shows that it is not optimal for her to do so. That is, the principal never gains by committing to approval rules that are not ex-post optimal.

Theorem 1 (Ex-post Optimal Decision Rules). *If* (OPT) *has a solution, then it has a solution* χ^* *such that each* $\chi^*(\theta)$ *is deterministic, so* $\chi^*(\theta) = (\pi^*_\theta, a^*_\theta)$*, and each* a^*_θ *is ex-post optimal for* π^*_θ*.*

By Theorem 1, the principal does not need to commit to an approval rule ex-ante. She can simply specify a mapping from types into experiments, approving only when her posterior belief lies above the threshold b. Therefore, the optimal mechanism χ^* can be implemented in a simple manner: the principal allows the agent to choose an experiment from a *delegation set*, or menu, $D \subseteq \Pi$ which contains the uninformative experiment (reflecting the fact that the agent has a participation constraint).

One might conjecture that the principal's use of ex-post optimal decision rules follows from the fact that her payoff is maximized by using the decision rule a^*_θ for any realization of θ. However, this reasoning is incorrect. A priori, the principal may wish to use ex-post suboptimal decision rules a_θ to provide incentives to the agent — e.g., to ensure that approval happens often enough that experimentation is individually rational for some type. The bite of Theorem 1 is that this is never optimal.

Proof Sketch for Theorem 1

The proof of Theorem 1 involves three steps. First, we establish a fundamental result that we use throughout the paper: the *continuity lemma* (Lemma 5). Informally, the continuity lemma says that, for a mechanism χ, if no incentive constraint binds for type θ, then the principal can achieve a higher payoff when interacting with type θ by offering an allocation where some constraint binds for type θ.[18] In standard delegation problems with discrete types, incentive constraints need not bind. In contrast, in mechanism design problems with transfers and discrete types, typically local downward incentive constraints bind.[19] In our setting, the continuity lemma arises as a consequence of the richness of the space of allocations, and can be used to establish connections between our problem and standard contracting problems with transfers.

The continuity lemma implies that some incentive constraint must bind for each type. We show that, for every type except the lowest, this binding constraint is the downward incentive constraint. That is, in the optimal mechanism, type θ_n is indifferent between reporting his type truthfully and reporting θ_{n-1} whenever $n \geq 1$. For the lowest type θ_0, it is instead the participation constraint that binds. In Section 3.2 we give additional intuition for these downward binding incentive constraints in the context of the limited

[18] In fact, the continuity lemma is much stronger than this. We also show that the allocation rule which causes the constraint to bind involves a binary experiment without loss of generality and has the same *Youden's index* (see p. 15) as the original experiment.

[19] Along with the participation constraint for the lowest type.

commitment problem.

Finally, we establish the optimality of ex-post optimal decision rules. Consider any allocation where a_{θ_n} is not ex-post optimal (without loss, let θ_n be the lowest type with this property). We show that the principal can do strictly better by modifying the menu in the following manner: if $n \geq 1$, assign type θ_n the allocation originally designed for type θ_{n-1}. If $n = 0$, assign θ_n the uninformative experiment, and never approve the project. This mechanism results in a higher payoff for the principal, but may fail to be incentive compatible. In particular, type θ_{n+1}'s downward binding incentive constraint may no longer hold. Using the continuity lemma, however, we can edit the allocation of type θ_{n+1} in such a way where the local downward incentive constraint binds and the principal is better off. Moreover, this can be done in such a way that does not affect the local downward incentive constraint of type θ_{n+1}.

3.2 Menus and Implementation

In light of Theorem 1, the principal need not specify an approval rule in the optimal mechanism. Hence, the optimal mechanism can be implemented by allowing the agent to choose from a menu $D^* \subseteq \Pi$ containing the uninformative experiment. Before writing the principal's problem, we find it useful to describe which *menus* $D = \{\pi_\theta\}_{\theta \in \Theta} \cup \{\pi_0\}$ are implementable (supposing that the principal approves if and only if it is ex-post optimal). We refer to π_θ as the experiment *designed* for type θ. Unlike the previous section, we often write incentive compatibility and individual rationality constraints in terms of distributions over posterior beliefs.

Since the agent's cost function is posterior-separable, his payoffs depend on the experiment he selects only through the distribution of posterior beliefs that the experiment induces. However, since the principal cannot infer the agent's private information about the state by observing his choice of experiment, she may have a different posterior than the agent does after observing the experiment's result. This is particularly relevant when considering an agent's payoff from deviating to the experiment designed for some other type.

Consequently, the same experiment can induce up to three different distributions that are relevant to our analysis: the distribution of the principal's posteriors, the distribution of the agent's posteriors, and the distribution of the principal's posteriors *from the perspective of the agent*.

The first two of these can be described using notation that we have already introduced: if the principal's interim belief after observing the agent's choice is α, then π induces the distribution $\langle \pi | \alpha \rangle$ for her and $\langle \pi | \theta \rangle$ for the agent. Lemma 1 characterizes the last of these distributions in the same terms.[20]

Lemma 1 (Agent's Distribution of Principal's Posteriors). *Suppose α is the principal's interim belief after she observes the agent choose the experiment π. Then the agent places probability $\int_B \left(\frac{\theta}{\alpha}\beta + \frac{1-\theta}{1-\alpha}(1-\beta)\right) d\langle\pi|\alpha\rangle(\beta)$ on the principal updating her belief to $\beta \in B$ after observing the result of π.*

Lemma 1 allows us to describe the set of implementable menus in terms of the distributions of posterior beliefs that they induce for the agent. Observe that since the principal's approval rule is ex-post optimal, when the principal's posterior belief is β, her interim expected payoff is given by $W(\beta)$ and the agent's gross payoff[21] is given by $U(\beta)$, where

$$W(\beta) := \begin{cases} 0, & \beta < b; \\ w_0 + (w_1 - w_0)\beta, & \beta \geq b, \end{cases} \quad \text{and} \quad U(\beta) := \begin{cases} 0, & \beta < b; \\ u, & \beta \geq b. \end{cases}$$

Then a menu $D = \{\pi_\theta\}_{\theta\in\Theta} \cup \{\pi_0\}$ is individually rational if

$$E_{\langle\pi_\theta|\theta\rangle}\left[U(\beta) - G(\beta|\theta)\right] \geq 0 \tag{IRθ}$$

holds for all θ, and is incentive compatible if

$$E_{\langle\pi_\theta|\theta\rangle}\left[U(\beta) - G(\beta|\theta)\right] \geq E_{\langle\pi_{\theta'}|\theta'\rangle}\left[\left(\tfrac{\theta}{\theta'}\beta + \tfrac{1-\theta}{1-\theta'}(1-\beta)\right)U(\beta) - G(\beta|\theta')\right] \tag{ICθ}$$

holds for all $\theta, \theta' \in \Theta$.[22]

Even though our setting does not involve transfers, appealing to the continuity lemma allows us to offer a characterization of the class of implementable menus that is reminiscent of standard transferable utility contracting results (e.g., Maskin and Riley (1984)).

Proposition 1 (Implementable Menus). *Let $D = \{\pi_\theta\}_{\theta\in\Theta} \cup \{\pi_0\}$ be a menu.*

i. If D satisfies

(EC) *Envelope Condition: For all θ,*

$$E_{\langle\pi_\theta|\theta\rangle}[U(\beta)] - C(\pi_\theta) = \sum_{\theta_i<\theta}(\theta_{i+1} - \theta_i)E_{\langle\pi_{\theta_i}|\theta_i\rangle}\left[\left(\tfrac{\beta-\theta_i}{\theta_i(1-\theta_i)}\right)U(\beta)\right]; \tag{ECθ}$$

[20] Lemma 1 is complementary to Proposition 1 in Alonso and Câmara (2016), who also study a setting where the priors of the sender and receiver may differ: they compute the receiver's belief as a function of the sender's belief, whereas we instead compute the probability that the sender places on the receiver having a certain belief.

[21] Before considering the cost of the experiment.

[22] That is, D satisfies (IRθ) (respectively, (ICθ)) if and only if the mechanism $\chi \in \mathcal{X}$ formed by pairing D with ex-post optimal decision rules a^*_θ (i.e., $\chi(\theta) = (\pi_\theta, a^*_\theta)$) satisfies (IR) (resp., (IC)).

(M) *Monotonicity: For all $\theta \geq \theta'$,*

$$E_{\langle \pi_\theta | \theta \rangle} \left[\left(\frac{\beta - \theta}{\theta(1-\theta)} \right) U(\beta) \right] \geq E_{\langle \pi_{\theta'} | \theta' \rangle} \left[\left(\frac{\beta - \theta'}{\theta'(1-\theta')} \right) U(\beta) \right] \geq 0, \qquad (\mathrm{M}(\theta, \theta'))$$

then D is implementable.

ii. If D is implementable, it satisfies (M) (monotonicity).

iii. If D is implementable, there exists an implementable menu $D' = \{\pi'_\theta\}_{\theta \in \Theta} \cup \{\pi_0\}$ that satisfies (EC) such that $E_{\langle \pi'_\theta | \theta \rangle}[W(\beta)] \geq E_{\langle \pi_\theta | \theta \rangle}[W(\beta)]$ for each $\theta \in \Theta$.

The key insight behind this characterization is that the law of iterated expectations allows us to decompose the expectations in (ICθ) into expectations of U *conditional on the state*. This decomposition has two implications. First, because experiments are multidimensional objects, these conditional expectations can vary separately, just like transfers and quantities in, e.g., Maskin and Riley (1984) and Mussa and Rosen (1978). And since types differ precisely in the probabilities that they assign to each state, they can be screened by offering a lower type an experiment that yields higher expected utility conditional on state 0 and by offering a higher type an experiment that yields higher expected utility conditional on state 1. Formally, every implementable menu must satisfy the following condition: if $\theta > \theta'$, then

$$E_{\langle \pi_\theta | \theta \rangle}[U(\beta) | \omega = 1] - C(\pi_\theta) \geq E_{\langle \pi_{\theta'} | \theta' \rangle}[U(\beta) | \omega = 1] - C(\pi_{\theta'});$$
$$E_{\langle \pi_\theta | \theta \rangle}[U(\beta) | \omega = 0] - C(\pi_\theta) \leq E_{\langle \pi_{\theta'} | \theta' \rangle}[U(\beta) | \omega = 0] - C(\pi_{\theta'}).$$

As in the standard transferable utility model, this screening is most effective when local incentive compatibility constraints bind. However, these local constraints — captured in (ECθ) — are not enough to ensure that a menu is incentive compatible. Instead, ensuring that global deviations in type reports are not advantageous requires an additional monotonicity condition. Intuitively, a monotonic menu is one in which the difference between the conditional expected payoff in state $\omega = 1$ and state $\omega = 0$ is increasing in the type report.

Second, the participation constraint for any type $\theta > \theta_0$ is redundant, just like in a transferable utility setting. Since U is increasing, its expectation conditional on $\omega = 1$ is always higher than its expectation conditional on $\omega = 0$, and so the expectations on the right hand side of (ICθ) are higher for $\theta > \theta_0$ than for $\theta = \theta_0$.

3.3 Binary Experiments

A *binary experiment* is (S, π) with $S = \{\underline{s}, \bar{s}\}$; without loss, we label S so that $\pi(\bar{s}|0) \leq \pi(\bar{s}|1)$. We often refer to $\pi(\bar{s}|1)$ as the true positive rate and $\pi(\bar{s}|0)$ as the false positive rate. If every informative experiment $\pi \not\sim_B \pi_0$ in a menu D is binary, we say that it is a *binary menu*. As observed by Doval and Skreta (2023), we cannot appeal to standard two-state information design arguments to conclude that binary menus are optimal, because the principal faces both incentive compatibility and individual rationality constraints. Nevertheless, we use the continuity lemma to show that binary experiments are without loss in our setting. Intuitively, if D is a menu containing non-binary experiments, replacing each non-binary experiment π_θ with the binary experiment π'_θ that leads the principal to approve the project with the same probabilities conditional on the state does not lower principal payoffs, but weakens the agents' incentive constraints since $C(\pi_\theta) > C(\pi'_\theta)$. The continuity lemma can then be used to construct an experiment π''_θ such that the agent's incentive constraints are satisfied in $D'' = (D \setminus \{\pi_\theta\}) \cup \{\pi''_\theta\}$ and such that the principal strictly prefers D'' to D.

Lemma 2 (Menus Are Binary Without Loss). *For any non-binary implementable menu D, there is an implementable binary menu D'' that gives the principal a higher ex-ante expected payoff than D.*

For binary tests, a common measure of the performance of a diagnostic test in the medical literature is *Youden's index* (Youden, 1950). Youden's index is the difference between the true positive rate and the false positive rate. In our notation, we denote the Youden's index for a binary experiment π as $\eta(\pi) := \pi(\bar{s}|1) - \pi(\bar{s}|0)$.

In the literature, tests with higher Youden's indices are often considered to be better performing. This is not entirely congruent with Blackwell's Theorem: If one binary experiment is Blackwell-more informative than another, it must have a higher Youden's index, but the converse is not true. However, a clarification reveals the relevance of the Youden's index to implementation: tests with higher Youden's indices are better performing *for* an agent with transparent motives *when* the hypothesis being tested is more likely to be true.

Observe that when a type-θ agent conducts a binary experiment π_θ, his expected payoff can be written

$$E_{\langle \pi_\theta | \theta \rangle}[U(\beta)] - C(\pi_\theta) = (\theta\eta(\pi_\theta) + \pi_\theta(\bar{s}|0))u - C(\pi_\theta), \tag{2}$$

so long as the principal approves after $\bar{s}$ but not after $\underline{s}$. Thus, since the cost of experimentation C is prior-independent, an experiment's Youden's index is a sufficient statistic for the way that the agent's payoff from conducting it depends on his type. In particular, if the Youden's index of one experiment is higher than that of another, then whenever some type chooses the former over the latter, all higher types do as well. That is, *the Youden's index is the aspect of the experiment in which the agent's payoff has the single crossing property in* θ*.* It should not be surprising, then, that the monotonicity constraint (M) is equivalent to requiring $\eta(\pi_\theta)$ to be non-decreasing in θ. Likewise, the envelope condition (EC) means that the menu gives each type a payoff equal to the weighted sum of the lower types' Youden indices.

Proposition 2 (Youden's Index and Implementation). *Suppose that* $D = \{\pi_\theta\}_{\theta\in\Theta} \cup \{\pi_0\}$ *is a binary menu that is individually rational.*

i. D satisfies the monotonicity condition (M) if and only if $\eta(\pi_\theta)$ *is non-decreasing in* θ*.*

ii. D satisfies the envelope condition (EC) if and only if $E_{\langle\pi_\theta|\theta\rangle}[U(\beta)] - C(\pi_\theta) = u\sum_{\theta_i<\theta}(\theta_{i+1} - \theta_i)\eta(\pi_{\theta_i})$ *for all* $\theta \in \Theta$*.*

4 Optimal Delegation

In Section 3, we showed that (without loss of optimality) the principal can restrict attention to deterministic mechanisms which *only* specify the experiment as a function of the agent's type, taking as given that the principal will approve if and only if it is ex-post optimal to do so. Now, we study the principal's problem of designing a menu of experiments for the agent to choose from. First, in Section 4.1, we show that an optimal menu exists, and characterize it. Sections 4.2 and 4.3 then show how asymmetric information makes the optimal menu inefficient.

4.1 Optimal Menus

By Proposition 1 and Theorem 1, we write the principal's program as the following constrained information design problem:

$$\max_{\{\pi_\theta\}_{\theta\in\Theta}\in\Pi^\Theta} \sum_{\theta\in\Theta} E_{\langle\pi_\theta|\theta\rangle}[W(\beta)]\sigma(\theta) \tag{COPT}$$

$$\text{s.t.} \quad \begin{array}{ll} (EC): & E_{\langle\pi_{\theta_n}|\theta_n\rangle}[U(\beta) - G(\beta|\theta_n)] = \sum_{i<n}(\theta_{i+1} - \theta_i)E_{\langle\pi_{\theta_i}|\theta_i\rangle}\left[\left(\frac{\beta-\theta_i}{\theta_i(1-\theta_i)}\right)U(\beta)\right] \;\; \forall n; \\ (M): & \begin{array}{l} E_{\langle\pi_{\theta_n}|\theta_n\rangle}\left[\left(\frac{\beta-\theta_n}{\theta_n(1-\theta_n)}\right)U(\beta)\right] \geq E_{\langle\pi_{\theta_{n-1}}|\theta_{n-1}\rangle}\left[\left(\frac{\beta-\theta_{n-1}}{\theta_{n-1}(1-\theta_{n-1})}\right)U(\beta)\right] \;\; \forall n > 0; \\ E_{\langle\pi_{\theta_0}|\theta_0\rangle}\left[\left(\frac{\beta-\theta_0}{\theta_0(1-\theta_0)}\right)U(\beta)\right] \geq 0. \end{array} \end{array}$$

A priori, (COPT) is a nonstandard constrained information design problem: The characterization results of Doval and Skreta (2023) cannot be immediately applied, since the constraint functions are not upper semi-continuous in posterior space. Nonetheless, Theorem 2 shows that we can write (COPT) as an unconstrained problem using the method of Lagrange.[23] As we show, the Lagrangian (COPT′) has a solution, and so (COPT) must as well.

Theorem 2 (Existence and Uniqueness). *A solution exists to* (COPT)*. This solution is unique up to Blackwell equivalence. Furthermore, there exists a family of non-negative Lagrange multipliers* $\{\lambda_n^*\}_{n=0}^N, \{\delta_n^*\}_{n=0}^N$ *such that* $D^* = \{\pi_\theta^*\}_{\theta\in\Theta} \cup \{\pi_0\}$ *solves* (COPT) *if and only if* $\{\pi_{\theta_n}^*\}_{n=0}^N$ *solves*

$$\max_{\{\pi_{\theta_n}\}_{n=0}^N \in \Pi^\Theta} \sum_{n=0}^N E_{\langle \pi_{\theta_n}|\theta_n\rangle}[W(\beta)]\sigma(\theta_n) \tag{COPT′}$$
$$+\lambda_n^* \left(E_{\langle \pi_{\theta_n}|\theta_n\rangle}[U(\beta) - G(\beta|\theta_n)] - \sum_{i<n} (\theta_{i+1} - \theta_i) E_{\langle \pi_{\theta_i}|\theta_i\rangle} \left[\left(\tfrac{\beta-\theta_i}{\theta_i(1-\theta_i)} \right) U(\beta) \right] \right)$$
$$+\delta_n^* \left(E_{\langle \pi_{\theta_n}|\theta_n\rangle} \left[\left(\tfrac{\beta-\theta_n}{\theta_n(1-\theta_n)} \right) U(\beta) \right] - E_{\langle \pi_{\theta_{n-1}}|\theta_{n-1}\rangle} \left[\left(\tfrac{\beta-\theta_{n-1}}{\theta_{n-1}(1-\theta_{n-1})} \right) U(\beta) \right] \mathbf{1}_{n>0} \right).$$

and $\pi_{\theta_n}^* \sim_B \pi_0$ *for each* n *with* $\lambda_n^* = 0$*. Furthermore, for all* n*,*

$$\delta_n^* \left(E_{\langle \pi_{\theta_n}^*|\theta_n\rangle} \left[\left(\tfrac{\beta-\theta_n}{\theta_n(1-\theta_n)} \right) U(\beta) \right] - E_{\langle \pi_{\theta_{n-1}}^*|\theta_{n-1}\rangle} \left[\left(\tfrac{\beta-\theta_{n-1}}{\theta_{n-1}(1-\theta_{n-1})} \right) U(\beta) \right] \mathbf{1}_{n>0} \right) = 0. \tag{CS}$$

Moreover, the maximized values of (COPT) *and* (COPT′) *are equal.*

By rearranging the terms in (COPT′), we can write the unique solution $D^* = \{\pi_\theta^*\}_{\theta\in\Theta} \cup \{\pi_0\}$ to (COPT) as the tuple of (unique) solutions to a set of type-by-type problems: For each n,

$$\pi_{\theta_n}^* \in \arg\max_{\pi\in\Pi} E_{\langle \pi|\theta_n\rangle} \left[\underbrace{W(\beta)\sigma(\theta_n) + \lambda_n^*(U(\beta) - G(\beta|\theta_n))}_{\text{Surplus Term } S(\beta,\theta_n)} - \underbrace{\rho_n^* \left(\frac{\beta - \theta_n}{\theta_n(1-\theta_n)} \right) U(\beta)}_{\text{Rent Term } R(\beta,\theta_n)} \right],$$
$$\text{where } \rho_n^* := \left(\delta_{n+1}^* + (\theta_{n+1} - \theta_n) \sum_{i=n+1}^N \lambda_i^* \right) \mathbf{1}_{n<N} - \delta_n^*. \tag{TBTθ_n}$$

These solutions thus maximize virtual surplus: that is, a weighted sum of the principal and agent's expected payoffs (the *surplus term* $E_{\langle \pi|\theta_n\rangle}[S(\beta, \theta_n)]$), minus a *rent term*

[23] See Amador and Bagwell (2013), Guo (2016), and Kartik et al. (2021) for other instances where the Lagrangian approach is useful in characterizing the solution to a delegation problem.

$E_{\langle\pi|\theta_n\rangle}[R(\beta,\theta_n)]$ (for a binary experiment, a scalar multiple of the Youden's index discussed in Section 3.3). Figure 1 depicts the objective function of these type-by-type problems. As we show in Section 4.3, the latter term distorts the problem's solution away from efficiency, except in the case of the highest type θ_N.

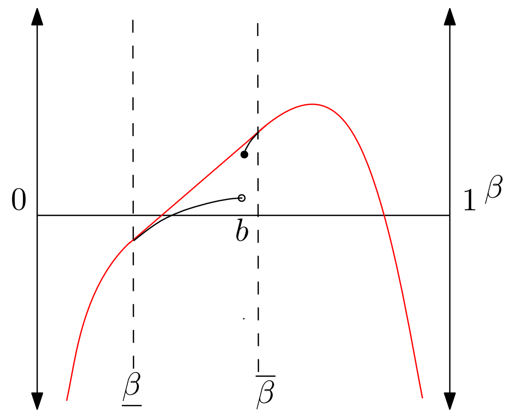

Figure 1: Solution to the Type-by-Type Problems. The type-by-type objective functions (black) are each strictly concave in β on $[0, b)$ and on $[b, 1]$, with a jump discontinuity at b.

Exclusion

The principal may find it optimal to *exclude* some types by not offering them an informative experiment that they are willing to conduct; i.e., we may have $\pi^*_\theta \sim_B \pi_0$ for some types. Lemma 3 shows that the set of *included types* $\tilde{\Theta} := \{\theta \in \Theta : \pi^*_\theta \not\sim_B \pi_0\}$ forms an upper set: that is, if it is optimal for the principal to exclude one type θ, it must also be optimal for her to exclude any type $\theta' \leq \theta$ with worse news about the state of the world.

Lemma 3 (Exclusion Is At The Low End). $\tilde{\Theta} = \{\theta \in \Theta : \theta \geq \underline{\theta}\}$ *for some* $\underline{\theta} \in \Theta$.

Intuitively, a binary experiment π is uninformative if and only if its Youden's index $\eta(\pi) = \pi(\bar{s}|1) - \pi(\bar{s}|0)$ is zero. Since (M) is equivalent to the monotonicity of the Youden's index, it follows that any implementable menu of binary experiments — including D^* — must exclude only at the low end of the distribution of private information.

The Monotonicity Constraint

By its nature, the monotonicity constraint (M) must bind on *intervals*: that is, if $(\mathrm{M}(\theta'', \theta'))$ binds for some $\theta'' > \theta'$, $(\mathrm{M}(\theta, \theta'))$ must also bind for all $\theta \in (\theta', \theta'')$. But unlike in standard contracting problems where the choice variable is one-dimensional, this does not guarantee that each type in such an interval conducts the same experiment — just that those experiments have the same Youden's index η. However, Lemma 4 shows that in the *optimal* menu D^*, if two types conduct experiments with the same Youden's index, the experiments must be identical.

Lemma 4 (Experiments Where Monotonicity Binds). *Suppose* $(M(\theta'', \theta'))$ *binds for some* $\theta'' > \theta'$. *Then for each* $\theta \in [\theta', \theta'']$, $\pi^*_\theta = \pi^*_{\theta''}$.

The intuition is as follows. Suppose monotonicity binds on the interval I, so that the experiments π^*_θ conducted by types $\theta \in I$ have identical Youden's indices η. Then on I, both sides of the envelope condition (ECθ) are linear in θ, with the same coefficient η. It follows that (ECθ) can only be satisfied on I if every experiment π^*_θ conducted by a type $\theta \in I$ yields the same payoff to any agent type. If multiple experiments have the same Youden's index and yield the same payoff to any agent type, the principal optimally includes in D^* the one with the lowest false positive rate.

True Positive Rates

Because any implementable menu must satisfy the monotonicity condition ($M(\theta, \theta')$), the difference between its experiments' true and false positive rates must be higher when those experiments are designed for higher types. However, we provide a stronger characterization regarding the solution to the principal's problem; in the optimal menu, the true positive rate is increasing in the agent's type. Since the principal and agent are aligned when the state of the world is high, the principal finds it most effective to design experiments for higher types with a larger true positive rate.

Proposition 3 (Increasing True Positive Rate). *If* $\theta' \geq \theta$, *then* $\pi^*_{\theta'}(\bar{s}|1) \geq \pi^*_\theta(\bar{s}|1)$. *Moreover, if* $\pi^*_{\theta'} \neq \pi^*_\theta$, *then* $\pi^*_{\theta'}(\bar{s}|1) > \pi^*_\theta(\bar{s}|1)$.

If $\pi^*_{\theta_n}$ has a higher true positive rate than $\pi^*_{\theta_{n+1}}$, it must be the case that $\pi^*_{\theta_n}$ is more expensive than $\pi^*_{\theta_{n+1}}$.[24] The principal would then have to compensate agents of type θ_n by giving $\pi^*_{\theta_n}$ a higher false positive rate than $\pi^*_{\theta_{n+1}}$, so that the former leads to approval in the bad state more frequently than the latter. But the continuity lemma (Lemma 6) shows that the principal can find a more effective way to give rents to agents of type θ_n by simultaneously reducing the true and false positive rates by equal amounts; this reduces costs while also reducing the probability of approval in the bad state. Incentive constraints are unchanged, since the Youden's index is left unchanged.

One might conjecture that, in the optimal menu, false positive rates are decreasing in the agent's type since the principal prefers to reject when $\omega = 0$ and higher type agents find $\omega = 0$ to be relatively less likely. This is not necessarily the case, however. To induce higher types to choose experiments with higher true positive rates, the principal may also raise the corresponding false positive rates. A higher false positive rate lowers the cost of the experiment and increases the probability of approval, both of which make the

[24]Otherwise, type θ_{n+1} would deviate to π^*_n, since it has a higher overall approval probability.

experiment more attractive to the agent. Nonetheless, we verify that in the optimal menu higher types always conduct more expensive experiments.

Corollary 1. *If $\theta' \geq \theta$, then $C(\pi^*_{\theta'}) \geq C(\pi^*_\theta)$. Moreover, if $\pi^*_{\theta'} \neq \pi^*_\theta$, then $C(\pi^*_{\theta'}) > C(\pi^*_\theta)$.*

Intuitively, since the true positive rate of $\pi^*_{\theta'}$ must exceed the true positive rate of π^*_θ, either $\pi^*_{\theta'}$ is Blackwell-more informative than π^*_θ, or $\pi^*_{\theta'}$ involves a higher false positive rate as well. In the former case, $\pi^*_{\theta'}$ is costlier than π^*_θ by the properties of C. In the latter case, to deter type θ from deviating to the experiment conducted by type θ', which gives a higher probability of approval in each state, $\pi^*_{\theta'}$ must be more expensive than π^*_θ.

Example (An Optimal Menu). The Lagrangian characterization of Theorem 2 allows us to numerically approximate the optimal menu. Since each experiment in the menu is binary, we can visualize the numerical solution to (COPT) on a two-dimensional graph, with false positive rate on one axis and true positive rate on the other. Figure 2 does so for a model with twenty types, evenly spaced on $[0.205, 0.485]$ with a uniform type distribution and parameters $u = 3$ and $b = 1 - w_1 = -w_0 = 0.5$; for other parameter values, see the widget linked here. We set $\theta_0 = 0.205$ to ensure (given the other parameters) that for every type, there is an informative experiment that satisfies their participation constraint.

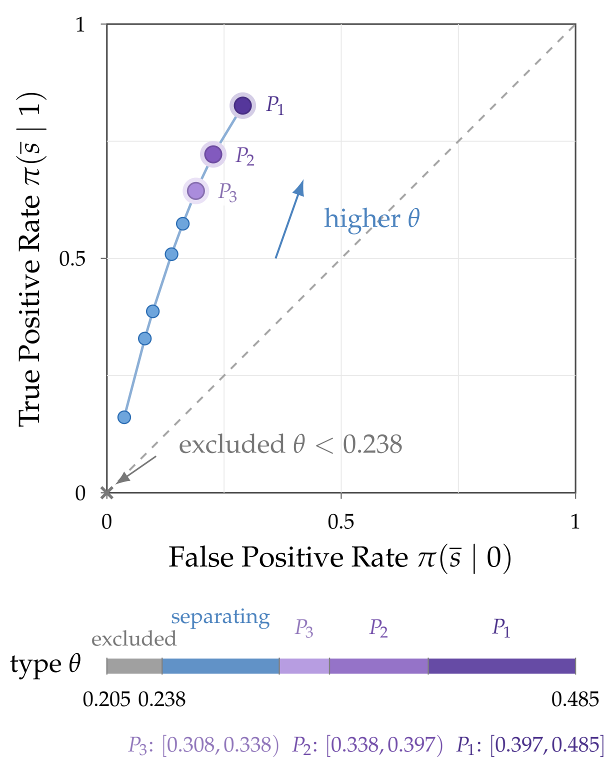

Figure 2: Simulated solution to (COPT). For twenty types, $u = 3$, and $b = 1/2$, the optimal menu excludes low types, separates an intermediate range of types, and pools high types on a sequence of experiments denoted by P_3, P_2, or P_1. The menu is not Blackwell monotone: higher types conduct experiments with higher true *and* false positive rates.

As we discuss above, it is not necessarily the case that the optimal menu is Blackwell monotone. While the principal induces a more expensive test from the more optimistic agents (Corollary 1), he directs this additional expenditure to raising the true positive rate (Proposition 3), and in fact redirects some resources from lowering the false positive rate. This is not a necessary consequence of the monotonicity constraint (M), which, in Figure 2, only requires that higher type experiments are on higher lines with slope -1 (Proposition 2(i)). Instead, the principal is more willing to trade a higher false positive rate for a higher true positive rate when the agent's private signal suggests the project is more likely to be good.

4.2 The Pareto Frontier

To see how experimentation is distorted by the need to induce the agent to voluntarily reveal his private information, we must first characterize what it is distorted away *from*. Hence, for each type of agent, we describe the set of experiments that are Pareto efficient, given the agent's type. Since transfers are not possible in our setting, these sets are not singletons. Instead, there is a *Pareto frontier* of efficient experiments for each type, with one end of that frontier being optimal for the principal, and the other being optimal for the agent.

Formally, an experiment π is *Pareto efficient for type* θ if there is no π' such that $E_{\langle \pi' | \theta \rangle}[W(\beta)] \geq E_{\langle \pi | \theta \rangle}[W(\beta)]$ and $E_{\langle \pi' | \theta \rangle}[U(\beta)] - C(\pi') \geq E_{\langle \pi | \theta \rangle}[U(\beta)] - C(\pi)$, with one of the inequalities strict. As one might expect, such experiments are precisely those that solve a social planner's problem.

Proposition 4 (Efficiency and the Social Planner's Problem). *π is Pareto efficient for type θ if and only if there exist $\lambda_p \geq 0$ and $\lambda_a \geq 0$, not both zero, such that*

$$\langle \pi | \theta \rangle \in \arg \max_{\tau \in \Delta(\Delta(\Omega))} \left\{ E_\tau \left[\lambda_p W(\beta) + \lambda_a (U(\beta) - G(\beta|\theta)) \right] \text{ s.t. } E_\tau \beta = \theta \right\}. \qquad \text{(SPP}\theta\text{)}$$

Together, Proposition 4 and Lemma OA.2 allow us to characterize the type-θ Pareto frontier geometrically. In particular, Lemma OA.2 shows that when $\lambda_a \neq 0$, (SPPθ) has a unique solution characterized by one of three alternatives. First, if there are posteriors $\underline{\beta}$ and $\overline{\beta}$ such that $\underline{\beta} < \theta < b < \overline{\beta}$ and

$$-\lambda_a G'(\underline{\beta}|\theta) = \lambda_p(w_1 - w_0) - \lambda_a G'(\overline{\beta}|\theta) = \frac{\lambda_p(w_0 + (w_1 - w_0)\overline{\beta}) + \lambda_a(u - G(\overline{\beta}|\theta)) + \lambda_a G(\underline{\beta}|\theta)}{\overline{\beta} - \underline{\beta}},$$

or equivalently, since $b \equiv \frac{-w_0}{w_1 - w_0}$,

$$\frac{\lambda_p(-w_0)}{\lambda_a b} = G'(\overline{\beta}|\theta) - G'(\underline{\beta}|\theta) \text{ and}$$
$$0 = \underbrace{u - G(\overline{\beta}|\theta) - G'(\overline{\beta}|\theta)(b - \overline{\beta})}_{\text{tangent line to } U - G(\cdot|\theta) \text{ at } \overline{\beta}} + \underbrace{G(\underline{\beta}|\theta) + G'(\underline{\beta}|\theta)(b - \underline{\beta})}_{\text{tangent line to } U - G(\cdot|\theta) \text{ at } \underline{\beta}}, \tag{3}$$

then (SPPθ)'s unique solution is given by the unique Bayes-plausible distribution that induces the posteriors $\underline{\beta}$ and $\overline{\beta}$. Geometrically, this condition means that the tangent lines to the agent's value function at $\underline{\beta}$ and $\overline{\beta}$ must cross at b, and the difference in their slopes (and since $G(\cdot|\theta)$ is strictly convex, the distance $\overline{\beta} - \underline{\beta}$) is increasing in the relative Pareto weight on the principal.[25] Figure 3 illustrates.

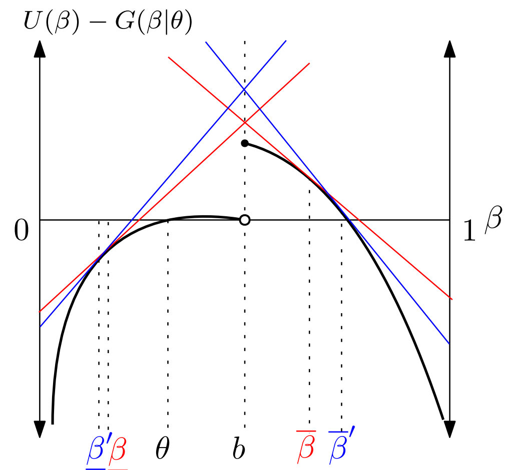

Figure 3: Characterization of the Pareto Frontier. Each experiment that is efficient for type θ induces a pair of posteriors such that the tangent lines to the agent's value function at those posteriors cross at b. When the Pareto weight on the principal is higher relative to the weight on the agent, those posteriors are further apart, and the tangent lines cross at a higher point: The experiment that induces $\underline{\beta}'$ and $\overline{\beta}'$ solves (SPPθ) for a higher value of λ_p/λ_a, while the experiment that induces $\underline{\beta}$ and $\overline{\beta}$ solves (SPPθ) for a lower value of λ_p/λ_a.

If the Pareto weight assigned to the principal is small enough relative to the Pareto weight on the agent, (3) becomes impossible to satisfy. Then, Lemma OA.2's second alternative[26] says that (3)'s unique solution is the unique Bayes-plausible distribution that

[25]To understand why, recall that W is continuous, affine to the right of the approval threshold b, and zero to the left of b. Then adding $\lambda_p/\lambda_a W(\beta)$ to the agent's value function (thus yielding the objective in (SPPθ)) essentially rotates the function — and thus its tangent lines — upward to the right of b. Consequently, the tangent lines to the value function in (SPPθ) at $\underline{\beta}$ and $\overline{\beta}$ *coincide* (as is necessary for the two posteriors to support a solution) exactly when tangent lines to the agent's value function $U - G(\cdot|\theta)$ *cross* at the approval threshold, and have slopes that differ by the relative weight λ_p/λ_a that the planner places on the principal multiplied by the slope of the principal's value function.

[26]Lemma OA.2's third alternative is ruled out: By assumption, there is an informative experiment that

induces the posterior beliefs $\underline{\beta}$ and b, where $\underline{\beta} < \theta$ is pinned down by

$$-G(\underline{\beta}|\theta) - G'(\underline{\beta}|\theta)(b - \underline{\beta}) = u - G(b|\theta), \tag{4}$$

Geometrically, this requires that at the approval threshold b, the agent's value function coincides with the tangent line to his value function at $\underline{\beta}$. Thus, in this case, (SPPθ)'s solution is precisely the distribution of posteriors that maximizes the agent's payoff when his type is common knowledge.

Corollary 2 summarizes.

Corollary 2. *π is Pareto efficient for type θ if and only if one of the following holds:*

i. $\text{supp}\langle\pi|\theta\rangle = \{\underline{\beta}, \overline{\beta}\}$, $\underline{\beta} < \theta < b < \overline{\beta}$, *and* (3) *holds for some* $\lambda_a, \lambda_p > 0$.

ii. $\text{supp}\langle\pi|\theta\rangle = \{\underline{\beta}, b\}$, $\underline{\beta} < \theta$, *and* (4) *holds.*

iii. *π is fully informative: $\pi \sim_B \pi_\infty$, where π_∞ is the binary experiment with $\pi_\infty(\overline{s}|1) = \pi_\infty(\underline{s}|0) = 1$.*

4.3 Distortion

Corollary 2 describes the experiments that are *on* each type's Pareto frontier. This allows us to describe how and why that type's experiment from the principal's optimal menu is distorted *away* from the frontier. In the case of the highest type, the answer is simple: it isn't. Thus, we arrive at the classical result that there is *no distortion at the top*.

Proposition 5 (No Distortion at the Top). *In the principal's optimal menu, the high type's experiment $\pi^*_{\theta_N}$ is Pareto efficient for type θ_N.*

The intuition is standard when the highest type's monotonicity constraint $M(\theta_N, \theta_{N-1})$ does not bind: Then, the Lagrange multiplier δ^*_N is zero, and so the type-by-type problem for the highest type coincides with the social planner's objective in (SPPθ), given appropriate Pareto weights.[27] If, on the other hand, the highest type's monotonicity constraint binds, the argument is more subtle: If $\pi^*_{\theta_N}$ was not on the Pareto frontier, then the principal can construct a strictly better implementable menu by replacing $\pi^*_{\theta_N}$ with some experiment π' for all types that conduct $\pi^*_{\theta_N}$. This can be done in a way that makes the principal strictly better off without affecting agent payoffs since $\pi^*_{\theta_N}$ is not on the Pareto frontier.

gives the agent a nonnegative payoff — and thus Pareto improves upon the uninformative experiment π_0.

[27] Specifically, for Pareto weights $\lambda_p = \sigma(\theta)$ and $\lambda_a = \lambda^*_\theta$.

Just like in optimal contracting problems with transfers, however, there *is* distortion for types other than the highest. When a type-θ agent chooses the experiment $\pi_{\theta'}$ designed for some other type $\theta' \neq \theta$, the principal shifts her interim belief to θ', rather than θ. Because of this, the type-θ incentive compatibility constraint — and hence the Lagrangian term in (COPT′) — contains an extra coefficient on $U(\beta)$ that depends on the posterior belief and is not present in the social planner's problem for *any* Pareto weights λ_a and λ_p. As Corollary 2 shows, this distorts the tangent line conditions in (3) and (4), and thus the experiment they pin down.

Corollary 3. *Suppose that* $\pi^*_{\theta_n} \not\sim_B \pi_0$. *Then* $\operatorname{supp}\langle \pi^*_{\theta_n} | \theta_n \rangle = \{\underline{\beta}, \overline{\beta}\}$, *where* $\underline{\beta} < \theta_n < b \leq \overline{\beta}$*;*

$$\underbrace{-G(\underline{\beta}|\theta_n) - G'(\underline{\beta}|\theta_n)(b - \underline{\beta})}_{\text{tangent line to } U - G(\cdot|\theta_n) \text{ at } \underline{\beta}} = \underbrace{u - G(\overline{\beta}|\theta_n) - G'(\overline{\beta}|\theta_n)(b - \overline{\beta})}_{\text{tangent line to } U - G(\cdot|\theta_n) \text{ at } \overline{\beta}} - \underbrace{R(b, \theta_n)/\lambda^*_n}_{\text{distortion term}}; \tag{5}$$

and when $\overline{\beta} > b$, $G'(\overline{\beta}|\theta_n) - G'(\underline{\beta}|\theta_n) = \frac{\sigma(\theta_n)(-w_0)}{\lambda^*_n b} + \frac{R_\beta(\overline{\beta}, \theta_n)}{\lambda^*_n}$.

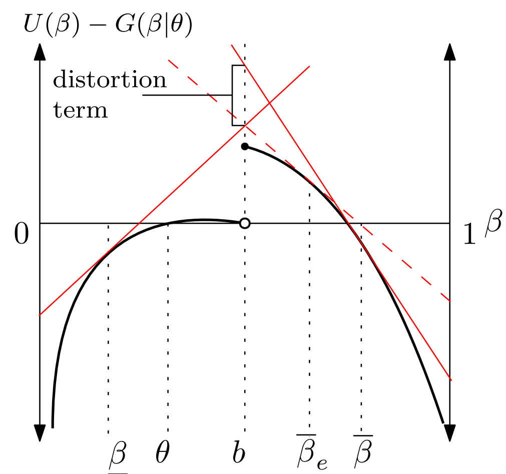

Figure 4: Distortion of the type-θ experiment. Suppose a binary experiment π_θ induces the posterior $\underline{\beta}$. For π^*_θ to be efficient for type θ, it would need to also induce $\overline{\beta}_e$: At the approval threshold b, the tangent line to the agent's value function at $\overline{\beta}_e$ crosses the tangent line to the value function at $\underline{\beta}$. But if π^*_θ solves the principal's type-by-type problem (TBTθ_n), it must induce $\overline{\beta}$ instead of $\overline{\beta}_e$, since the tangent lines to the value function at $\overline{\beta}_e$ and $\underline{\beta}$ differ by the distortion term at β.

Observe that there is an additional "distortion term" in (5) that is not present in (3). This ensures that instead of being characterized by tangent lines that *cross* at the approval threshold, the posteriors $\underline{\beta}$ and $\overline{\beta}$ that support $\langle \pi^*_\theta | \theta_n \rangle$ are characterized by tangent lines that *have a gap equal to the distortion term*. When type θ_n's downward monotonicity constraint $M(\theta_n, \theta_{n-1})$ does not bind, we can show that the distortion term is positive

(Lemma OA.5 in the appendix); consequently, $\underline{\beta}$ is inefficiently high given $\overline{\beta}$, and $\overline{\beta}$ is inefficiently high given $\underline{\beta}$. In fact, this is also true for types whose downward monotonicity constraint *does* bind, since their experiment from the optimal menu is the same as the highest type below them for whom such constraints do not bind (Lemma 4).[28]

This geometric characterization captures the key way that the type-θ agent's experiment π^*_θ from the optimal menu differs from *every* experiment that is Pareto efficient for him. But not all of these differences can be accurately described as *distortion*: Since there are no transfers, the fact that an experiment is efficient does not imply that it makes *both* the principal and the agent better off than π^*_θ. Hence, we focus on comparing the experiments from the optimal menu to *Pareto improving* experiments on the Pareto frontier.

Theorem 3 shows that when faced with the optimal menu, every included type other than the highest θ_N chooses an experiment whose true positive rate is too low. Intuitively, since each posterior induced by the optimal experiment is inefficiently high given the other, one can show that if π_{θ_n} Pareto improves upon $\pi^*_{\theta_n}$, it either (a) is Blackwell-more informative than $\pi^*_{\theta_n}$ or (b) corresponds to a *rightward shift* of the posteriors induced by $\pi^*_{\theta_n}$. In either case, it must have a higher true positive rate than $\pi^*_{\theta_n}$. Thus, private information results in an inefficiently low chance of approval in the state where the project is worth approving.

Theorem 3 (Distortion Everywhere Else). *Let $\{\pi^*_\theta\}_{\theta\in\Theta}$ be the principal's optimal menu.*

i. For each included type $\theta \in \tilde{\Theta}$ lower than the highest θ_N, π^*_θ is not Pareto efficient for type θ.

ii. For each included type $\theta \in \tilde{\Theta}$ other than the highest θ_N, π^*_θ has an inefficiently low true positive rate: For any π_θ on the Pareto frontier for type θ that Pareto-improves upon π^*_θ for type θ,

$$\pi_\theta(\bar{s}|1) > \pi^*_\theta(\bar{s}|1).$$

iii. Exclusion is inefficient: For every excluded type, there is a Pareto efficient experiment that Pareto improves upon the totally uninformative experiment π_0.

5 Conclusion

This paper studies the optimal delegation of experiments to a privately informed agent. Based on the characterization results, we offer three main conclusions.

[28] In Theorem 3 (i), we show something stronger: whether or not a type's downward monotonicity constraint binds, the principal's optimal menu yields an experiment with an inefficiently low false negative rate.

First, when the principal can commit ex-ante to an approval rule, she never finds it optimal to do so. The principal always incentivizes the agents through the design of experiments included in the menu alone. Moreover, since the optimal menu of experiments is binary, one can reduce the principal's choice over direct mechanisms from type reports to distributions over experiment/approval rule pairs to a simple choice over menus of binary experiments. This effectively turns a question of delegation with an infinite-dimensional allocation space to a simpler two-dimensional delegation problem.

Second, menus screen agent types by offering different payoffs in each state of the world. In our drug approval application, pharmaceutical companies with better news select clinical trials from menus that offer a higher expected payoff conditional on the drug being effective than their lower confidence counterparts. That is, the difference between the true and false positive rate (the Youden's index) must be increasing in the agent's type report. Moreover, in the optimal menu the true positive rate is increasing in the agent's report. That is, higher types conduct higher-powered tests.

Finally, in the optimal menu there is no distortion for the highest type. However, there is distortion for all other types. Lower types are either excluded, which is inefficient, or conduct an experiment with an inefficiently high false negative rate. This results in an inefficiently low approval rate when the state of the world is high. In the context of drug regulation, the optimal menu rejects more good drugs than is efficient.

Appendix

Lemma 5 (Continuity Lemma). *If $\alpha < b$, $((S,\pi),a) \in \Pi \times \mathcal{A}$, $h_1, h_0 \in \mathbb{R}$, and*

$$u\left[\int_S a(\pi,s)h_1 d\pi(s|1) + \int_S a(\pi,s)h_0 d\pi(s|0)\right] - C(\pi) \geq y \tag{6}$$

for some $y \in \mathbb{R}_+$, then there exists $(\pi', a') \in \Pi \times \mathcal{A}$ such that

(i) $\operatorname{supp}\pi' = \{\underline{s}, \bar{s}\}$, *$a'(\pi',\underline{s}) = 0$, and $a'(\pi',\bar{s}) = 1$.*

(ii) The principal prefers (π', a') to (π, a), and strictly if (6) is strict. That is,

$$w_1\pi'(\bar{s}|1)\alpha + w_0\pi'(\bar{s}|0)(1-\alpha) \geq \alpha\int_S a(\pi,s)w_1 d\pi(s|1) + (1-\alpha)\int_S a(\pi,s)w_0 d\pi(s|0),$$

and this inequality is strict if (6) is strict.

(iii) $u(h_1\pi'(\bar{s}|1) + h_0\pi'(\bar{s}|0)) - C(\pi') = y$.

(iv) $u(\pi'(\bar{s}|1) - \pi'(\bar{s}|0)) = \int_S ua(\pi,s)d\pi(s|1) - \int_S ua(\pi,s)d\pi(s|0)$.

Proof. First, we construct a pair $(\tilde{\pi}, \tilde{a})$ that satisfies (i) and (iv). Let $\tilde{\pi}$ be a binary experiment with $\tilde{\pi}(\bar{s}|\omega) = \int_S a(\pi, s)\pi(s|\omega)ds$ for each $\omega \in \{0, 1\}$. That is, $\tilde{\pi}$ generates the high outcome $\bar{s}$ in state ω with the same probability with which the principal approves a result in (π, a) in state ω. Define $\tilde{a}$ such that $\tilde{a}(\tilde{\pi}, \bar{s}) = 1$, $\tilde{a}(\tilde{\pi}, \underline{s}) = 0$, and $\tilde{a}(\pi, s) = 0$ for each $\pi \neq \tilde{\pi}$ and each s. It follows immediately that $(\tilde{\pi}, \tilde{a})$ satisfies (i) and (iv).

Next, observe that the principal is indifferent between (π, a) and $(\tilde{\pi}, \tilde{a})$ since the state-conditional approval probabilities are the same under (π, a) and $(\tilde{\pi}, \tilde{a})$. That is,

$$w_1\tilde{\pi}(\bar{s}|1)\alpha + w_0\tilde{\pi}(\bar{s}|0)(1-\alpha) = \alpha \int_S a(\pi, s)w_1 d\pi(s|1) + (1-\alpha)\int_S a(\pi, s)w_0 d\pi(s|0). \quad (7)$$

Moreover, observe that

$$\begin{aligned} u\left[h_1\tilde{\pi}(\bar{s}|1) + h_0\tilde{\pi}(\bar{s}|0)\right] - C(\tilde{\pi}) &\geq u\left[h_1\tilde{\pi}(\bar{s}|1) + h_0\tilde{\pi}(\bar{s}|0)\right] - C(\pi) \\ &= u\left[h_1\int_S a(\pi, s)d\pi(s|1) + h_0\int_S a(\pi, s)d\pi(s|0)\right] - C(\pi) \geq y, \end{aligned} \quad (8)$$

where the first inequality is from the fact that $C(\tilde{\pi}) \leq C(\pi)$ (since $\tilde{\pi}$ is a garbling of π and C is increasing in the Blackwell order), the equality follows from the definition of $\tilde{\pi}$, and the last inequality is strict whenever (6) is strict.

To proceed further, we first suppose that $\tilde{\pi}(\bar{s}|0) = 0$ or $\tilde{\pi}(\bar{s}|1) = 0$. Then either (a) $\tilde{\pi}(\bar{s}|0) = \tilde{\pi}(\bar{s}|1) = 0$, or (b) there exists $s \in \operatorname{supp} \pi(\cdot|\omega) \setminus \operatorname{supp} \pi(\cdot|1-\omega)$ for some $\omega \in \{0, 1\}$. In the latter case (b), $C(\pi) = \infty$, contradicting (6). Suppose instead that (a) holds. Then by construction, we must have $a(\pi, s) = 0$ for all $s \in \operatorname{supp} \pi(\cdot|1) \cup \operatorname{supp} \pi(\cdot|0)$. Then since $y \geq 0$, (6) must hold with equality, and $C(\pi) = y = 0$. And by (7), $\tilde{\pi}$ satisfies (ii). Since $\tilde{\pi} \sim_B \pi_0$, $C(\tilde{\pi}) = 0$; then $\tilde{\pi}$ satisfies (iii), and the claim follows by letting $\pi' = \tilde{\pi}$.

Now suppose instead that $\tilde{\pi}(\bar{s}|0) > 0$ and $\tilde{\pi}(\bar{s}|1) > 0$. We then construct, from $(\tilde{\pi}, \tilde{a})$, a pair (π', a') such that (i),(ii),(iii), and (iv) hold. To this end, observe that a binary $\hat{\pi}$ satisfies (iv) iff

$$\hat{\pi}(\bar{s}|1) = \phi(\hat{\pi}(\bar{s}|0)) := \hat{\pi}(\bar{s}|0) + \tilde{\pi}(\bar{s}|1) - \tilde{\pi}(\bar{s}|0). \quad (9)$$

Let $P = [0, \tilde{\pi}(\bar{s}|0)] \cap \phi^{-1}([0, 1])$ be the set of $p \leq \tilde{\pi}(\bar{s}|0)$ such that there is a binary experiment $\hat{\pi}$ satisfying (iv) with $\hat{\pi}(\bar{s}|1) = \phi(p)$ and $\hat{\pi}(\bar{s}|0) = p$; then

$$P = \left[\max\left\{0, \tilde{\pi}(\bar{s}|0) - \tilde{\pi}(\bar{s}|1)\right\}, \tilde{\pi}(\bar{s}|0)\right],$$

and $\min P < \tilde{\pi}(\bar{s}|0)$. Now define

$$f(p) = u(h_1\phi(p) + h_0 p) - \left(\begin{array}{c} \log\left(\frac{\phi(p)}{p}\right)\phi(p) + \log\left(\frac{1-\phi(p)}{1-p}\right)(1-\phi(p)) \\ +\log\left(\frac{p}{\phi(p)}\right)p + \log\left(\frac{1-p}{1-\phi(p)}\right)(1-p) \end{array} \right)$$
$$= u(h_1\phi(p) + h_0 p) - (\phi(p) - p)\log\left(\frac{\phi(p)(1-p)}{p(1-\phi(p))}\right).$$

Observe that if a binary $\hat{\pi}$ satisfies (9), then $f(\hat{\pi}(\bar{s}|0)) = u(h_1\hat{\pi}(\bar{s}|1) + h_0\hat{\pi}(\bar{s}|0)) - C(\hat{\pi})$. Since $\tilde{\pi}$ satisfies (9), it follows that $f(\tilde{\pi}(\bar{s}|0)) \geq y$, and that this inequality is strict if (6) is strict. Moreover, observe that either $\min P = 0$ (if $\int_S a(\pi,s)(\pi(s|0) - \pi(s|1))ds \geq 0$), or $\phi(\min P) = 0$ (otherwise); in either case, $f(\min P) = -\infty$. Then since P is connected and f is continuous, it follows that there exists $p' \in P$ such that $f(p') = y$, and that $p' < \tilde{\pi}(\bar{s}|0)$ if (6) is strict.

Then define π' as the binary experiment with $(\pi'(\bar{s}|1), \pi'(\bar{s}|0)) = (\phi(p'), p')$, and define a' such that $a'(\pi', \bar{s}) = 1$, $a'(\pi', \underline{s}) = 0$, $a'(\pi, s) = 0$ for each $\pi \neq \pi'$ and each s. By construction, π' satisfies (i), (iii), and (iv); it remains to show (ii).

Observe that for any binary experiment $\hat{\pi}$ satisfying (iv),

$$w_1\hat{\pi}(\bar{s}|1)\alpha + w_0\hat{\pi}(\bar{s}|0)(1-\alpha) = (\alpha w_1 + (1-\alpha)w_0)\hat{\pi}(\bar{s}|0) + w_1\alpha\int_S a(\pi,s)(\pi(s|1) - \pi(s|0))ds.$$

It follows that

$$\begin{array}{r} w_1\pi'(\bar{s}|1)\alpha + w_0\pi'(\bar{s}|0)(1-\alpha) \\ -(w_1\tilde{\pi}(\bar{s}|1)\alpha + w_0\tilde{\pi}(\bar{s}|0)(1-\alpha)) \end{array} = (\alpha w_1 + (1-\alpha)w_0)(p' - \tilde{\pi}(\bar{s}|0)) \geq 0,$$

since $\alpha < b$ and $p' \leq \tilde{\pi}(\bar{s}|0)$, and that if (6) is strict, then the inequality is strict since $p' < \tilde{\pi}(\bar{s}|0)$. (ii) then follows from (7). □

Corollary 4. *There exists $(\pi', a') \in \Pi \times \mathcal{A}$ such that π' is binary, $a'(\pi', \bar{s}) = 1$ and $a'(\pi', \underline{s}) = 0$, and for each $\theta \in \Theta$, $u(\theta\pi'(\bar{s}|1) + (1-\theta)\pi'(\bar{s}|0)) - C(\pi') > 0$ and $w_1\pi'(\bar{s}|1)\theta + w_0\pi'(\bar{s}|0)(1-\theta) > 0$.*

Proof. By assumption, there exists $(\pi, a) \in \Pi \times \mathcal{A}$ such that $\int_S (a(\pi,s)u - C(\pi))d\pi(s|\theta_0) > 0$ and $\theta_0 \int_S a(\pi,s)w_1 d\pi(s|1) + (1-\theta_0)\int_S a(\pi,s)w_0 d\pi(s|0) > 0$. Since $w_0 < 0 < w_1$, we have $\int_S a(\pi,s)d\pi(s|1) > \int_S a(\pi,s)d\pi(s|0)$. Then $\int_S (a(\pi,s)u - C(\pi))d\pi(s|\theta) > 0$ and $\theta\int_S a(\pi,s)w_1 d\pi(s|1) + (1-\theta)\int_S a(\pi,s)w_0 d\pi(s|0) > 0$ for all $\theta \in \Theta$. The claim then follows by Lemma 5 with $h_1 = \theta$, $h_0 = 1-\theta$, and $\alpha = \theta$. □

Lemma 6 (Characterizing Incentive Compatibility). *Let $\chi : \Theta \to \Pi \times \mathcal{A}$ be a deterministic* *mechanism, with* $\chi(\theta) = (\pi_\theta, a_\theta)$. *For each* $\theta \in \Theta$ *and* $\omega \in \{0,1\}$, *let*

$$q(\theta,\omega) := \int_{S_\theta} a_\theta(\pi_\theta, s) d\pi_\theta(s|\omega), \qquad r(\theta) := q(\theta,1) - q(\theta,0),$$

where S_θ *is the signal space of* π_θ, *and let*

$$V(\theta) := u\left[\theta q(\theta,1) + (1-\theta)q(\theta,0)\right] - C(\pi_\theta)$$

denote type θ*'s truthful payoff. Then:*

(i) If χ *satisfies the envelope condition*

$$V(\theta) = \sum_{\theta_i < \theta} (\theta_{i+1} - \theta_i) u r(\theta_i) \quad \text{for every } \theta, \tag{EC-FC}$$

and the monotonicity condition

$$r(\theta) \geq r(\theta') \quad \text{for every } \theta' \leq \theta, \tag{M-FC}$$

then χ *is incentive compatible. If, in addition,*

$$r(\theta_0) \geq 0, \tag{NL}$$

then χ *is implementable.*

(ii) If χ *is implementable, then there exists an implementable deterministic mechanism* χ' *satisfying* (EC-FC), (M-FC), *and* (NL) *such that each* π'_θ *is binary,* $a'_\theta(\pi'_\theta, \bar{s}) = 1$, $a'_\theta(\pi'_\theta, \underline{s}) = 0$, *and which gives the principal a weakly higher payoff type-by-type:*

$$\theta \pi'_\theta(\bar{s}|1) w_1 + (1-\theta)\pi'_\theta(\bar{s}|0) w_0 \geq \theta q(\theta,1) w_1 + (1-\theta) q(\theta,0) w_0$$

for every $\theta \in \Theta$.

Proof. **(i):** Suppose that χ satisfies (EC-FC) and (M-FC). For any report θ_m, type θ_n obtains

$$V(\theta_m) + (\theta_n - \theta_m) u r(\theta_m).$$

Thus, when $n > m$, the gain from truth-telling over reporting θ_m is

$$V(\theta_n) - V(\theta_m) - (\theta_n - \theta_m) u r(\theta_m) = \sum_{i=m}^{n-1} (\theta_{i+1} - \theta_i) u (r(\theta_i) - r(\theta_m)) \geq 0,$$

where the equality uses (EC-FC) and the inequality uses (M-FC). Similarly, when $m > n$,

$$V(\theta_n) - V(\theta_m) - (\theta_n - \theta_m)ur(\theta_m) = \sum_{i=n}^{m-1}(\theta_{i+1} - \theta_i)u(r(\theta_m) - r(\theta_i)) \geq 0.$$

Hence, χ is incentive compatible.

If, additionally, (NL) holds, then since (EC-FC) gives $V(\theta_0) = 0$, and since (M-FC) and (NL) imply $r(\theta_i) \geq 0$ for all i, every type's payoff is nonnegative; it follows that χ is implementable.

(ii): Suppose χ is implementable. Then by incentive compatibility, for any $\theta_n > \theta_m$,

$$V(\theta_n) \geq V(\theta_m) + (\theta_n - \theta_m)ur(\theta_m), \qquad \text{and} \qquad V(\theta_m) \geq V(\theta_n) + (\theta_m - \theta_n)ur(\theta_n).$$

Combining these inequalities yields

$$ur(\theta_n) \geq \frac{V(\theta_n) - V(\theta_m)}{\theta_n - \theta_m} \geq ur(\theta_m),$$

so r is monotone.

Let $T(\theta_0) = 0$ and, for $n > 0$,

$$T(\theta_n) := \sum_{i<n}(\theta_{i+1} - \theta_i)ur(\theta_i).$$

Since χ is individually rational, $V(\theta_0) \geq T(\theta_0) = 0$. Since χ is incentive compatible, $V(\theta_n) \geq V(\theta_{n-1}) + (\theta_n - \theta_{n-1})ur(\theta_{n-1})$ for all $n > 0$. Combining these inductively implies $V(\theta_n) \geq T(\theta_n)$ for every n. Then for each n, by Lemma 5 (with $\alpha = \theta_n < b$, $y = T(\theta_n)$, $h_1 = \theta_n$, and $h_0 = 1 - \theta_n$), we can construct $\chi'(\theta_n) = (\pi'_{\theta_n}, a'_{\theta_n})$ such that π'_{θ_n} is binary, $a'_{\theta_n}(\pi'_{\theta_n}, \bar{s}) = 1$ and $a'_{\theta_n}(\pi'_{\theta_n}, \underline{s}) = 0$, and

$$\pi'_{\theta_n}(\bar{s}|1) - \pi'_{\theta_n}(\bar{s}|0) = r(\theta_n); \qquad u[\theta_n\pi'_{\theta_n}(\bar{s}|1) + (1-\theta_n)\pi'_{\theta_n}(\bar{s}|0)] - C(\pi'_{\theta_n}) = T(\theta_n); \tag{10}$$

$$w_1\pi'_{\theta_n}(\bar{s}|1)\theta_n + w_0\pi'_{\theta_n}(\bar{s}|0)(1-\theta_n) \geq w_1 q(\theta_n, 1)\theta_n + w_0 q(\theta_n, 0)(1-\theta_n).$$

It remains to show that χ' is implementable. By (10), χ' satisfies (EC-FC), (M-FC) (since r is monotone), and (NL) (since $T(\theta_0) = 0$). Thus, by part (i), χ' is implementable. □

Proof of Theorem 1 By Lemma OA.4, we can restrict attention to deterministic mechanisms $\chi : \Theta \to \Pi \times \mathcal{A}$ denoted by $\chi(\theta) = (\pi_\theta, a_\theta)$. By Lemma 6(ii), it is further without loss to restrict attention to mechanisms such that each π_θ is binary, $a_\theta(\pi_\theta, \bar{s}) = 1$, $a_\theta(\pi_\theta, \underline{s}) = 0$, and (EC-FC), (M-FC), and (NL) hold. Let $\mathcal{B}$ denote the set of all such mechanisms. We prove the following claim:

Claim T1.1: For any $\chi \in \mathcal{B}$, if there is some θ such that a_θ is not ex-post optimal for π_θ, then there exists another implementable $\chi' \in \mathcal{B}$ which yields a strictly higher value of the objective in (OPT). Suppose that for such a mechanism χ, there exists some $\theta \in \Theta$ such that a_θ is not ex-post optimal. Then either $w_1\pi_\theta(\bar{s}|1)\theta + w_0\pi_\theta(\bar{s}|0)(1-\theta) < 0$ or $\underline{s} \in \operatorname{supp} \pi_\theta(\cdot|1) \cup \operatorname{supp} \pi_\theta(\cdot|0)$ and $w_1\pi_\theta(\underline{s}|1)\theta + w_0\pi_\theta(\underline{s}|0)(1-\theta) \geq 0$. Suppose that the latter is true. Since (M-FC) and (NL) imply $\pi_\theta(\bar{s}|1) \geq \pi_\theta(\bar{s}|0)$, we have

$$\theta w_1 \pi_\theta(\underline{s}|1) + (1-\theta) w_0 \pi_\theta(\underline{s}|0) \leq [\theta w_1 + (1-\theta) w_0] \pi_\theta(\underline{s}|0) \leq 0,$$

with equality only if $\underline{s} \notin \operatorname{supp} \pi_\theta(\cdot|1) \cup \operatorname{supp} \pi_\theta(\cdot|0)$. It follows that we must have

$$\theta w_1 \pi_\theta(\bar{s}|1) + (1-\theta) w_0 \pi_\theta(\bar{s}|0) < 0. \tag{11}$$

Given such a mechanism χ, we recursively construct a sequence of mechanisms $\chi^0, \ldots, \chi^{N+1}$ as follows. Let $\chi^0 = \chi$, and for each χ^n, write q^n, r^n, V^n, and T^n for the corresponding objects from Lemma 6. For each $n \in \{0, \ldots, N\}$, suppose χ^n satisfies (EC-FC), (M-FC), and (NL), and that

$$\theta_m w_1 \pi^n_{\theta_m}(\bar{s}|1) + (1-\theta_m) w_0 \pi^n_{\theta_m}(\bar{s}|0) \geq 0$$

for every $m < n$. If the same inequality also holds for $m = n$, set $\chi^{n+1} = \chi^n$. If not, then construct χ^{n+1} as follows.

- If $n = 0$, set $\chi^1(\theta_0) = (\pi_0, a^1_{\theta_0})$, where $a^1_{\theta_0}(s, \pi) = 0$ for all $s \in S$, $\pi \in \Pi$. Then

$$r^1(\theta_0) = 0 \leq r^0(\theta_0), \qquad V^1(\theta_0) = T^1(\theta_0) = 0.$$

- If $n > 0$, set $\chi^{n+1}(\theta_m) = \chi^n(\theta_m)$ for every $m < n$, and set $\chi^{n+1}(\theta_n) = \chi^n(\theta_{n-1})$. Then

$$V^{n+1}(\theta_n) = V^n(\theta_{n-1}) + (\theta_n - \theta_{n-1}) u r^n(\theta_{n-1}) = T^n(\theta_n) = T^{n+1}(\theta_n),$$

  and, using the induction hypothesis for $m = n-1$,

$$\theta_n w_1 \pi^{n+1}_{\theta_n}(\bar{s}|1) + (1-\theta_n) w_0 \pi^{n+1}_{\theta_n}(\bar{s}|0) = \begin{array}{c} \theta_{n-1} w_1 \pi^n_{\theta_{n-1}}(\bar{s}|1) + (1-\theta_{n-1}) w_0 \pi^n_{\theta_{n-1}}(\bar{s}|0) \\ +(\theta_n - \theta_{n-1}) \left[w_1 \pi^n_{\theta_{n-1}}(\bar{s}|1) - w_0 \pi^n_{\theta_{n-1}}(\bar{s}|0) \right] \end{array} \geq 0.$$

  Observe that $r^{n+1}(\theta_i) \leq r^n(\theta_i)$ for each $i \leq n$: $r^{n+1}(\theta_i) = r^n(\theta_i)$ for $i < n$, and $r^{n+1}(\theta_n) = r^n(\theta_{n-1}) \leq r^n(\theta_n)$ since χ^n satisfies (M-FC).

- For each $m > n$, define $\chi^{n+1}(\theta_m)$ recursively as follows: Given $\{\chi^{n+1}(\theta_i)\}_{i<m}$ such

that $r^{n+1}(\theta_i) \leq r^n(\theta_i)$ for each $i < m$, observe that $T^{n+1}(\theta_m) \leq T^n(\theta_m) = V^n(\theta_m)$. Then applying Lemma 5 to $\chi^n(\theta_m)$ with $\alpha = \theta_m$, $y = T^{n+1}(\theta_m)$, $h_1 = \theta_m$, and $h_0 = 1 - \theta_m$ yields $\chi^{n+1}(\theta_m) = (\pi^{n+1}_{\theta_m}, a^{n+1}_{\theta_m})$ such that $\pi^{n+1}_{\theta_m}$ is binary, $a^{n+1}_{\theta_m}(\pi^{n+1}_{\theta_m}, \bar{s}) = 1$ and $a^{n+1}_{\theta_m}(\pi^{n+1}_{\theta_m}, \underline{s}) = 0$, and

$$V^{n+1}(\theta_m) = T^{n+1}(\theta_m), \qquad r^{n+1}(\theta_m) = r^n(\theta_m),$$
$$\theta_m w_1 \pi^{n+1}_{\theta_m}(\bar{s}|1) + (1-\theta_m) w_0 \pi^{n+1}_{\theta_m}(\bar{s}|0) \geq \theta_m w_1 \pi^n_{\theta_m}(\bar{s}|1) + (1-\theta_m) w_0 \pi^n_{\theta_m}(\bar{s}|0).$$

By construction, χ^{n+1} satisfies (EC-FC). Since r^{n+1} is obtained from r^n by replacing either $r^n(\theta_0)$ with 0 or $r^n(\theta_n)$ with $r^n(\theta_{n-1})$, and $r^{n+1}(\theta) = r^n(\theta)$ for each $\theta \neq \theta_n$, χ^{n+1} satisfies (M-FC) and (NL). Then by Lemma 6(i), χ^{n+1} is implementable. Moreover,

$$\theta_m w_1 \pi^{n+1}_{\theta_m}(\bar{s}|1) + (1-\theta_m) w_0 \pi^{n+1}_{\theta_m}(\bar{s}|0) \geq \theta_m w_1 \pi^n_{\theta_m}(\bar{s}|1) + (1-\theta_m) w_0 \pi^n_{\theta_m}(\bar{s}|0)$$

for every m, with strict inequality for $m = n$ whenever $\chi^{n+1} \neq \chi^n$. Now let $\chi' = \chi^{N+1}$. By induction, χ' is implementable and

$$\theta_m w_1 \pi'_{\theta_m}(\bar{s}|1) + (1-\theta_m) w_0 \pi'_{\theta_m}(\bar{s}|0) \geq 0$$

for every m. By (11), $\chi^{n+1} \neq \chi^n$ for some n; hence, since $\sigma(\theta) > 0$ for all $\theta \in \Theta$,

$$\sum_{m=0}^{N} \sigma(\theta_m) \left[\theta_m w_1 \pi'_{\theta_m}(\bar{s}|1) + (1-\theta_m) w_0 \pi'_{\theta_m}(\bar{s}|0)\right]$$
$$> \sum_{m=0}^{N} \sigma(\theta_m) \left[\theta_m w_1 \pi^0_{\theta_m}(\bar{s}|1) + (1-\theta_m) w_0 \pi^0_{\theta_m}(\bar{s}|0)\right],$$

proving Claim T1.1. ■

It follows immediately from Lemma OA.4, Lemma 6(ii), and Claim T1.1 that if (OPT) admits a solution, then it has a deterministic solution χ^* with ex-post optimal decision rules, as desired. □

Lemma 7. *Let $\theta \in \Theta$ and suppose that π is binary. Then*

i. $\langle \pi|\alpha\rangle([b,1]) = \begin{cases} \alpha\eta(\pi) + \pi(\bar{s}|0), & \frac{\pi(\bar{s}|1)\alpha}{\eta(\pi)\alpha + \pi(\bar{s}|0)} \geq b; \\ 0, & \textit{otherwise.} \end{cases}$

ii. $E_{\langle \pi|\alpha\rangle}\left[\left(\frac{\theta}{\alpha}\beta + \frac{1-\theta}{1-\alpha}(1-\beta)\right) U(\beta)\right] = \begin{cases} (\theta\eta(\pi) + \pi(\bar{s}|0))u, & \frac{\pi(\bar{s}|1)\alpha}{\eta(\pi)\alpha + \pi(\bar{s}|0)} \geq b; \\ 0, & \textit{otherwise.} \end{cases}$.

iii. *If* $E_{\langle \pi|\alpha\rangle}\left[\left(\frac{\theta}{\alpha}\beta + \frac{1-\theta}{1-\alpha}(1-\beta)\right)U(\beta)\right] - C(\pi) \geq 0$ *and* $\pi \not\sim_B \pi_0$, *then* $\text{supp}\langle \pi|\alpha\rangle \cap [b,1] = \left\{\frac{\pi(\bar{s}|1)\alpha}{\alpha\eta(\pi)+\pi(\bar{s}|0)}\right\}$;

iv. $E_{\langle \pi|\alpha\rangle}\left[\left(\frac{\beta-\alpha}{\alpha(1-\alpha)}\right)U(\beta)\right] = \begin{cases} \eta(\pi), & \text{if } \frac{\pi(\bar{s}|1)\alpha}{\pi(\bar{s}|1)\alpha+\pi(\bar{s}|0)(1-\alpha)} \geq b; \\ 0, & \text{otherwise.} \end{cases}$.

Proof. (i) follows immediately from Lemma 1 and by Bayes' rule. (ii) then follows from (i). (iii) follows from (ii) since $C(\pi) > 0$ for all $\pi \not\sim_B \pi_0$ and by Bayes' rule. Finally, when $\frac{\pi(\bar{s}|1)\alpha}{\eta(\pi)\alpha+\pi(\bar{s}|0)} \geq b$, we have

$$E_{\langle \pi|\alpha\rangle}\left[\left(\frac{\beta-\alpha}{\alpha(1-\alpha)}\right)U(\beta)\right] = (\alpha\eta(\pi) + \pi(\bar{s}|0))\left(\frac{\frac{\pi(\bar{s}|1)\alpha}{\alpha\eta(\pi)+\pi(\bar{s}|0)} - \alpha}{\alpha(1-\alpha)}\right)u$$
$$= \left(\frac{\pi(\bar{s}|1) - \alpha\eta(\pi) - \pi(\bar{s}|0)}{1-\alpha}\right) = \eta(\pi),$$

as desired. □

Proof of Proposition 1 (Implementable Menus) Follows immediately from Lemma 6(i).

Proof of Lemma 2 (Menus Are Binary Without Loss) Follows immediately from Lemma 6(ii).

Proof of Proposition 2 (Youden's Index and Implementation) Let $D = \{\pi_\theta\}_{\theta\in\Theta} \cup \{\pi_0\}$ be a binary menu that is individually rational. Then each π_θ is binary and satisfies (IRθ). Claims (i) and (ii) thus follow immediately from Lemma 7 (iii) and (iv). □

In what follows, we consider problems where the principal is restricted to binary experiments which result in approval after a positive result $\bar{s}$. To this end, denote the agent's cost of performing a binary experiment π by $C_b : [0,1]^2 \to \mathbb{R}_+ \cup \{\infty\}$, where

$$C_b(\pi(\bar{s}|1), \pi(\bar{s}|0)) := \begin{pmatrix} \log\left(\frac{\pi(\bar{s}|1)}{\pi(\bar{s}|0)}\right)\pi(\bar{s}|1) + \log\left(\frac{1-\pi(\bar{s}|1)}{1-\pi(\bar{s}|0)}\right)(1-\pi(\bar{s}|1)) \\ +\log\left(\frac{\pi(\bar{s}|0)}{\pi(\bar{s}|1)}\right)\pi(\bar{s}|0) + \log\left(\frac{1-\pi(\bar{s}|0)}{1-\pi(\bar{s}|1)}\right)(1-\pi(\bar{s}|0)) \end{pmatrix}$$
$$= (\pi(\bar{s}|1) - \pi(\bar{s}|0))\log\left(\frac{\pi(\bar{s}|1)(1-\pi(\bar{s}|0))}{\pi(\bar{s}|0)(1-\pi(\bar{s}|1))}\right),$$

if $\pi(\bar{s}|1) \neq \pi(\bar{s}|0)$, and $C_b(\pi(\bar{s}|1), \pi(\bar{s}|0)) := 0$ otherwise. Further, let

$$U_b(x,y,\theta) = (\theta(x-y)+y)u - C_b(x,y); \qquad W_b(x,y,\theta) = \theta w_1 x + (1-\theta)w_0 y.$$

Lemma 8. *There exists a family of non-negative Lagrange multipliers* $\{\lambda_n^*\}_{n=0}^N$ *and* $\{\delta_n^*\}_{n=0}^N$ *such that* $\{(x_n^*, y_n^*)\}_{n=0}^N$ *solves*

$$\max_{\{(x_n,y_n)\}_{n=0}^N \in [0,1]^{2(N+1)}} \sum_{n=0}^N W_b(x_n, y_n, \theta_n)\sigma(\theta_n) \tag{OPT2}$$

$$\text{s.t. } U_b(x_n, y_n, \theta_n) = \sum_{i<n}(\theta_{i+1} - \theta_i)(x_i - y_i)u \quad \forall n;$$

$$x_n - y_n \geq x_{n-1} - y_{n-1} \quad \forall n > 0; \qquad x_0 - y_0 \geq 0,$$

if and only if it solves

$$\max_{\{(x_n,y_n)\}_{n=0}^N \in [0,1]^{2(N+1)}} \sum_{n=0}^N W_b(x_n, y_n, \theta_n)\sigma(\theta_n) + \lambda_n^* \left(U_b(x_n, y_n, \theta_n) - \sum_{i<n}(\theta_{i+1} - \theta_i)(x_i - y_i)u \right)$$
$$+ \delta_n^* \left((x_n - y_n) - (x_{n-1} - y_{n-1})\right) \mathbf{1}_{n>0} + \delta_0^*(x_0 - y_0) \tag{OPT2'}$$

and satisfies the complementary slackness conditions

$$\lambda_n^* \left(U_b(x_n, y_n, \theta_n) - \sum_{i<n}(\theta_{i+1} - \theta_i)(x_i - y_i)u \right) = 0 \quad \forall n$$
$$\delta_n^* \left((x_n - y_n) - (x_{n-1} - y_{n-1})\right) = 0 \quad \forall n > 0; \quad \delta_0^*(x_0 - y_0) = 0. \tag{CS2}$$

Moreover, this solution exists and is unique, and $(x_n^*, y_n^*) = (0, 0)$ *for every* n *such that* $\lambda_n^* = 0$.

Proof. By Lemma 5, we can relax (OPT2) to allow the envelope constraints to hold with weak inequality: For any $\{(x_n, y_n)\}_{n=0}^N \in [0,1]^{2(N+1)}$ such that $x_n - y_n \geq x_{n-1} - y_{n-1}$ for all $n > 0$ and $(\theta_n x_n + (1-\theta_n)y_n)u - C_b(x_n, y_n) \geq \sum_{i<n}(\theta_{i+1} - \theta_i)(x_i - y_i)u$ for all n, there is a $\{(x_n', y_n')\}_{n=0}^N \in [0,1]^{2N}$ that is feasible in (OPT2) and achieves a weakly greater value of the objective function.

Moreover, we can restrict attention to a convex, compact region of $[0,1]^{2N}$ in which the constraint functions are finite: By (EC), $(\theta_n x_n + (1-\theta_n)y_n)u - C_b(x_n, y_n) \geq -(\theta_n - \theta_0)u$ for every feasible (x_n, y_n). Hence, we can restrict attention to $\{(x_n, y_n)\}_{n=0}^N \in B := \prod_{n=0}^N B_n$, where $B_n = \{(x, y) \in [0,1]^2 : (\theta_n x + (1-\theta_n)y)u - C_b(x, y) \geq -u\}$. Since C_b is convex, the function $(x, y) \mapsto (\theta_n x + (1-\theta_n)y)u - C_b(x, y)$ is concave, and so each B_n is convex. Hence B is convex. Moreover, each B_n is closed and bounded, and so B is compact.

Then the solutions of (OPT2) coincide with the solutions to the following program:

$$\max_{\{(x_n,y_n)\}_{n=0}^N \in B} \sum_{n=0}^{N} W_b(x_n,y_n,\theta_n)\sigma(\theta_n) \tag{OPT2R}$$
$$\text{s.t. } U_b(x_n,y_n,\theta_n) \geq \sum_{i<n}(\theta_{i+1}-\theta_i)(x_i-y_i)u \quad \forall n$$
$$x_n - y_n \geq x_{n-1} - y_{n-1} \quad \forall n > 0; \qquad x_0 - y_0 \geq 0.$$

The objective function in (OPT2R) is continuous and the constraint set is compact. Hence, (OPT2R) (and thus (OPT2)) has a solution. Moreover, since its objective function is affine, its inequality constraint functions are concave and finite on B, and its equality constraint functions are affine, (OPT2R) is an ordinary convex program (in the sense of Rockafellar (1970)). By Corollary 4, there exists (x', y') such that $(\theta_0 x' + (1-\theta_0)y')u - C_b(x', y') > 0$ and $w_1 x'\theta_0 + w_0 y'(1-\theta_0) > 0$; since $\theta_0 < b$, we have $x' - y' > 0$. Then for each n, we have

$$U_b(x',y',\theta_n) - \sum_{i<n}(\theta_{i+1}-\theta_i)(x'-y')u = (\theta_0 x' + (1-\theta_0)y')u - C_b(x',y') > 0.$$

Hence, by continuity, we can choose an increasing sequence $\{\epsilon_n\}_{n=0}^N$ such that $\{(x_n,y_n)\}_{n=0}^N = (x' + \epsilon_n, y')$ satisfies all constraints in (OPT2R) strictly.

Then by Corollary 28.2.1 and Theorem 28.1 of Rockafellar (1970), there exist non-negative $\{\lambda_n^*\}_{n=0}^N$ and $\{\delta_n^*\}_{n=0}^N$ such that the maximized values of (OPT2) and

$$\max_{\{(x_n,y_n)\}_{n=0}^N \in B} \sum_{n=0}^{N} W_b(x_n,y_n,\theta_n)\sigma(\theta_n) + \lambda_n^* \left(U_b(x_n,y_n,\theta_n) - \sum_{i<n}(\theta_{i+1}-\theta_i)(x_i-y_i)u \right)$$
$$+ \delta_n^* \left((x_n-y_n) - (x_{n-1}-y_{n-1})\right) \mathbf{1}_{n>0} + \delta_0^*(x_0-y_0) \tag{OPT2R′}$$

are equal, and $\{(x_n,y_n)\}_{n=0}^N$ solves (OPT2) if and only if it solves (OPT2R′) and satisfies (CS2).

Finally, we show that the set of solutions of (OPT2R′) that satisfy (CS2) coincides with the set of solutions of (OPT2′) that satisfy (CS2). Consider a solution $d^* = \{(x_n^*, y_n^*)\}_{n=0}^N$ to (OPT2R′) that satisfies (CS2), and suppose toward a contradiction that there exists $d' = \{(x_n', y_n')\}_{n=0}^N \in [0,1]^{2(N+1)} \setminus B$ that achieves a weakly greater value of the objective in (OPT2′) and satisfies (CS2). The objective in the two problems is quasiconcave, so it follows that if $\alpha d^* + (1-\alpha)d'$ is in B for some $\alpha \in (0,1)$, then it solves (OPT2R′). Let $M' = \{n : (\theta_n x_n' + (1-\theta_n)y_n')u - C_b(x_n', y_n') < -u\}$. Since $d' \notin B$, M' is nonempty. Since $(\theta_n x + (1-\theta_n)y)u - C_b(x,y)$ is continuous in (x,y), it follows that for each $n \in M'$, there exists α_n such that $(\theta_n(\alpha x_n + (1-\alpha)x_n') + (1-\theta_n)(\alpha y_n + (1-\alpha)y_n'))u - C_b(\alpha x_n + (1 -$

$\alpha)x'_n, \alpha y_n + (1-\alpha)y'_n) = -u$. Let $n^* = \arg\max_{n \in M'} \alpha_n$. Then since each B_n is convex, $\alpha_{n^*} d^* + (1-\alpha_{n^*})d' \in B$, and so it solves (OPT2R′).

Moreover, it satisfies (CS2): The δ_n^* conditions are linear, so they are satisfied by $\alpha_{n^*} d^* + (1-\alpha_{n^*})d'$ if they are satisfied by d^* and d'. For the λ_n^* conditions, observe that we must have $\lambda_n^* = 0$ for any n such that $(x_n^*, y_n^*) \neq (x'_n, y'_n)$ and either $x_n^* \neq y_n^*$ or $x'_n \neq y'_n$; otherwise, by strict concavity of C_b, $\alpha_{n^*} d^* + (1-\alpha_{n^*})d'$ achieves a strictly greater value than d^* does in (OPT2R′), contradicting the optimality of d^*. Then when applied to $\alpha_{n^*} d^* + (1-\alpha_{n^*})d'$, either the λ_n^* comparative slackness condition is vacuous (if $\lambda_n^* = 0$), linear in α_{n^*} (if $x_n^* = y_n^*$ and $x'_n = y'_n$), or satisfied because $(x_n^*, y_n^*) = (x'_n, y'_n)$ and d^* and d' each satisfy (CS2).

Then by the equivalence between (OPT2R) and (OPT2R′), $\alpha_{n^*} d^* + (1-\alpha_{n^*})d'$ solves (OPT2R). But $(\theta_{n^*}(\alpha_{n^*} x_{n^*} + (1-\alpha_{n^*})x'_{n^*}) + (1-\theta_{n^*})(\alpha_{n^*} y_{n^*} + (1-\alpha_{n^*})y'_{n^*}))u - C_b(\alpha_{n^*} x_{n^*} + (1-\alpha_{n^*})x'_{n^*}, \alpha_{n^*} y_{n^*} + (1-\alpha_{n^*})y'_{n^*}) = -u < -u(\theta_{n^*} - \theta_0)$, so $\alpha_{n^*} d^* + (1-\alpha_{n^*})d'$ is not feasible in (OPT2R), a contradiction.

Uniqueness. Observe that (OPT2′) is additively separable in (x_n, y_n), so that $\{(x_n^*, y_n^*)\}_{n=0}^N$ solves it if and only if for each n,

$$(x_n^*, y_n^*) \in \arg \max_{x,y \in [0,1]^2} \begin{array}{l} \sigma(\theta_n) W_b(x,y,\theta_n) + \lambda_n^* U_b(x,y,\theta_n) \\ +(x-y)\left(\delta_n^* - \delta_{n+1}^* \mathbf{1}_{n<N} - (\theta_{n+1} - \theta_n)u \sum_{i>n} \lambda_i^*\right). \end{array} \tag{TBT2θ_n}$$

For each n with $\lambda_n^* > 0$, the objective in (TBT2θ_n) is strictly concave in (x,y), and so its solution is unique. For each n with $\lambda_n^* = 0$, the set of solutions is given by

$$\left\{ (x,y) \in [0,1]^2 \;\middle|\; \begin{array}{rl} x = 1, y = 0, & \left(\delta_n^* - \delta_{n+1}^* \mathbf{1}_{n<N} - (\theta_{n+1} - \theta_n)u \sum_{i>n} \lambda_i^*\right) > -\theta_n w_1; \\ y = 0, & \left(\delta_n^* - \delta_{n+1}^* \mathbf{1}_{n<N} - (\theta_{n+1} - \theta_n)u \sum_{i>n} \lambda_i^*\right) = -\theta_n w_1; \\ x = y = 0, & \left(\delta_n^* - \delta_{n+1}^* \mathbf{1}_{n<N} - (\theta_{n+1} - \theta_n)u \sum_{i>n} \lambda_i^*\right) \in ((1-\theta_n)w_0, -\theta_n w_1); \\ x = 0, & \left(\delta_n^* - \delta_{n+1}^* \mathbf{1}_{n<N} - (\theta_{n+1} - \theta_n)u \sum_{i>n} \lambda_i^*\right) = (1-\theta_n)w_0; \\ x = 0, y = 1, & \left(\delta_n^* - \delta_{n+1}^* \mathbf{1}_{n<N} - (\theta_{n+1} - \theta_n)u \sum_{i>n} \lambda_i^*\right) < (1-\theta_n)w_0. \end{array} \right\}$$

But if $x = 0$ and $y > 0$, or $x > 0$ and $y = 0$, then $C_b(x,y) = \infty$, and any $\{(x_i, y_i)\}_{i=0}^N$ with $(x_n, y_n) = (x,y)$ is not feasible in (OPT2). Thus, from the equivalence between solutions to (OPT2) and (OPT2′) that satisfy (CS2) it follows that for any n with $\lambda_n^* = 0$, any solution $\{(x_i^*, y_i^*)\}_{i=0}^N$ to (OPT2′) either has $(x_n^*, y_n^*) = (0,0)$ or does not satisfy (CS2). Thus, there is a unique solution to (OPT2′) which satisfies (CS2) — and thus a unique solution to (OPT2). □

Proof of Theorem 2 (Existence and Uniqueness) We prove a series of intermediate claims that together establish Theorem 2.

Claim T2.1: Given $\{\lambda_n^*\}_{n=0}^N$ and $\{\delta_n^*\}_{n=0}^N$, (TBTθ_n) has a binary solution for each n, and this solution is unique (up to Blackwell equivalence) whenever $\lambda_n^* > 0$. By Lemma OA.2, for each n with $\lambda_n^* > 0$, (TBTθ_n) has a unique (up to Blackwell equivalence) solution, and this solution is binary. For n with $\lambda_n^* = 0$, (TBTθ_n) has a binary solution by the Fenchel-Bunt theorem. ■

Claim T2.2: There exist nonnegative Lagrange multipliers $\{\lambda_n^*\}_{n=0}^N$ and $\{\delta_n^*\}_{n=0}^N$ such that the unique (by Lemma 8) solution to (OPT2) is equivalent to

- **i. a deterministic mechanism $\chi^*(\theta) = (\pi_\theta^*, a_\theta^*)$ that solves (OPT) such that each π_θ^* is binary, each a_θ^* is ex-post optimal for π_θ^*, $a_\theta^*(\pi_\theta^*, \bar{s}) = 1$ and $a_\theta^*(\pi_\theta^*, \underline{s}) = 0$ for each $\theta \in \Theta$, and (M-FC), (EC-FC), and (NL) hold;**
- **ii. a binary menu $D^* = \{\pi_{\theta_n}^*\}_{n=0}^N \cup \{\pi_0\}$ that solves (COPT), and**
- **iii. the unique solution $\{\pi_{\theta_n}^*\}_{n=0}^N$ to (COPT′) with $\pi_{\theta_n}^* \sim_B \pi_0$ for each n with $\lambda_n^* = 0$.**

Moreover, $\{\pi_{\theta_n}^*\}_{n=0}^N$ satisfies (CS).

Let $\{(x_n^*, y_n^*)\}_{n=0}^N$ be the unique (by Lemma 8) solution to (OPT2); let $\pi_{\theta_n}^*$ be the binary experiment with $\pi_{\theta_n}^*(\bar{s}|1) = x_n^*$ and $\pi_{\theta_n}^*(\bar{s}|0) = y_n^*$.

By Lemma 6(ii) it is without loss in (OPT) to restrict attention to deterministic mechanisms $\chi : \Theta \to \Pi \times \mathcal{A}$ denoted by $\chi(\theta) = (\pi_\theta, a_\theta)$ such that each π_θ is binary, $a_\theta(\pi_\theta, \bar{s}) = 1$, $a_\theta(\pi_\theta, \underline{s}) = 0$, and (EC-FC), (M-FC), and (NL) hold. When we do, (OPT) reduces to the finite-dimensional program (OPT2). It follows that the mechanism χ^* defined by $\chi^*(\theta_n) = (\pi_{\theta_n}^*, a_{\theta_n}^*)$, where for each n, $a_{\theta_n}^*(\pi_{\theta_n}^*, \bar{s}) = 1$ and $a_{\theta_n}^*(\pi_{\theta_n}^*, \underline{s}) = 0$, is a solution to (OPT) (Claim T2.2.i). By Claim T1.1, each $a_{\theta_n}^*$ must be ex-post optimal for $\pi_{\theta_n}^*$: we must have $\theta_n x_n^* w_1 + (1-\theta_n) y_n^* w_0 \geq 0$ for all n. Therefore, $D^* = \{\pi_{\theta_n}^*\}_{n=0}^N \cup \{\pi_0\}$ solves (COPT) (Claim T2.2.ii).

By Lemma 8, there exist nonnegative Lagrange multipliers $\{\lambda_n^*\}_{n=0}^N$ and $\{\delta_n^*\}_{n=0}^N$ such that $\{(x_n^*, y_n^*)\}_{n=0}^N$ solves (OPT2R′) and satisfies (CS2), and such that $(x_n^*, y_n^*) = (0,0)$ for all n with $\lambda_n^* = 0$. Since $\theta_n x_n^* w_1 + (1-\theta_n) y_n^* w_0 \geq 0$ for all n, it follows that given $\{\lambda_n^*\}_{n=0}^N$ and $\{\delta_n^*\}_{n=0}^N$, $\{\pi_{\theta_n}^*\}_{n=0}^N$ satisfies (CS) and maximizes (COPT′) among all tuples of binary experiments. By Claim 2.1, (COPT′) has a binary solution, so $\{\pi_{\theta_n}^*\}_{n=0}^N$ solves (COPT′). Moreover, by Claim 2.1, this must be the unique solution to (COPT′) with $\pi_{\theta_n}^* \sim_B \pi_0$ for each n with $\lambda_n^* = 0$, since the solution to (TBTθ_n) for each n with $\lambda_n^* > 0$ is unique. (Claim T2.2.iii). ■

Claim T2.3: D^* is the unique solution to (COPT) up to Blackwell equivalence. Suppose toward a contradiction that there is another menu $D' = \{\pi'_{\theta_n}\}_{n=0}^N \cup \{\pi'_0\} \not\sim_B D^*$ which solves (COPT). Then $\sum_{n=0}^N E_{\langle \pi'_{\theta_n} | \theta_n \rangle}[W(\beta)]\sigma(\theta_n) = \sum_{n=0}^N E_{\langle \pi^*_{\theta_n} | \theta_n \rangle}[W(\beta)]\sigma(\theta_n)$.

Since each δ_n^* is nonnegative, and D' satisfies (ECθ) and (M(θ, θ')), it follows that $\{\pi'_{\theta_n}\}_{n=0}^N$ achieves a weakly higher value of the objective function in (COPT′) than $\{\pi^*_{\theta_n}\}_{n=0}^N$, and so also solves (COPT′). By Claim 2.1, then, we must have $\pi^*_{\theta_n} \sim_B \pi'_{\theta_n}$ for all n with $\lambda_n^* > 0$. It follows that there exists some k with $\lambda_k^* = 0$ such that $\pi'_{\theta_k} \nsim_B \pi^*_{\theta_k} \sim_B \pi_0$.

Since D^* satisfies (M), if $\lambda_n^* = 0$ for some n, then since $E_{\langle \pi^*_{\theta_n} | \theta_n \rangle}\left[\left(\frac{\beta - \theta_n}{\theta_n(1-\theta_n)}\right) U(\beta)\right] = 0$, it follows that for all $m < n$, $E_{\langle \pi^*_{\theta_m} | \theta_m \rangle}\left[\left(\frac{\beta - \theta_m}{\theta_m(1-\theta_m)}\right) U(\beta)\right] = 0$. Let $n^* = \max\{n : \lambda_n^* = 0\} \geq k$. If $n^* < N$, then since D^* satisfies (EC), we have $E_{\langle \pi^*_{\theta_{n^*+1}} | \theta_{n^*+1} \rangle}\left[U(\beta) - G(\beta | \theta_{n^*+1})\right] = 0$. Then since D' satisfies (EC), and (by the previous step) $\pi^*_{\theta_{n^*+1}} \sim_B \pi'_{\theta_{n^*+1}}$, we must have $E_{\langle \pi'_{\theta_m} | \theta_m \rangle}\left[\left(\frac{\beta - \theta_m}{\theta_m(1-\theta_m)}\right) U(\beta)\right] = 0$, and thus $E_{\langle \pi'_{\theta_m} | \theta_m \rangle}\left[U(\beta)\right] = 0$, for all $m \leq n^*$. But since $\pi'_{\theta_k} \nsim_B \pi^*_{\theta_k} \sim_B \pi_0$, we must have $C(\pi'_{\theta_k}) > 0$, and therefore $E_{\langle \pi'_{\theta_k} | \theta_k \rangle}\left[U(\beta)\right] - C(\pi'_{\theta_k}) < 0$, a contradiction since D' satisfies (EC) and thus (ECθ) for $\theta = \theta_k$.

Then it must be that $n^* = N$, and so $\pi^*_\theta \sim_B \pi_0$ for all $\theta \in \Theta$, and the maximized value of (COPT) is zero. But by assumption, there is π with $E_{\langle \pi | \theta_0 \rangle}[U(\beta)] > C(\pi)$ and $E_{\langle \pi | \theta_0 \rangle}[W(\beta)] > 0$. Since Bayes' rule is monotone in prior belief, it follows that $E_{\langle \pi | \theta \rangle}[U(\beta)] > C(\pi)$ and $E_{\langle \pi | \theta \rangle}[W(\beta)] > 0$ for all $\theta \in \Theta$. Then the mechanism χ' with $\chi'(\theta) = (\pi, a^*)$ with $a^*(\pi, \bar{s}) = 1$ and $a^*(\pi, \underline{s}) = 0$ is feasible in (OPT) and achieves a strictly higher value of the objective function than χ^*, contradicting Claim T2.2.i. The claim (and thus the proposition) follows. ■□

Proof of Proposition 3 (Increasing True Positive Rate) For $(x, y) \in [0,1]^2$, define

$$I(x, y, \theta) = \{(\hat{x}, \hat{y}) \in [0,1]^2 : U_b(\hat{x}, \hat{y}, \theta) = U_b(x, y, \theta)\}.$$

We first prove two intermediate claims.

Claim P3.1: For any $\theta', \theta \in \Theta$, if $(\hat{x}, \hat{y}) \in I(x, y, \theta')$ and $\hat{x} - \hat{y} = x - y$, then $(\hat{x}, \hat{y}) \in I(x, y, \theta)$. We have

$$\begin{aligned} (\theta'(\hat{x} - \hat{y}) + \hat{y})u - C_b(\hat{x}, \hat{y}) &= (\theta'(x - y) + y)u - C_b(x, y) \\ \iff \quad \hat{y}u - C_b(\hat{x}, \hat{y}) &= yu - C_b(x, y) \\ \iff \quad (\theta(\hat{x} - \hat{y}) + \hat{y})u - C_b(\hat{x}, \hat{y}) &= (\theta(x - y) + y)u - C_b(x, y). \end{aligned} \quad ■$$

Claim P3.2: Suppose $\theta, \theta' \in \Theta$, $(x, y) \in [0,1]^2$, $(x', y') \in I(x, y, \theta')$, $x' < x$, and $x' - y' > x - y \geq 0$. Then there exists $(\hat{x}, \hat{y}) \in I(x, y, \theta')$ such that $\hat{x} < x'$ and $\hat{x} - \hat{y} = x - y$. If $x - y = 0$, then $x' - y' > x - y = 0$; since $y' \geq 0$, we must have $x' > 0$. Then choose $(\hat{x}, \hat{y}) = (0, 0)$: we have $U_b(x, y, \theta') = 0 = U_b(0, 0, \theta')$.

Now suppose that $x > y$. Since $x' - y' > x - y$, we have $x' - (x - y) > y'$. Holding x'

fixed, $U_b(x', \cdot, \theta')$ is strictly increasing on $[0, x']$: for $0 < z < x' < 1$,

$$\frac{\partial}{\partial z} U_b(x', z, \theta') = (1 - \theta')u + \log\left(\frac{x'(1-z)}{z(1-x')}\right) + \frac{x' - z}{z(1-z)} > 0.$$

Thus,

$$U_b(x', x' - (x - y), \theta') > U_b(x', y', \theta') = U_b(x, y, \theta').$$

Define $\hat{U}(s) = U_b(s, s - (x - y), \theta')$ for $s \in (x - y, 1)$. Then $\hat{U}$ is continuous, $\hat{U}(s) \to -\infty$ as $s \downarrow (x - y)$, and $\hat{U}(x') > U_b(x, y, \theta')$. Hence, by the intermediate value theorem, there exists $\hat{x} \in (x - y, x')$ such that $\hat{U}(\hat{x}) = U_b(x, y, \theta')$. Let $\hat{y} = \hat{x} - (x - y)$. Then $(\hat{x}, \hat{y}) \in I(x, y, \theta')$ and $\hat{x} - \hat{y} = x - y$. The claim follows. ■

Let $d^* = \{(x_n^*, y_n^*)\}_{n=0}^N$ be the (unique, by Lemma 8) solution to (OPT2); by Claims T2.2(ii) and T2.3, $\pi_{\theta_n}^*(\bar{s}|1) = x_n^*$ and $\pi_{\theta_n}^*(\bar{s}|0) = y_n^*$ for each $n \in \{0, \dots, N\}$.

It is enough to prove the proposition for adjacent types. Fix $n \in \{0, \dots, N-1\}$. By feasibility in (OPT2),

$$\begin{aligned} U_b(x_{n+1}^*, y_{n+1}^*, \theta_{n+1}) &= u \sum_{i<n+1} (\theta_{i+1} - \theta_i)(x_i^* - y_i^*) \\ &= U_b(x_n^*, y_n^*, \theta_n) + u(\theta_{n+1} - \theta_n)(x_n^* - y_n^*) \\ &= U_b(x_n^*, y_n^*, \theta_{n+1}). \end{aligned} \tag{12}$$

so $(x_{n+1}^*, y_{n+1}^*) \in I(x_n^*, y_n^*, \theta_{n+1})$.

Suppose toward a contradiction that $x_n^* > x_{n+1}^*$. By (M), $x_{n+1}^* - y_{n+1}^* = \eta(\pi_{\theta_{n+1}}^*) \geq \eta(\pi_{\theta_n}^*) = x_n^* - y_n^* \geq 0$. If equality held, Lemma 4 would imply $\pi_{\theta_{n+1}}^* \sim_B \pi_{\theta_n}^*$, contradicting $x_n^* > x_{n+1}^*$. Thus $x_{n+1}^* - y_{n+1}^* > x_n^* - y_n^*$.

By Claim P3.2, there exists $(\hat{x}_n, \hat{y}_n) \in I(x_n^*, y_n^*, \theta_{n+1})$ such that $\hat{x}_n < x_{n+1}^*$ and $\hat{x}_n - \hat{y}_n = x_n^* - y_n^*$. By Claim P3.1, $(\hat{x}_n, \hat{y}_n) \in I(x_n^*, y_n^*, \theta_n)$ as well.

Then $\hat{d} = \{\{(x_i^*, y_i^*)\}_{i \neq n}, (\hat{x}_n, \hat{y}_n)\}$ is feasible in (OPT2), since d^* is feasible, $\hat{x}_n - \hat{y}_n = x_n^* - y_n^*$, and $U_b(\hat{x}_n, \hat{y}_n, \theta_n) = U_b(x_n^*, y_n^*, \theta_n)$. But since $\hat{y}_n = \hat{x}_n - (x_n^* - y_n^*) < x_n^* - (x_n^* - y_n^*) = y_n^*$ and $\theta_n < b$,

$$W_b(\hat{x}_n, \hat{y}_n, \theta_n) - W_b(x_n^*, y_n^*, \theta_n) = (\hat{y}_n - y_n^*)(\theta_n w_1 + (1 - \theta_n) w_0) > 0.$$

and so $\hat{d}$ yields a strictly higher value of the objective in (OPT2), a contradiction.

Then $x_{n+1}^* \geq x_n^*$. Suppose that $\pi_{\theta_{n+1}}^* \neq \pi_{\theta_n}^*$. Then Lemma 4 and (M) imply $x_{n+1}^* - y_{n+1}^* = \eta(\pi_{\theta_{n+1}}^*) > \eta(\pi_{\theta_n}^*) = x_n^* - y_n^* \geq 0$. If $x_{n+1}^* = x_n^*$, then $y_{n+1}^* < y_n^* \leq x_n^*$. Since d^* is feasible in (OPT2), we must have $C_b(x_n^*, y_n^*) < \infty$. If $x_n^* = 1$, finite cost would force $y_n^* = 1$; then by (M) and feasibility in (OPT2), $u = U_b(x_n^*, y_n^*, \theta_n) = 0$, a contradiction.

Thus $x_n^* < 1$. Then by strict monotonicity of $U_b(x_n^*, \cdot, \theta_{n+1})$ on $[0, x_n^*]$, we have

$$U_b(x_n^*, y_n^*, \theta_{n+1}) > U_b(x_n^*, y_{n+1}^*, \theta_{n+1}) = U_b(x_{n+1}^*, y_{n+1}^*, \theta_{n+1}),$$

contradicting (12).

The statement follows when $\theta' = \theta_{n+1}$ and $\theta = \theta_n$ for some n; it follows for arbitrary $\theta' > \theta$ by chaining the argument for adjacent pairs. □

Proof of Corollary 1 It suffices to consider a pair $\theta' \geq \theta$ such that $\theta' = \theta_{n+1}$ and $\theta = \theta_n$. Suppose that $\pi_{\theta'}^* \neq \pi_\theta^*$. Then, by Proposition 3, $\pi_{\theta'}^*(\bar{s}|1) > \pi_\theta^*(\bar{s}|1)$. If $\pi_{\theta'}^*(\bar{s}|0) \leq \pi_\theta^*(\bar{s}|0)$, then $\pi_{\theta'}^* \succ_B \pi_\theta^*$ and so $C(\pi_{\theta'}^*) > C(\pi_\theta^*)$. If $\pi_{\theta'}^*(\bar{s}|0) \geq \pi_\theta^*(\bar{s}|0)$, by (EC$\theta'$),

$$\begin{aligned}
&(\theta'(x'-y')+y')u - C_b(x',y') = (\theta(x-y)+y)u - C_b(x,y) \\
&\quad \iff C_b(x',y') - C_b(x,y) = (\theta'((x'-y')-(x-y)) + (y'-y))u > 0,
\end{aligned}$$

where $(x, y) = (\pi_\theta^*(\bar{s}|1), \pi_\theta^*(\bar{s}|0))$ and $(x', y') = (\pi_{\theta'}^*(\bar{s}|1), \pi_{\theta'}^*(\bar{s}|0))$. This follows since, by Lemma 4, $x' - y' > x - y$.

Proof of Proposition 4 (Efficiency and the Social Planner's Problem) (If) Suppose that $\langle \pi|\theta \rangle$ solves (SPPθ). We consider three cases.

If $\lambda_a > 0$ and $\lambda_p > 0$, then there can be no $\pi' \neq \pi$ with $E_{\langle \pi'|\theta \rangle}[W(\beta)] \geq E_{\langle \pi|\theta \rangle}[W(\beta)]$ and $E_{\langle \pi'|\theta \rangle}[U(\beta) - G(\beta|\theta)] \geq E_{\langle \pi|\theta \rangle}[U(\beta) - G(\beta|\theta)]$ with one inequality strict; if there were, $\langle \pi'|\theta \rangle$ would achieve a higher value in (SPPθ). Thus, π is Pareto efficient for type θ.

If $\lambda_a = 0$, then (SPPθ) is equivalent to

$$\langle \pi|\theta \rangle \in \arg \max_{\tau \in \Delta(\Delta(\Omega))} \{E_\tau [W(\beta)] \text{ s.t. } E_\tau \beta = \theta\}. \tag{13}$$

Since W is convex, its concavification can be written $\overline{W}(\beta) = \beta w_1$. Hence, by Lemma 3 in Yoder (2022), the solution to (13) is the unique Bayes-plausible distribution with support $\{0, 1\}$, which is induced by the fully informative experiment π_∞. It follows that there can be no $\pi' \not\sim_B \pi_\infty$ with $E_{\langle \pi'|\theta \rangle}[W(\beta)] \geq E_{\langle \pi_\infty|\theta \rangle}[W(\beta)]$, and so π is Pareto efficient for type θ.

If $\lambda_p = 0$, then (SPPθ) is equivalent to

$$\langle \pi|\theta \rangle \in \arg \max_{\tau \in \Delta(\Delta(\Omega))} \{E_\tau [U(\beta) - G(\beta|\theta)] \text{ s.t. } E_\tau \beta = \theta\}. \tag{14}$$

By Lemma OA.2, the solution to (14) is unique. It follows that there can be no π' with $E_{\langle \pi'|\theta \rangle}[U(\beta) - G(\beta|\theta)] \geq E_{\langle \pi|\theta \rangle}[U(\beta) - G(\beta|\theta)]$, and so π is Pareto efficient for type θ.

(Only if) Suppose that π is Pareto efficient for type θ. If $E_{\langle \pi|\theta \rangle}[U(\beta) - G(\beta|\theta)] = -\infty$ then (13) must hold, and so $\langle \pi|\theta \rangle$ solves (SPPθ) for $\lambda_a = 0$ and $\lambda_p > 0$. Suppose instead that $E_{\langle\pi|\theta\rangle}[U(\beta) - G(\beta|\theta)] = y^* \in \mathbb{R}$, and let $E_{\langle\pi|\theta\rangle}[W(\beta)] = x^*$.

Let

$$Z_\theta = \{(x,y) \in \mathbb{R}^2 \mid \exists \tau \in \Delta(\Delta(\Omega)) : x \leq E_\tau [W(\beta)],\ E_\tau [U(\beta) - G(\beta|\theta)] \geq y,\ E_\tau \beta = \theta\}.$$

Z_θ is convex: Let $(x,y), (x',y') \in Z_\theta$ and let $\lambda \in (0,1)$. Then for some τ, τ' with $E_\tau \beta = E_{\tau'} \beta = \theta$,

$$x \leq E_\tau [W(\beta)],\quad y \leq E_\tau [U(\beta) - G(\beta|\theta)],\quad x' \leq E_{\tau'} [W(\beta)],\quad y' \leq E_{\tau'} [U(\beta) - G(\beta|\theta)].$$

Let $\tau_\lambda = \lambda\tau + (1-\lambda)\tau'$. Then $E_{\tau_\lambda}\beta = \lambda E_\tau \beta + (1-\lambda) E_{\tau'}\beta = \theta$. Moreover,

$$\begin{aligned}
&\lambda x + (1-\lambda)x' \leq \lambda E_\tau [W(\beta)] + (1-\lambda) E_{\tau'} [W(\beta)] = E_{\tau_\lambda} [W(\beta)], \\
&\lambda y + (1-\lambda)y' \leq \lambda E_\tau [U(\beta) - G(\beta|\theta)] + (1-\lambda) E_{\tau'} [U(\beta) - G(\beta|\theta)] = E_{\tau_\lambda} [U(\beta) - G(\beta|\theta)].
\end{aligned}$$

Hence $(\lambda x + (1-\lambda)x', \lambda y + (1-\lambda)y') \in Z_\theta$.

Now $(x^*, y^*) \in \operatorname{Bd} Z_\theta$: If not, then there exists $\epsilon > 0$ such that the ϵ-ball about (x^*, y^*) is contained in Z_θ. But then $(x^* + \epsilon/2, y^* + \epsilon/2) \in Z_\theta$, a contradiction since π is Pareto efficient for type θ.

Then by the supporting hyperplane theorem, there exists $(\lambda_p^*, \lambda_a^*) \in \mathbb{R}^2 \setminus \{0\}$ such that $\lambda_p^* x^* + \lambda_a^* y^* \geq \lambda_p^* x + \lambda_a^* y$ for all $(x,y) \in Z_\theta$. Since $(x,y) \in Z_\theta$ implies $(x',y') \in Z_\theta$ for all $x' \leq x$, $y' \leq y$, it follows that $\lambda_p^*, \lambda_a^* \geq 0$.

If $\lambda_a^* > 0$, the claim follows immediately. If $\lambda_a^* = 0$, then

$$\begin{aligned}
x^* &= \sup\{x \mid x = E_\tau [W(\beta)] \text{ for some } \tau \text{ with } E_\tau [U(\beta) - G(\beta|\theta)] > -\infty \text{ and } E_\tau \beta = \theta\} \\
&= \sup\{x \mid x = E_\tau [W(\beta)] \text{ for some } \tau \text{ with } \operatorname{supp} \tau \subseteq (0,1) \text{ and } E_\tau \beta = \theta\} \\
&= \theta W(1) + (1-\theta) W(0) = \max_{\tau \in \Delta(\Delta(\Omega))} \{E_\tau [W(\beta)] \text{ s.t. } E_\tau \beta = \theta\},
\end{aligned}$$

where the third equality follows since W is convex and continuous, and since we have $\tau_n \to^{w^*} \tau_\infty$, where τ_∞ is the fully informative Bayes-plausible distribution over posteriors and

$$\begin{aligned}
\tau_n(\{\theta/n\}) &= 1-\theta, &\qquad \tau_n(\{1-(1-\theta)/n\}) &= \theta; \\
\tau_\infty(\{0\}) &= 1-\theta, &\qquad \tau_\infty(\{1\}) &= \theta.
\end{aligned}$$

The claim then follows by letting $\lambda_a = 0$ and $\lambda_p > 0$ in (SPPθ). □

Proof of Corollary 2 Follows immediately from Proposition 4 and Lemma OA.2. ■

Proof of Corollary 3 By Lemma OA.2, since $\pi^*_{\theta_n} \not\sim_B \pi_0$ and $\pi^*_{\theta_n}$ solves (TBTθ_n), we have

$$-\lambda_n^* G'(\underline{\beta}|\theta_n) = \frac{\sigma(\theta_n)W(\overline{\beta}) + \lambda_n^* u - R(\overline{\beta},\theta_n) - \lambda_n^*(G(\overline{\beta}|\theta_n) - G(\underline{\beta}|\theta_n))}{\overline{\beta} - \underline{\beta}}; \tag{15}$$

$$-\lambda_n^* G'(\underline{\beta}|\theta_n) = \sigma(\theta_n)(w_1 - w_0) + R_\beta(\overline{\beta},\theta_n) - \lambda_n^* G'(\overline{\beta}|\theta_n), \text{ if } \overline{\beta} > b. \tag{16}$$

When $\overline{\beta} = b$, (5) follows immediately from (15); when $\overline{\beta} > b$, application of (16) yields

$$\begin{aligned}
-G(\underline{\beta}|\theta_n) - G'(\underline{\beta}|\theta_n)(b - \underline{\beta}) &= -G(\overline{\beta}|\theta_n) - G'(\overline{\beta}|\theta_n)(b - \overline{\beta}) + u \\
&\quad + \frac{1}{\lambda_n^*}\begin{pmatrix} \sigma(\theta_n)W(\overline{\beta}) - R(\overline{\beta},\theta_n) \\ -\left(\sigma(\theta_n)(w_1 - w_0) - R_\beta(\overline{\beta},\theta_n)\right)(\overline{\beta} - b) \end{pmatrix} \\
&= -G(\overline{\beta}|\theta_n) - G'(\overline{\beta}|\theta_n)(b - \overline{\beta}) - R(b,\theta_n)/\lambda_n^*,
\end{aligned}$$

as desired. ■

Proof of Theorem 3 (Distortion Everywhere Else) Let $d^* = \{(x_i^*, y_i^*)\}_{i=0}^N$ be the (unique, by Lemma 8) solution to (OPT2); let $\{\lambda_i^*\}_{i=0}^N$ and $\{\delta_i^*\}_{i=0}^N$ be the nonnegative Lagrange multipliers pinned down in Lemma 8. By Claims T2.2(ii) and T2.3, $\pi^*_{\theta_i}(\overline{s}|1) = x_i^*$ and $\pi^*_{\theta_i}(\overline{s}|0) = y_i^*$ for each $i \in \{0, \ldots, N\}$. By Lemma 8, $\lambda_i^* = 0$ implies $(x_i^*, y_i^*) = (0,0)$, so $\lambda_i^* > 0$ for each $\theta_i \in \tilde{\Theta}$.

By Lemma 8, (x_i^*, y_i^*) solves the type-i problem (TBT2θ_i), whose objective is

$$\sigma(\theta_i)W_b(x,y,\theta_i) + \lambda_i^* U_b(x,y,\theta_i) - \rho_i^*(x - y),$$

where $\rho_i^* := \left(\delta_{i+1}^* + (\theta_{i+1} - \theta_i)u \sum_{j>i} \lambda_j^*\right) \mathbf{1}_{i<N} - \delta_i^*$. Now fix $\theta = \theta_n \in \tilde{\Theta}$ with $\theta < \theta_N$, and let $z = \max\{i \leq n \mid \delta_i^* = 0 \text{ or } i = 0\}$. That is, θ_z is the largest type below θ_n such that the monotonicity constraint does not bind. Then $\delta_i^* > 0$ for all $i \in \{z+1, \ldots, n\}$. By Lemma 8, d^* satisfies (CS2); it follows that $x_z^* - y_z^* = x_n^* - y_n^* > 0$. Then by Lemma 4, $(x_z^*, y_z^*) = (x_n^*, y_n^*)$, and so $\theta_z \in \tilde{\Theta}$.

If $z > 0$, then $\delta_z^* = 0$ by construction; if $z = 0$, then $\delta_0^* = 0$ by (CS2) since $x_z^* - y_z^* > 0$. Hence $\rho_z^* \geq 0$.

(ii): Let π be an experiment on the Pareto frontier for type θ_n that Pareto improves upon $\pi^*_{\theta_n}$ for type θ_n. By Corollary 2, we may take π to be binary; write $x = \pi(\overline{s}|1)$ and $y = \pi(\overline{s}|0)$. Suppose toward a contradiction that $x \leq x_n^*$.

Since π Pareto improves upon $\pi^*_{\theta_n}$, we have $W_b(x,y,\theta_n) \geq W_b(x_n^*, y_n^*, \theta_n)$. By Claim T2.2(i), $W_b(x_n^*, y_n^*, \theta_n) \geq 0$ for all n. Then since $\theta_n < b$, we must have $x > y$: if $x \leq y$, then $W_b(x,y,\theta_n) \leq y(\theta_n w_1 + (1-\theta_n)w_0) < 0$, except at $(x,y) = (0,0)$. But if $x = y = 0$, then $\pi \sim_B \pi_0$, and π cannot be on the Pareto frontier for type θ_n by Corollary 4.

Moreover, we must have $y < y_n^*$: Otherwise, we either have $y > y_n^*$, or (since $\pi \neq \pi^*_{\theta_n}$) $y = y_n^*$ and $x < x_n^*$. In either case, since $x \leq x_n^*$, we would have $W_b(x, y, \theta_n) < W_b(x_n^*, y_n^*, \theta_n)$, a contradiction.

Let

$$\bar{y} = \frac{(x - x_n^*)\theta_n w_1}{-(1-\theta_n)w_0} + y_n^* = \frac{x\theta_n w_1 - W_b(x_n^*, y_n^*, \theta_n)}{-(1-\theta_n)w_0} \leq x \frac{\theta_n w_1}{-(1-\theta_n)w_0} < x,$$

where the last inequality holds since $\theta_n < b$. Then $W_b(x, \bar{y}, \theta_n) = W_b(x_n^*, y_n^*, \theta_n)$. Moreover, since $w_0 < 0$ and $x \leq x_n^*$, we have $\bar{y} \leq y_n^*$. Furthermore, we have

$$\bar{y} = \frac{x\theta_n w_1 - W_b(x_n^*, y_n^*, \theta_n)}{-(1-\theta_n)w_0} \geq \frac{x\theta_n w_1 - W_b(x, y, \theta_n)}{-(1-\theta_n)w_0} = y,$$

and this inequality is strict if $W_b(x_n^*, y_n^*, \theta_n) < W_b(x, y, \theta_n)$.

Now observe that we have

$$\frac{\partial C_b(x,t)}{\partial t} = -\log\left(\frac{x(1-t)}{t(1-x)}\right) - \frac{x-t}{t(1-t)} < 0$$

for all $t \in (0, x)$. It follows that $C_b(x, y) \geq C_b(x, \bar{y})$, and this inequality is strict if $W_b(x_n^*, y_n^*, \theta_n) < W_b(x, y, \theta_n)$. Then since π Pareto improves on $\pi^*_{\theta_n}$, we have

$$\begin{aligned} U_b(x, \bar{y}, \theta_n) &= u(\theta_n x + (1-\theta_n)\bar{y}) - C_b(x, \bar{y}) \\ &\geq u(\theta_n x + (1-\theta_n)y) - C_b(x, y) = U_b(x, y, \theta_n) \geq U_b(x_n^*, y_n^*, \theta_n), \end{aligned}$$

where the first inequality is strict if $W_b(x_n^*, y_n^*, \theta_n) < W_b(x, y, \theta_n)$, and the second is strict if $U_b(x, y, \theta_n) > U_b(x_n^*, y_n^*, \theta_n)$. Since π Pareto improves on $\pi^*_{\theta_n}$, it follows that $U_b(x, \bar{y}, \theta_n) > U_b(x_n^*, y_n^*, \theta_n)$. Since $(x_z^*, y_z^*) = (x_n^*, y_n^*)$, we then have

$$\begin{aligned} U_b(x, \bar{y}, \theta_z) - U_b(x_z^*, y_z^*, \theta_z) &= U_b(x, \bar{y}, \theta_n) - U_b(x_n^*, y_n^*, \theta_n) \\ &\quad - u(\theta_n - \theta_z)\big((x - \bar{y}) - (x_n^* - y_n^*)\big) > 0; \\ W_b(x, \bar{y}, \theta_z) - W_b(x_z^*, y_z^*, \theta_z) &= W_b(x, \bar{y}, \theta_n) - W_b(x_n^*, y_n^*, \theta_n) \\ &\quad + (\theta_z - \theta_n)\big(w_1(x - x_n^*) - w_0(\bar{y} - y_n^*)\big) \geq 0, \end{aligned}$$

since $W_b(x, \bar{y}, \theta_n) = W_b(x_n^*, y_n^*, \theta_n)$, $(\theta_z - \theta_n) \leq 0$, $w_1 > 0$, $w_0 < 0$, $x \leq x_n^*$, and $\bar{y} \leq y_n^*$.

Now since $W_b(x, \bar{y}, \theta_n) = W_b(x_n^*, y_n^*, \theta_n)$ and $\theta_n < b$, it follows that

$$(x - \bar{y}) - (x_n^* - y_n^*) = \left(1 - \frac{\theta_n w_1}{-(1-\theta_n)w_0}\right)(x - x_n^*) \leq 0,$$

and strictly if $x \neq x_n^*$. Therefore replacing (x_z^*, y_z^*) by $(x, \bar{y})$ in the type-θ_z problem

(TBT2θ_z) increases its objective by

$$\sigma(\theta_z)\big(W_b(x,\bar{y},\theta_z) - W_b(x_z^*,y_z^*,\theta_z)\big) + \lambda_z^*\big(U_b(x,\bar{y},\theta_z) - U_b(x_z^*,y_z^*,\theta_z)\big) \\ - \rho_z^*\big((x-\bar{y}) - (x_z^* - y_z^*)\big) > 0,$$

because $\lambda_z^* > 0$, $\rho_z^* \geq 0$, and $(x-\bar{y}) - (x_z^* - y_z^*) \leq 0$. Then (x_z^*, y_z^*) does not solve (TBT2θ_z), a contradiction.

(iii)**:** Follows immediately from Corollary 4.

(i): We consider two cases.

Case 1: $\delta_n^* = 0$**, or** $n = 0$**.** Suppose toward a contradiction that $\pi_{\theta_n}^*$ is Pareto efficient for type θ_n. By Proposition 4, there exist $\lambda_p, \lambda_a \geq 0$, not identically zero, such that $\langle \pi_{\theta_n}^* | \theta_n \rangle$ solves (SPPθ).

If $\lambda_a = 0$, then $\langle \pi_{\theta_n}^* | \theta_n \rangle$ solves (13). Since W is convex, its concavification can be written $\overline{W}(\beta) = \beta w_1$. Hence, by Lemma 3 in Yoder (2022), the unique solution to (13) is the unique Bayes-plausible distribution with support $\{0,1\}$. Then $\pi_{\theta_n}^*$ is Blackwell-equivalent to the fully informative binary experiment π_∞ with $\pi_\infty(\bar{s}|1) = \pi_\infty(\underline{s}|0) = 1$. But $E_{\langle \pi_\infty | \theta_n \rangle}[S(\beta,\theta_n) - R(\beta,\theta_n)] = -\infty < S(\theta_n,\theta_n) - R(\theta_n,\theta_n)$, so $\pi_{\theta_n}^*$ does not solve (TBTθ_n). Hence, D^* does not solve (COPT), contradicting Theorem 2.

If $\lambda_a > 0$, then by Lemma OA.2, we have supp $\pi_{\theta_n}^* = \{\underline{\beta}^*, \overline{\beta}^*\}$ for $\underline{\beta}^*, \overline{\beta}^*$ that solve (3). By Corollary 3, $\underline{\beta}^*, \overline{\beta}^*$ also solve (5). But by Lemma OA.5, $R(b,\theta_n)/\lambda_n^* > 0$; then $\underline{\beta}^*, \overline{\beta}^*$ cannot solve both (5) and (3), a contradiction.

Case 2: $n > 0$ **and** $\delta_n^* \neq 0$**.** By Lemma 4, $\pi_{\theta_n}^* = \pi_{\theta_z}^*$. By Case 1, $\pi_{\theta_z}^*$ is inefficient for type θ_z. Then there exists π_{θ_z} on the Pareto frontier for type θ_z which Pareto dominates $\pi_{\theta_z}^*$ for type θ_z; by Corollary 2, π_{θ_z} is (without loss) binary; and by (ii), $\pi_{\theta_z}(\underline{s}|1) < \pi_{\theta_z}^*(\underline{s}|1)$. Then by Lemma OA.6, π_{θ_z} Pareto dominates $\pi_{\theta_z}^* = \pi_{\theta_n}^*$ for type θ_n as well. □

References

ADUSUMILLI, K. AND A. VEMULAPATI (2026): "Designing Persuasive Experiments," *arXiv preprint arXiv:2605.16703*.

ALONSO, R. AND O. CÂMARA (2016): "Bayesian Persuasion with Heterogeneous Priors," *Journal of Economic Theory*, 165, 672–706.

ALONSO, R. AND N. MATOUSCHEK (2008): "Optimal Delegation," *The Review of Economic Studies*, 75, 259–293.

AMADOR, M. AND K. BAGWELL (2013): "The Theory of Optimal Delegation With an Application to Tariff Caps," *Econometrica*, 81, 1541–1599.

BLOEDEL, A. AND W. ZHONG (2021): "The Cost of Optimally Acquired Information," *Working Paper*.

BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM (2015): "Supervisory Letter SR 15-18: Federal Reserve Supervisory Assessment of Capital Planning and Positions for Firms Subject to Category I Standards," https://www.federalreserve.gov/supervisionreg/srletters/sr1518.htm, accessed April 18, 2026.

DOVAL, L. AND V. SKRETA (2023): "Constrained Information Design," *Mathematics of Operations Research*.

GUO, Y. (2016): "Dynamic Delegation of Experimentation," *American Economic Review*, 106, 1969–2008.

HANCART, N. (2025): "The Optimal Menu of Tests," Working Paper.

HEDLUND, J. (2017): "Bayesian Persuasion by a Privately Informed Sender," *Journal of Economic Theory*, 167, 229–268.

HOLMSTRÖM, B. (1977): "On Incentives and Control in Organizations," Ph.D. thesis, Stanford University.

JAGADEESAN, R. AND D. VIVIANO (2024): "Publication Design with Incentives in Mind," Working Paper.

KAMENICA, E. AND M. GENTZKOW (2011): "Bayesian Persuasion," *American Economic Review*, 101, 2590–2615.

KARTIK, N., A. KLEINER, AND R. VAN WEELDEN (2021): "Delegation in Veto Bargaining," *American Economic Review*, 111, 4046–87.

KOESSLER, F. AND D. MARTIMORT (2012): "Optimal Delegation with Multi-Dimensional Decisions," *Journal of Economic Theory*, 147, 1850–1881.

KOSENKO, A. (2023): "Constrained Persuasion with Private Information," *The B.E. Journal of Theoretical Economics*, 23, 345–370.

MASKIN, E. AND J. RILEY (1984): "Monopoly with Incomplete Information," *The RAND Journal of Economics*, 15, 171–196.

MCCLELLAN, A. (2017): "Experimentation and Approval Mechanisms," Unpublished paper, New York University.

——— (2022): "Experimentation and Approval Mechanisms," *Econometrica*, 90, 2215–2247.

MORRIS, S. AND P. STRACK (2019): "The Wald Problem and the Relation of Sequential Sampling and Ex-Ante Information Costs," *Available at SSRN 2991567*.

MUSSA, M. AND S. ROSEN (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301–317.

MYERSON, R. B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

POMATTO, L., P. STRACK, AND O. TAMUZ (2023): "The Cost of Information: The Case of Constant Marginal Costs," *American Economic Review*.

RAPPOPORT, D. AND V. SOMMA (2017): "Incentivizing Information Design," *Available at SSRN 3001416*.

ROCKAFELLAR, R. T. (1970): *Convex Analysis*, Princeton Mathematical Series, Princeton, N. J.: Princeton University Press.

SHARMA, S., E. TSAKAS, AND M. VOORNEVELD (2024): "Procuring Unverifiable Information," *Mathematics of Operations Research*.

TETENOV, A. (2016): "An Economic Theory of Statistical Testing," CeMMAP working papers CWP50/16, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

U.S. FOOD AND DRUG ADMINISTRATION (2009): "End-of-Phase 2A Meetings: Guidance for Industry," https://www.fda.gov/media/72211/download, accessed May 27, 2026.

——— (2018): "Special Protocol Assessment: Guidance for Industry," https://www.fda.gov/regulatory-information/search-fda-guidance-documents/special-protocol-assessment-guidance-industry, accessed April 18, 2026.

——— (2023): "Clinical Trial Considerations to Support Accelerated Approval of Oncology Therapeutics: Draft Guidance for Industry," https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-trial-considerations-support-accelerated-approval-oncology-therapeutics, accessed May 27, 2026.

——— (2024a): "IND Application Procedures: Clinical Hold," https://www.fda.gov/drugs/investigational-new-drug-ind-application/ind-application-procedures-clinical-hold, accessed April 18, 2026.

——— (2024b): "IND Applications for Clinical Investigations: Clinical Protocols," https://www.fda.gov/drugs/investigational-new-drug-ind-application/ind-applications-clinical-investigations-clinical-protocols, accessed April 18, 2026.

WALD, A. (1947): *Sequential Analysis*, John Wiley.

WANG, H. (2023): "Contracting with Heterogeneous Researchers," *arXiv preprint arXiv:2307.07629*.

WHITMEYER, M. AND K. ZHANG (2022): "Buying Opinions," *arXiv preprint arXiv:2202.05249*.

YODER, N. (2022): "Designing Incentives for Heterogeneous Researchers," *Journal of Political Economy*, 130, 2018–2054.

YOUDEN, W. J. (1950): "Index for Rating Diagnostic Tests," *Cancer*, 3, 32–35.