Online Appendix for "Designing Incentives for Heterogeneous Researchers"

Nathan Yoder*

October 10, 2021

S.1 Intrinsic Researcher Motivation

In some applications, researchers have an intrinsic motivation to gather information. For instance, a pharmaceutical firm may have a significant interest in the regulatory approval of their vaccine, and so may undertake clinical trials even in the absence of a government subsidy. This section extends my model to accommodate these applications, by relaxing the assumption that the researcher is disinterested in the result of her experiment.

Suppose that when the researcher's experiment induces the belief p, the researcher receives indirect utility of $\pi(p)$, in addition to any compensation she might receive from the researcher and any costs associated with her experiment. As with the principal, this indirect utility might represent a payoff from persuasion, from optimally choosing an action, or from some non-instrumental benefit of information. Then the value that the researcher receives from the experiment τ, exclusive of cost and transfers from the principal, is given by $\Pi(\tau) \equiv E_\tau[\pi(p) - \pi(p_0)]$. Hence, when a type-$\theta$ researcher conducts the experiment τ and receives t from the principal, she now receives a payoff of $t + \Pi(\tau) - C(\tau, \theta)$. We say that a contract *implements* $\chi : \Theta \to X$ *when the researcher is intrinsically motivated* if a type-θ intrinsically motivated researcher who accepts that contract and conducts the experiment $\chi(\theta)$ gets an expected payoff which is at least as large as the expected payoff from conducting any other experiment, either after accepting the contract or after rejecting it.

This affects both the incentive compatibility and participation constraints facing the principal. Define $\overline{\chi}(\theta) \in \arg\max_\tau\{E_\tau[\pi(p) - \theta c(p)] \text{ s.t. } E_\tau p = p_0\}$. Then when the re-

*University of Georgia, Terry College of Business, Department of Economics; E-mail: nathan.yoder@uga.edu.

searcher is intrinsically motivated, the methods-based contracting problem (1) becomes

$$\max_{\chi,t} E_F[E_{\chi(\theta)}[w(p)] - t(\theta)] \tag{S.1}$$

$$\text{s.t. } E_{\chi(\theta)}p = p_0 \qquad \forall\theta \in \Theta,$$

$$t(\theta) + \Pi(\chi(\theta)) - C(\chi(\theta),\theta) \geq t(\theta') + \Pi(\chi(\theta')) - C(\chi(\theta'),\theta) \qquad \forall\theta,\theta' \in \Theta, \qquad (\widetilde{\text{IC}\theta})$$

$$t(\theta) + \Pi(\chi(\theta)) - C(\chi(\theta),\theta) \geq \Pi(\overline{\chi}(\theta)) - C(\overline{\chi}(\theta),\theta) \qquad \forall\theta \in \Theta. \qquad (\widetilde{\text{IR}\theta})$$

If the only difference between (S.1) and the contracting problem (1) from the main text was the replacement of the incentive compatibility constraint (ICθ) with ($\widetilde{\text{IC}\theta}$), the difference between their solutions would be minimal: By subtracting $\Pi(\chi(\theta))$ from the transfers $t(\theta)$ found in (2) and (3), and adding $\pi(p)$ to the principal's objective function in the reduced problem (4), the analysis from Section 3 would go through essentially unchanged. The replacement of the participation constraint (IRθ) by ($\widetilde{\text{IR}\theta}$) is far more consequential: Different types now have different outside options, as in Jullien (2000), and so participation constraints for types other than θ_1 may be relevant. If so, we cannot pin down t and reduce the principal's contracting problem the way we did in Section 3.

However, there is a condition under which the results from the basic model continue to apply. Proposition S.1 shows that as long as the principal wants each researcher type to conduct a more costly experiment than the *next-highest* cost type would conduct in the absence of a subsidy, only the highest-cost type's participation constraint can bind.

Proposition S.1 (Reducing the Principal's Problem with an Intrinsically Motivated Researcher)**.** *Suppose that $E_{\chi(\theta_{i-1})}[c(p)] \geq E_{\overline{\chi}(\theta_i)}[c(p)]$ for each $i \in \{2,\dots,N\}$. Then (χ, t) solves the principal's problem when the researcher is intrinsically motivated (S.1) if and only if t is given by*

$$t(\theta_1) = \theta_1 E_{\chi(\theta_1)}[c(p)] - \Pi(\chi(\theta_1)) + \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1),\theta_1), \tag{S.2}$$

$$t(\theta_i) = \begin{cases} \theta_i E_{\chi(\theta_i)}[c(p)] + \sum_{j=1}^{i-1}(\theta_j - \theta_{j+1})E_{\chi(\theta_j)}[c(p)] \\ \qquad -\Pi(\chi(\theta_i)) + \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1),\theta_1) \end{cases} \quad \forall i \in \{2,\dots,N\}, \tag{S.3}$$

and χ solves

$$\max_{\chi} E_F\left[E_{\chi(\theta)}\left[w(p) + \pi(p) - g(\theta)c(p)\right]\right] \tag{S.4}$$

$$s.t.\ E_{\chi(\theta)}p = p_0 \qquad \forall\theta \in \Theta,$$

$$E_{\chi(\theta_{i-1})}[c(p)] \leq E_{\chi(\theta_i)}[c(p)] \qquad \forall i \in \{2,\dots,N\}. \tag{S.5}$$

Corollary S.1 (Methods-Based Contracting with an Intrinsically Motivated Researcher)**.**

When the researcher is intrinsically motivated, and the principal's reduced problem (S.4) has a solution $\hat{\chi}$ such that $E_{\hat{\chi}(\theta_{i-1})}[c(p)] \geq E_{\overline{\chi}(\theta_i)}[c(p)]$ for each $i \in \{2, \ldots, N\}$:

i. *(Theorem 1) The principal has an optimal methods-based contract which implements a binary, Blackwell-monotone experiment choice function χ^*; and*

ii. *(Theorem 2) If the researcher's virtual type $g(\theta)$ is increasing in θ, then χ^* exhibits no distortion at the top, garbling everywhere else, and distortion lowers cost to the researcher and value for the principal.*

This condition also ensures that results-based implementation works largely the same way as in the basic model, with minor modifications to the implementing contract ψ_χ constructed in Section 4. These are simple accounting changes: First, increase compensation for each result by the amount of the highest-cost researcher's reservation payoff. Then, reduce compensation for each result by the change that that result causes in the researcher's intrinsic payoff. More formally, offer the contract $\overline{\psi}_\chi$, where

$$\overline{\psi}_\chi(p) \equiv \psi_\chi(p) - (\pi(p) - \pi(p_0)) + \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta_1).$$

Proposition S.2 shows that with an intrinsically motivated researcher, this contract implements χ, and does so at an expected cost no greater than that of the lowest-cost implementing methods-based contract.

Proposition S.2 (Results-Based Implementation with an Intrinsically Motivated Researcher)**.** *If $\chi : \Theta \to X$ is binary and Blackwell-monotone, and $E_{\chi(\theta_{i-1})}[c(p)] \geq E_{\overline{\chi}(\theta_i)}[c(p)]$ for each $i \in \{2, \ldots, N\}$, then χ can be implemented with a results-based contract $\overline{\psi}_\chi$ when the researcher is intrinsically motivated. Furthermore, $\overline{\psi}_\chi$ cannot be outperformed by a methods-based contract T: If T implements χ when the researcher is intrinsically motivated, then $E_{\chi(\theta)}\overline{\psi}_\chi(p) \leq T(\chi(\theta))$ for each $\theta \in \Theta$.*

S.1.1 Proofs for Section S.1

Lemma S.1. *Suppose that $\chi : \Theta \to X$ has the property that $E_{\chi(\theta_{i-1})}[c(p)] \geq E_{\overline{\chi}(\theta_i)}[c(p)]$ for each $i \in \{2, \ldots, N\}$. Then if (χ, t) satisfies ($\widetilde{IR\theta}$) for $\theta = \theta_1$, and for each $i \in \{2, \ldots, N\}$, satisfies ($\widetilde{IC\theta}$) for $(\theta, \theta') = (\theta_i, \theta_{i-1})$, then (χ, t) satisfies ($\widetilde{IR\theta}$).*

Proof. First note that by definition of $\overline{\chi}$, $\Pi(\overline{\chi}(\theta)) - C(\overline{\chi}(\theta), \theta) \geq \Pi(\overline{\chi}(\theta')) - C(\overline{\chi}(\theta'), \theta)$ for each $\theta, \theta' \in \Theta$. By assumption, ($\widetilde{IR\theta}$) holds for $\theta = \theta_1$. For $i \in \{2, \ldots, N\}$, suppose ($\widetilde{IR\theta}$) is satisfied for $\theta = \theta_{i-1}$. Then since ($\widetilde{\text{IC}\theta}$) holds for $(\theta, \theta') = (\theta_i, \theta_{i-1})$, we have

$$\begin{aligned} t(\theta_i) + \Pi(\chi(\theta_i)) - C(\chi(\theta_i), \theta_i) &\geq t(\theta_{i-1}) + \Pi(\chi(\theta_{i-1})) - C(\chi(\theta_{i-1}), \theta_i) \\ &= t(\theta_{i-1}) + \Pi(\chi(\theta_{i-1})) - C(\chi(\theta_{i-1}), \theta_{i-1}) + (\theta_{i-1} - \theta_i) E_{\chi(\theta_{i-1})}[c(p)] \\ &\geq \Pi(\overline{\chi}(\theta_{i-1})) - C(\overline{\chi}(\theta_{i-1}), \theta_{i-1}) + (\theta_{i-1} - \theta_i) E_{\chi(\theta_{i-1})}[c(p)] \\ &\geq \Pi(\overline{\chi}(\theta_i)) - C(\overline{\chi}(\theta_i), \theta_{i-1}) + (\theta_{i-1} - \theta_i) E_{\chi(\theta_{i-1})}[c(p)] \\ &= \Pi(\overline{\chi}(\theta_i)) - C(\overline{\chi}(\theta_i), \theta_i) + (\theta_{i-1} - \theta_i)(E_{\chi(\theta_{i-1})}[c(p)] - E_{\overline{\chi}(\theta_i)}[c(p)]) \\ &\geq \Pi(\overline{\chi}(\theta_i)) - C(\overline{\chi}(\theta_i), \theta_i), \end{aligned}$$

and ($\widetilde{\text{IR}\theta}$) is satisfied for $\theta = \theta_i$. Then by induction, ($\widetilde{\text{IR}\theta}$) is satisfied for all θ. □

Lemma S.2. *Suppose that $\chi : \Theta \to X$ has the property that $E_{\chi(\theta_{i-1})}[c(p)] \geq E_{\overline{\chi}(\theta_i)}[c(p)]$ for each $i \in \{2, \ldots, N\}$.*

i. If χ : satisfies the monotonicity constraints (S.5) and $t : \Theta \to \mathbb{R}$ is given by (S.2) and (S.3), then (χ, t) satisfies ($\widetilde{\text{IC}\theta}$) and ($\widetilde{\text{IR}\theta}$).

ii. For any direct revelation contract (χ, t) that satisfies ($\widetilde{\text{IC}\theta}$) and ($\widetilde{\text{IR}\theta}$),

(a) $E_{\chi(\theta_{i-1})}[c(p)] \leq E_{\chi(\theta_i)}[c(p)]$ for all $i \in \{2, \ldots, N\}$.

(b) If $\hat{t}$ is derived from χ according to (S.2) and (S.3), then $t(\theta) \geq \hat{t}(\theta)$ for each $\theta \in \Theta$. Moreover, $(\chi, \hat{t})$ satisfies ($\widetilde{\text{IC}\theta}$) and ($\widetilde{\text{IR}\theta}$).

Proof. The proof proceeds analogously to the proof of Lemma 4.

i. For each $i \in \{2, \ldots, N\}$ and $\ell < i$, (S.2) and (S.3) yield

$$\begin{aligned} &t(\theta_i) + \Pi(\chi(\theta_i)) - C(\chi(\theta_i), \theta_i) \\ &- (t(\theta_\ell) + \Pi(\chi(\theta_\ell)) - C(\chi(\theta_\ell), \theta_i)) \end{aligned} = \sum_{j=\ell}^{i-1} (\theta_j - \theta_{j+1}) E_{\chi(\theta_j)}[c(p)] - (\theta_\ell - \theta_i) E_{\chi(\theta_\ell)}[c(p)]$$
$$= \sum_{j=\ell}^{i-1} (\theta_j - \theta_{j+1}) \left(E_{\chi(\theta_j)}[c(p)] - E_{\chi(\theta_\ell)}[c(p)] \right) \geq 0, \quad \text{(by (S.5))}$$

and for each $i \in \{1, \ldots, N-1\}$ and $\ell > i$,

$$\begin{aligned} &t(\theta_i) + \Pi(\chi(\theta_i)) - C(\chi(\theta_i), \theta_i) \\ &- (t(\theta_\ell) + \Pi(\chi(\theta_\ell)) - C(\chi(\theta_\ell), \theta_i)) \end{aligned} = (\theta_i - \theta_\ell) E_{\chi(\theta_\ell)}[c(p)] - \sum_{j=i}^{\ell-1} (\theta_j - \theta_{j+1}) E_{\chi(\theta_j)}[c(p)]$$
$$= \sum_{j=i}^{\ell-1} (\theta_j - \theta_{j+1}) \left(E_{\chi(\theta_\ell)}[c(p)] - E_{\chi(\theta_j)}[c(p)] \right) \geq 0, \quad \text{(by (S.5))}$$

implying ($\widetilde{\text{IC}\theta}$). From (S.2), (χ, t) satisfies ($\widetilde{\text{IR}\theta}$) for $\theta = \theta_1$; then by Lemma S.1, (χ, t) satisfies ($\widetilde{\text{IR}\theta}$).

ii. (a) From ($\widetilde{\text{IC}\theta}$), for each $i \in \{2, \ldots, N\}$,

$$
\begin{aligned}
\theta_{i-1}\left(E_{\chi(\theta_i)}[c(p)] - E_{\chi(\theta_{i-1})}[c(p)]\right) &\geq t(\theta_i) + \Pi(\chi(\theta_i)) - t(\theta_{i-1}) - \Pi(\chi(\theta_{i-1})), \\
\theta_i\left(E_{\chi(\theta_i)}[c(p)] - E_{\chi(\theta_{i-1})}[c(p)]\right) &\leq t(\theta_i) + \Pi(\chi(\theta_i)) - t(\theta_{i-1}) - \Pi(\chi(\theta_{i-1})); \\
\Rightarrow (\theta_{i-1} - \theta_i)\left(E_{\chi(\theta_i)}[c(p)] - E_{\chi(\theta_{i-1})}[c(p)]\right) &\geq 0; \\
\Leftrightarrow E_{\chi(\theta_i)}[c(p)] - E_{\chi(\theta_{i-1})}[c(p)] &\geq 0,
\end{aligned}
$$

as desired.

(b) Since $\hat{t}$ is derived from (S.2) and (S.3), we have

$$\hat{t}(\theta_1) = C(\chi(\theta_1), \theta_1) - \Pi(\chi(\theta_1)) + \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta_1),$$

and for each $i \in \{2, \ldots, N\}$,

$$
\begin{aligned}
\hat{t}(\theta_i) - \hat{t}(\theta_{i-1}) &= \theta_i\left(E_{\chi(\theta_i)}[c(p)] - E_{\chi(\theta_{i-1})}[c(p)]\right) + \Pi(\chi(\theta_i)) - \Pi(\chi(\theta_{i-1})) \\
&= C(\chi(\theta_i), \theta_i) - C(\chi(\theta_{i-1}), \theta_i) + \Pi(\chi(\theta_i)) - \Pi(\chi(\theta_{i-1})).
\end{aligned}
$$

Since (χ, t) satisfies ($\widetilde{\text{IR}\theta}$) for $\theta = \theta_1$, we have

$$t(\theta_1) \geq C(\chi(\theta_1), \theta_1) - \Pi(\chi(\theta_1)) + \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta_1) = \hat{t}(\theta_1).$$

For each $i \in \{2, \ldots, N\}$, since (χ, t) satisfies ($\widetilde{\text{IC}\theta}$) for $\theta = \theta_i$ and $\theta' = \theta_{i-1}$, we have

$$
\begin{aligned}
t(\theta_i) - t(\theta_{i-1}) &\geq C(\chi(\theta_i), \theta_i) - C(\chi(\theta_{i-1}), \theta_i) + \Pi(\chi(\theta_i)) - \Pi(\chi(\theta_{i-1})) \\
&= \hat{t}(\theta_i) - \hat{t}(\theta_{i-1}).
\end{aligned}
$$

Then for each $j \in \{2, \ldots, N\}$, we have

$$t(\theta_j) = t(\theta_1) + \sum_{i=2}^{N}(t(\theta_i) - t(\theta_{i-1})) \geq \hat{t}(\theta_1) + \sum_{i=2}^{N}(\hat{t}(\theta_i) - \hat{t}(\theta_{i-1})) = \hat{t}(\theta_j).$$

Finally, from parts (iia) and (i), $(\chi, \hat{t})$ satisfies ($\widetilde{\text{IC}\theta}$) and ($\widetilde{\text{IR}\theta}$).

□

Proof of Proposition S.1 (Reducing the Principal's Problem with an Intrinsically Motivated Researcher) Follows identically to the proof of Corollary 1, replacing Lemma 4 with Lemma S.2. □

Proof of Corollary S.1 (Methods-Based Contracting with an Intrinsically Motivated Researcher) Observe that the principal's reduced problem when the researcher is intrinsically motivated (S.4) is just the principal's reduced problem from the basic model (4) with w replaced by $w + \pi$. Then by Proposition 3, if we replace w with $w + \pi$ in the ironed problem (12), it is solved by $\hat{\chi}$.

For each θ, $\hat{\chi}(\theta)$ is a Borel distribution, so $\operatorname{supp} \hat{\chi}(\theta)$ is closed, and $\min \operatorname{supp} \hat{\chi}(\theta)$ and $\max \operatorname{supp} \hat{\chi}(\theta)$ exist. Then $\min \operatorname{supp} \hat{\chi}(\theta) \leq p \leq \max \operatorname{supp} \hat{\chi}(\theta)$ for each $p \in \operatorname{supp} \hat{\chi}(\theta)$, and so $\min \operatorname{supp} \hat{\chi}(\theta) \leq E_{\hat{\chi}(\theta)} p = p_0 \leq \max \operatorname{supp} \hat{\chi}(\theta)$. Then there exists a Bayes-plausible distribution $\chi^b(\theta)$ with $\operatorname{supp} \chi^b(\theta) \subseteq \{\min \operatorname{supp} \hat{\chi}(\theta), \max \operatorname{supp} \hat{\chi}(\theta)\}$; this $\chi^b(\theta)$ is a mean-preserving spread of $\hat{\chi}$ in the sense of Machina and Pratt (1997), and by Machina and Pratt (1997) Theorem 3, in the usual stochastic dominance sense as well.

By Lemma 3, $w(p) + \pi(p) - \bar{g}(\theta)c(p)$ must coincide on $\operatorname{supp} \hat{\chi}(\theta)$ with its concavification, which is affine on $[\min \operatorname{supp} \hat{\chi}(\theta), \max \operatorname{supp} \hat{\chi}(\theta)]$. Since this interval contains p_0, the expectation of $w(p) + \pi(p) - \bar{g}(\theta)c(p)$ under $\chi^b(\theta)$ must be equal to its concavification's value at p_0. It follows from Kamenica and Gentzkow (2011) Online Appendix Proposition 3 that if we replace w with $w + \pi$ in the ironed problem (12), it is solved by $\chi^b : \Theta \to X$.

Then define $\chi^*(\theta) = \chi^b(\min\{\theta' | \bar{g}(\theta') = \bar{g}(\theta)\})$. By definition, χ^* is binary, and $\chi^*(\theta_i) = \chi^*(\theta_{i-1})$ whenever $\bar{g}(\theta_i) = \bar{g}(\theta_{i-1})$. Since the type-specific ironed problems in (12) are the same for each θ, θ' with $\bar{g}(\theta) = \bar{g}(\theta')$, it follows that if we replace w with $w + \pi$ in the ironed problem (12), it is solved by χ^*. Then by Proposition 3, if we replace w with $w + \pi$ in the principal's reduced problem from the basic model (4), it is solved by χ^*; equivalently, χ^* solves (S.4). Moreover, by Lemma 1 (i), χ^* is Blackwell-monotone.

Now by Proposition 1, $\overline{\chi}$ is Blackwell-monotone. Then since for each $i \in \{2, \ldots, N\}$, we have $E_{\chi^*(\theta_{i-1})}[c(p)] = E_{\chi^b(\theta_{j-1})}[c(p)]$ for some $j \geq i$, it follows that

$$E_{\chi^*(\theta_{i-1})}[c(p)] = E_{\chi^b(\theta_{j-1})}[c(p)] \geq E_{\hat{\chi}(\theta_{j-1})}[c(p)] \geq E_{\overline{\chi}(\theta_j)}[c(p)] \geq E_{\overline{\chi}(\theta_i)}[c(p)],$$

since $\chi^b(\theta_{j-1})$ is a mean-preserving spread of $\hat{\chi}(\theta_{j-1})$. Then by Proposition S.1, χ^* solves the principal's problem when the researcher is intrinsically motivated; since χ^* is binary and Blackwell-monotone, (i) follows. And since χ^* solves the principal's reduced problem (4) from the basic model with w replaced by $w + \pi$, (ii) follows identically to the proof of Theorem 2. □

Proof of Proposition S.2 (Results-Based Implementation with and Intrinsically Motivated Researcher)

By Theorem 3, when the researcher is not intrinsically motivated, there exists a methods-based contract ψ_χ which implements χ and such that $E_{\chi(\theta)}[\psi_\chi(p)] \leq \hat{T}(\chi(\theta))$ for any results-based contract $\hat{T}$ that implements χ. Then define $\overline{\psi}_\chi(p) \equiv \psi_\chi(p) - (\pi(p) - \pi(p_0)) +$

$\Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta_1)$. When a type-$\theta$ intrinsically motivated researcher accepts the results-based contract $\overline{\psi}_\chi$, she solves

$$\max_{\tau \in \Delta([0,1])} \{E_\tau[\overline{\psi}_\chi(p) + \pi(p) - \pi(p_0) - \theta c(p)] \text{ s.t. } E_\tau p = p_0\}$$
$$= \max_{\tau \in \Delta([0,1])} \{E_\tau[\psi_\chi(p) - \theta c(p)] + \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta_1) \text{ s.t. } E_\tau p = p_0\}.$$

Now by Theorem 3,

$$\chi(\theta) \in \arg \max_{\tau \in \Delta([0,1])} \{E_\tau[\psi_\chi(p) - \theta c(p)] \text{ s.t. } E_\tau p = p_0\}$$
$$= \arg \max_{\tau \in \Delta([0,1])} \{E_\tau[\psi_\chi(p) - \theta c(p)] + \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta_1) \text{ s.t. } E_\tau p = p_0\}.$$

It follows that when a type-θ intrinsically motivated researcher accepts the results-based contract $\overline{\psi}_\chi$, her expected payoff from conducting the experiment $\chi(\theta)$ is weakly larger than her expected payoff from conducting any other experiment.

Now let $r(\theta) = E_{\chi(\theta)}[\overline{\psi}_\chi(p)]$; it follows immediately that (χ, r) satisfies ($\widetilde{\text{IC}\theta}$). Moreover, since ψ_χ implements χ when the researcher is not intrinsically motivated, we have $E_{\chi(\theta_1)}[\psi_\chi(p)] \geq C(\chi(\theta_1), \theta_1)$; then

$$\begin{aligned} r(\theta_1) + \Pi(\chi(\theta_1)) - C(\chi(\theta_1), \theta_1) &= E_{\chi(\theta_1)}[\psi_\chi(p) - (\pi(p) - \pi(p_0))] + \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta_1) \\ &\quad + \Pi(\chi(\theta_1)) - C(\chi(\theta_1), \theta_1) \\ &= E_{\chi(\theta_1)}[\psi_\chi(p)] - C(\chi(\theta_1), \theta_1) + \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta_1) \\ &\geq \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta_1) \end{aligned}$$

and (χ, r) satisfies ($\widetilde{\text{IR}\theta}$) for $\theta = \theta_1$. Then by Lemma S.1, (χ, r) satisfies ($\widetilde{\text{IR}\theta}$). It follows that when a type-θ intrinsically motivated researcher accepts the results-based contract $\overline{\psi}_\chi$, her expected payoff from conducting the experiment $\chi(\theta)$ is weakly larger than her payoff from conducting any experiment when she refuses the contract.

Then $\overline{\psi}_\chi$ implements χ when the researcher is intrinsically motivated. Now suppose that T implements χ when the researcher is intrinsically motivated. Then by the revelation principle, so does the direct revelation contract (χ, t) with $t(\theta) = T(\chi(\theta))$. Then (χ, t) must satisfy ($\widetilde{\text{IC}\theta}$) and ($\widetilde{\text{IR}\theta}$). Now define $t_0 : \Theta \to \mathbb{R}$ by $t_0(\theta) = t(\theta) + \Pi(\chi(\theta)) - (\Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta_1))$. Then since (χ, t) satisfies ($\widetilde{\text{IC}\theta}$), (χ, t_0) satisfies ($\text{IC}\theta$). By definition of $\overline{\chi}$, for any $\theta \in \Theta$,

$$\begin{aligned}\Pi(\overline{\chi}(\theta)) - C(\overline{\chi}(\theta), \theta) &\geq \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta) \\ &= \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta_1) + (\theta_1 - \theta) E_{\overline{\chi}(\theta_1)}[c(p)] \\ &\geq \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta_1).\end{aligned}$$

Then since (χ, t) satisfies ($\widetilde{\text{IR}\theta}$), (χ, t_0) satisfies (IRθ). Then (χ, t_0) implements χ when the researcher is not intrinsically motivated; by the taxation principle, so does $T_0 : X \to \mathbb{R}$ with

$$T_0(\tau) = \begin{cases} t_0(\theta), & \tau = \chi(\theta); \\ 0, & \tau \notin \chi(\Theta). \end{cases}$$

Then by Theorem 3, we have $E_{\chi(\theta)}[\psi_\chi(p)] \leq T_0(\chi(\theta)) = t_0(\theta)$. Then by construction,

$$\begin{aligned}E_{\chi(\theta)}[\overline{\psi}_\chi(p)] &= E_{\chi(\theta)}[\psi_\chi(p)] - \Pi(\chi(\theta)) + \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta_1) \\ &\leq t_0(\theta) - \Pi(\chi(\theta)) + \Pi(\overline{\chi}(\theta_1)) - C(\overline{\chi}(\theta_1), \theta_1) \\ &= t(\theta) = T(\chi(\theta)),\end{aligned}$$

as desired. □

S.2 Contingency Fees and Failed Settlement Negotiations

As discussed in Section 5.2, plaintiff's attorneys are often compensated for their work using a contingency fee schedule which offers payment as a function of the amount recovered from the defendant. Here, I extend the example from the main text to allow for some probability that settlement negotiations break down and the parties proceed to trial.

Just as before, adopt the accounting convention that observable costs are assigned to the principal, and let the net settlement amount σ be a strictly increasing function of the posterior probability p that the defendant is liable. In addition, let the probability r that no settlement will be reached, the probability q that the plaintiff will win at trial, and the attorney's observable costs c_o also be real-valued functions of p, with $r(p) < 1$ for each p. Finally, suppose that when liability is certain, the probability $q(1)$ of winning at trial is one, the gross settlement $\sigma(1) + c_o(1)$ is equal to the damages that would be awarded at trial, and the attorney's expenses $c_o(1)$ are at least as high as the expenses she would incur for any other result.

Then the expected payment (net of observable costs) to an attorney who accepts the contingency fee schedule ϕ and produces result p is

$$\psi_\phi(p) \equiv r(p)(q(p)\phi(\sigma(1) + c_o(1) - c_o(p)) - (1 - q(p))c_o(p)) + (1 - r(p))\phi(\sigma(p)). \quad \text{(S.6)}$$

Thus, as in the example from the main text, the contingency fee schedule ϕ offers the same incentives for pretrial information acquisition as the results-based contract ψ_ϕ. Proposition S.3 shows that just as before, the fact that the net settlement σ is one-to-one ensures that the two types of contracts are equivalent.

Proposition S.3 (Equivalence of Results-Based and Contingency Fee Contracting)**.** *In the setting of Section S.2, the classes of results-based contracts and contingency fee schedules are equivalent: For every contingency fee ϕ, there is a results-based contract ψ_ϕ that yields the same expected payment for any experiment, and for every results-based contract ψ, there is a contingency fee ϕ_ψ that yields the same expected payment for any experiment.*

Proof. The first part follows immediately from letting ψ_ϕ be as in (S.6). For the second part, choose

$$\phi_\psi(x) \equiv \begin{cases} \frac{\psi(\sigma^{-1}(x)) - r(\sigma^{-1}(x))(q(\sigma^{-1}(x))\psi(1) - (1-q(\sigma^{-1}(x)))c_o(\sigma^{-1}(x)))}{1 - r(\sigma^{-1}(x))}, & x \leq \sigma(1) \\ \psi(1), & x > \sigma(1). \end{cases}$$

Since $c_o(1) \geq c_o(p)$ for each $p \in [0,1]$, $\phi_\psi(\sigma(1) + c_o(1) - c_o(p)) = \psi(1)$. Then from (S.6), if an attorney produces result p, she receives expected payment

$$\begin{aligned} r(p)(q(p)\phi_\psi(\sigma(1) + c_o(1) - c_o(p)) - (1 - q(p))c_o(p)) + (1 - r(p))\phi_\psi(\sigma(p)) \\ = r(p)\left(q(p)\psi(1) - (1 - q(p))c_o(p)\right) \\ + (1 - r(p))\left(\frac{\psi(p) - r(p)(q(p)\psi(1) - (1 - q(p))c_o(p))}{1 - r(p)}\right) \\ = \psi(p), \end{aligned}$$

as desired. □

S.3 Characterization Without Semicontinuity

The next two lemmas allow the assumption of upper semicontinuity to be dropped in Proposition 1. First, Lemma S.3 shows in the binary-state context that any solution to a persuasion problem remains a solution when the value function is replaced by its upper semicontinuous hull. Second, Lemma S.4 shows that the sum of a function's upper semicontinuous hull and another continuous function is the upper semicontinuous hull of the sum of both functions. This ensures that greater additive concavity is preserved by taking upper semicontinuous hulls. Together, Lemmas S.3 and S.4 allow us to apply Proposition 4 to the upper semicontinuous hulls of the value functions in the two problems whose solutions we wish to compare in Proposition 1.

Lemma S.3 (Characterization Without Semicontinuity in Binary-State Problems)**.** *Let $\bar{v}$ be the upper semicontinuous hull of $v : [0,1] \to \mathbb{R}$, i.e., $\bar{v}(p) \equiv \max\{z|(p,z) \in \mathrm{cl}(\mathrm{hypo}(v))\}$; and let $\bar{V}$ be the concavification of $\bar{v}$. Suppose that $p_0 \in (0,1)$. If $\tau^* \in \arg\max_{\tau \in \Delta([0,1])}\{E_\tau v(p)$ s.t. $E_\tau p = p_0\}$, then $\tau^* \in \arg\max_{\tau\in\Delta([0,1])}\{E_\tau \bar{v}(p)$ s.t. $E_\tau p = p_0\}$ and $E_{\tau^*} v(p) = \bar{V}(p_0)$.*

Proof. By Lemma 2 and Kamenica and Gentzkow (2011) Online Appendix Propositions 3 and 4, either (i) $\bar{V}(p_0) = \bar{v}(p_0)$, or (ii) there exist $\underline{p}, \overline{p} \in [0,1]$ with $\underline{p} < p_0 < \overline{p}$ such that $\bar{V}(p_0) = \frac{\overline{p}-p_0}{\overline{p}-\underline{p}}\bar{v}(\underline{p}) + \frac{p_0-\underline{p}}{\overline{p}-\underline{p}}\bar{v}(\overline{p})$.

First suppose that (i) holds. Since $(p_0, \bar{v}(p_0)) \in \mathrm{cl}(\mathrm{hypo}(v))$, there exists a sequence $\{(y_n, z_n)\}_{n=1}^\infty \subset \mathrm{hypo}(v)$ such that $(y_n, z_n) \to (p_0, \bar{v}(p_0))$. Then either (i(a)) there exists M such that $y_n < p_0$ for all $n \geq M$; (i(b)) there exists M such that $y_n > p_0$ for all $n \geq M$; (i(c)) there does not exist M such that $y_n \neq p_0$ for all $n \geq M$; or (i(d)) there exists M such that $y_n \neq p_0$ for all $n \geq M$, but neither (i(a)) nor (i(b)) hold.

If (i(c)) holds, there exists a subsequence $\{(y_{n_k}, z_{n_k})\}_{k=1}^\infty$ such that $p_{n_k} = p_0$ for all k. Then by Kamenica and Gentzkow (2011) Online Appendix Proposition 3 we have

$$\begin{aligned}
&\bar{V}(p_0) \geq E_{\tau^*}\bar{v}(p) \geq E_{\tau^*}v(p) = \max_{\tau\in\Delta([0,1])}\{E_\tau v(p) \text{ s.t. } E_\tau p = p_0\} \geq v(p_0) \geq z_{n_k} \\
\Rightarrow\ &\bar{V}(p_0) \geq E_{\tau^*}\bar{v}(p) \geq E_{\tau^*}v(p) = \max_{\tau\in\Delta([0,1])}\{E_\tau v(p) \text{ s.t. } E_\tau p = p_0\} \geq v(p_0) \geq \bar{V}(p_0) \\
\Rightarrow\ &\bar{V}(p_0) = E_{\tau^*}\bar{v}(p) = E_{\tau^*}v(p).
\end{aligned}$$

If (i(a)) or (i(d)) holds, there exists a subsequence $\{(y_{n_k}, z_{n_k})\}_{k=1}^\infty$ such that $y_{n_k} < p_0$ for all k; then let $\underline{y} = p_0$ and $\overline{y} = \frac{1+p_0}{2}$ and define $(\underline{y}_k, \underline{z}_k) = (y_{n_k}, z_{n_k})$ and $(\overline{y}_k, \overline{z}_k) = (\overline{y}, v(\overline{y}))$. Then $\lim_{k\to\infty}(\underline{y}_k, \underline{z}_k) = (\underline{y}, \bar{v}(\underline{y}))$ and $\frac{\overline{y}-p_0}{\overline{y}-\underline{y}}\bar{v}(\underline{y}) + \frac{p_0-\underline{y}}{\overline{y}-\underline{y}}\bar{v}(\overline{y}) = \bar{v}(p_0) = \bar{V}(p_0)$.

If (i(b)) holds, there exists a subsequence $\{(y_{n_k}, z_{n_k})\}_{k=1}^\infty$ such that $y_{n_k} > p_0$ for all k; then let $\overline{y} = p_0$ and $\underline{y} = \frac{p_0}{2}$ and define $(\overline{y}_k, \overline{z}_k) = (y_{n_k}, z_{n_k})$ and $(\underline{y}_k, \underline{z}_k) = (\underline{y}, v(\underline{y}))$. Then $\lim_{k\to\infty}(\overline{y}_k, \overline{z}_k) = (\overline{y}, \bar{v}(\overline{y}))$ and $\frac{\overline{y}-p_0}{\overline{y}-\underline{y}}\bar{v}(\underline{y}) + \frac{p_0-\underline{y}}{\overline{y}-\underline{y}}\bar{v}(\overline{y}) = \bar{v}(p_0) = \bar{V}(p_0)$.

Suppose instead that (ii) holds. Then let $\overline{y} = \overline{p}$ and $\underline{y} = \underline{p}$; it follows that $\frac{\overline{y}-p_0}{\overline{y}-\underline{y}}\bar{v}(\underline{y}) + \frac{p_0-\underline{y}}{\overline{y}-\underline{y}}\bar{v}(\overline{y}) = \bar{V}(p_0)$. By definition of $\bar{v}$, $(\underline{y}, \bar{v}(\underline{y})) \in \mathrm{cl}(\mathrm{hypo}(v))$; since $\underline{y} < p_0$, $(\underline{y}, \bar{v}(\underline{y})) \in \mathrm{cl}(\mathrm{hypo}(v)) \cap ([0,p_0) \times \mathbb{R})$; since $([0,p_0) \times \mathbb{R})$ is open as a subset of $[0,1] \times \mathbb{R}$, $\mathrm{cl}(\mathrm{hypo}(v)) \cap ([0,p_0) \times \mathbb{R}) = \mathrm{cl}(\mathrm{hypo}(v) \cap ([0,p_0) \times \mathbb{R}))$. Then there exists a sequence $\{(\underline{y}_n, \underline{z}_n)\}_{n=1}^\infty \subset \mathrm{hypo}(v) \cap ([0,p_0) \times \mathbb{R})$ such that $(\underline{y}_n, \underline{z}_n) \to (\underline{y}, \bar{v}(\underline{y}))$. By symmetry, there also exists a sequence $\{(\overline{y}_n, \overline{z}_n)\}_{n=1}^\infty \subset \mathrm{hypo}(v) \cap ((p_0, 1] \times \mathbb{R})$ such that $(\overline{y}_n, \overline{z}_n) \to (\overline{y}, \bar{v}(\overline{y}))$.

It follows that if either (i(a)), (i(b)), (i(d)), or (ii) hold, since $\bar{v}(p) \geq v(p)$ for all p by definition, we have

$$
\begin{aligned}
\tfrac{\overline{y}_n - p_0}{\overline{y}_n - \underline{y}_n}\bar{v}(\underline{y}_n) + \tfrac{p_0 - \underline{y}_n}{\overline{y}_n - \underline{y}_n}\bar{v}(\overline{y}_n) &\geq \tfrac{\overline{y}_n - p_0}{\overline{y}_n - \underline{y}_n}v(\underline{y}_n) + \tfrac{p_0 - \underline{y}_n}{\overline{y}_n - \underline{y}_n}v(\overline{y}_n) \geq \tfrac{\overline{y}_n - p_0}{\overline{y}_n - \underline{y}_n}\underline{z}_n + \tfrac{p_0 - \underline{y}_n}{\overline{y}_n - \underline{y}_n}\overline{z}_n \\
\limsup_{n\to\infty} \tfrac{\overline{y}_n - p_0}{\overline{y}_n - \underline{y}_n}\bar{v}(\underline{y}_n) + \tfrac{p_0 - \underline{y}_n}{\overline{y}_n - \underline{y}_n}\bar{v}(\overline{y}_n) &\geq \limsup_{n\to\infty} \tfrac{\overline{y}_n - p_0}{\overline{y}_n - \underline{y}_n}v(\underline{y}_n) + \tfrac{p_0 - \underline{y}_n}{\overline{p}_n - \underline{y}_n}v(\overline{y}_n) \geq \limsup_{n\to\infty} \tfrac{\overline{y}_n - p_0}{\overline{y}_n - \underline{y}_n}\underline{z}_n + \tfrac{p_0 - \underline{y}_n}{\overline{y}_n - \underline{y}_n}\overline{z}_n; \\
\limsup_{n\to\infty} \tfrac{\overline{y}_n - p_0}{\overline{y}_n - \underline{y}_n}\bar{v}(\underline{y}_n) + \tfrac{p_0 - \underline{y}_n}{\overline{y}_n - \underline{y}_n}\bar{v}(\overline{y}_n) &= \left(\lim_{n\to\infty} \tfrac{\overline{y}_n - p_0}{\overline{y}_n - \underline{y}_n}\right)\left(\limsup_{n\to\infty} \bar{v}(\underline{y}_n)\right) + \lim_{n\to\infty}\left(\tfrac{p_0 - \underline{y}_n}{\overline{y}_n - \underline{y}_n}\right)\left(\limsup_{n\to\infty} \bar{v}(\overline{y}_n)\right) \\
&\leq \tfrac{\overline{y} - p_0}{\overline{y} - \underline{y}}\bar{v}(\underline{y}) + \tfrac{p_0 - \underline{y}}{\overline{y} - \underline{y}}\bar{v}(\overline{y}) \text{ (by upper semicontinuity of } \bar{v}); \\
\Rightarrow \tfrac{\overline{y} - p_0}{\overline{y} - \underline{y}}\bar{v}(\underline{y}) + \tfrac{p_0 - \underline{y}}{\overline{y} - \underline{y}}\bar{v}(\overline{y}) &\geq \limsup_{n\to\infty} \tfrac{\overline{y}_n - p_0}{\overline{y}_n - \underline{y}_n}v(\underline{y}_n) + \tfrac{p_0 - \underline{y}_n}{\overline{y}_n - \underline{y}_n}v(\overline{y}_n) \geq \tfrac{\overline{y} - p_0}{\overline{y} - \underline{y}}\bar{v}(\underline{y}) + \tfrac{p_0 - \underline{y}}{\overline{y} - \underline{y}}\bar{v}(\overline{y}) \\
\Rightarrow \bar{V}(p_0) &= \limsup_{n\to\infty} \tfrac{\overline{y}_n - p_0}{\overline{y}_n - \underline{y}_n}v(\underline{y}_n) + \tfrac{p_0 - \underline{y}_n}{\overline{p}_n - \underline{y}_n}v(\overline{y}_n).
\end{aligned}
$$

By Kamenica and Gentzkow (2011) Online Appendix Proposition 3, $\bar{V}(p_0) \geq E_{\tau^*}\bar{v}(p)$. Then we have

$$
\begin{aligned}
E_{\tau^*}v(p) = \sup_{\tau \in \Delta([0,1])} \{E_\tau v(p) \text{ s.t. } E_\tau p = p_0\} &\geq \sup_{m \geq n}\left\{\tfrac{\overline{y}_m - p_0}{\overline{y}_m - \underline{y}_m}v(\underline{y}_m) + \tfrac{p_0 - \underline{y}_m}{\overline{y}_m - \underline{y}_m}v(\overline{y}_m)\right\} \text{ for all } n; \\
\Rightarrow \bar{V}(p_0) \geq E_{\tau^*}\bar{v}(p) \geq E_{\tau^*}v(p) &\geq \limsup_{n\to\infty} \tfrac{\overline{y}_n - p_0}{\overline{y}_n - \underline{y}_n}v(\underline{y}_n) + \tfrac{p_0 - \underline{y}_n}{\overline{y}_n - \underline{y}_n}v(\overline{y}_n) = \bar{V}(p_0), \\
\Leftrightarrow \bar{V}(p_0) = E_{\tau^*}\bar{v}(p) &= E_{\tau^*}v(p).
\end{aligned}
$$

The claim follows. □

Lemma S.4. *Let $S \subseteq \mathbb{R}^n$, and for $v, u : S \to \mathbb{R}$, let $\overline{v}$ and $\overline{v+u}$ be the upper semicontinuous hulls of v and $v+u$, respectively. If u is continuous, then $\overline{v+u} = \overline{v} + u$.*

Proof. By definition, $\overline{v}$ (resp., $\overline{v+u}$) is the smallest upper semicontinuous function that majorizes v (resp., $v+u$). Since u is continuous, $\overline{v}+u$ is upper semicontinuous; since $\overline{v}$ majorizes v, $\overline{v}+u$ majorizes $v+u$. It follows that $\overline{v}+u$ majorizes $\overline{v+u}$. Then for each $p \in S$, we have $v(p)+u(p) \leq \overline{v+u}(p) \leq \overline{v}(p)+u(p)$. It follows from the first inequality that $v_* \equiv \overline{v+u} - u$ majorizes v, and from the second that $\overline{v}$ majorizes v_*. Moreover, since $\overline{v+u}$ is upper semicontinuous and u is continuous, v_* is upper semicontinuous. Then we must have $\overline{v} = v_*$, since by definition, $\overline{v}$ is the smallest upper semicontinuous function that majorizes v. It follows that $\overline{v}+u = v_*+u = \overline{v+u}$, as desired. □

S.4 Extended-Real Valued Costs

In many settings, the researcher's experimentation technology may not allow her to learn the state of the world with certainty. In particular, as mentioned in the main text, an infinite cost of certainty is a feature of several posterior-separable cost functions discussed in the literature. These include the cost of Wald (1945) sequential sampling (Morris and Strack, 2019), log-likelihood ratio costs (Pomatto et al., 2018), and the *total information* cost function of Bloedel and Zhong (2021). Each is defined by a measure of uncertainty which takes values in the affinely extended real numbers $\overline{\mathbb{R}}$, but is finite on the interior of the set of posterior beliefs. Consequently, when these costs are part of a Bayesian persuasion problem, that problem's value function will also take values in $\overline{\mathbb{R}}$.

In Section S.4.1, I show that the characterization results for general persuasion problems from the main text (Propositions 1, 2 and 4) each hold in the extended real-valued setting. In doing so, I rely on Yoder (2021), which shows that extended real-valued persuasion problems are well behaved. It follows that allowing for an infinite cost of certainty has only a minor effect on my main results: The optimal contract implements (and results-based implementation requires) a choice function which is *feasible*, in the sense that it only produces results which have finite costs.

Formally, suppose we modify the model by letting H be a function from $[0,1]$ to $\mathbb{R} \cup \{-\infty\}$ which is finite on $(0,1)$. As before, H is continuous and strictly concave, and is differentiable on $(0,1)$. Say that an experiment choice function $\chi : \Theta \to X$ is *feasible* if for each θ, H is finite on the support of $\chi(\theta)$. Then we have the following:

Proposition S.4. *Suppose that the cost of certainty may be infinite: H may take the value $-\infty$ at 0 and 1, but is finite on $(0,1)$. Then*

i. (Theorem 1) The principal has an optimal methods-based contract which implements a binary, feasible, Blackwell-monotone experiment choice function χ^*;

ii. (Theorem 2) χ^* exhibits no distortion at the top, garbling everywhere else, and distortion lowers cost to the researcher and value for the principal;

iii. (Theorem 3) Any binary, feasible, Blackwell-monotone experiment choice function can be implemented with a results-based contract, and that results-based contract cannot be outperformed by a methods-based contract; and

iv. (Theorem 4) If χ is not Blackwell-monotone and feasible, it cannot be implemented by a results-based contract.

S.4.1 Characterization Results for Extended-Real Valued Persuasion Problems

Proposition S.5 extends the general additive concavity comparative static from the main text (Proposition 4) to extended real-valued problems. Lemmas S.5 and S.6 do the same for Lemmas S.3 and S.4, respectively. Finally, Propositions S.6 and S.7 use these results to give versions of Propositions 1 and 2 that accommodate an infinite cost of certainty.

For a set S, let $\operatorname{aff}(S)$ denote its affine hull. For a function $v : S \to \overline{\mathbb{R}}$, let $\operatorname{dom}(v) \equiv \{s \in S | v(s) \in \mathbb{R}\}$ denote its effective domain.

Proposition S.5 (Additive Convexity and Monotonicity in Extended Real-Valued Problems)**.** *Suppose that $S \subseteq \mathbb{R}^n$, that $v_0, v_1 : S \to \mathbb{R} \cup \{-\infty\}$ are upper semicontinuous, and that v_0 is additively more concave than v_1. Then for any $p_0 \in \operatorname{ri}(\operatorname{dom} v_0) \cap \operatorname{ri}(\operatorname{dom} v_1)$, and any solutions to the persuasion problems*

$$\tau_0^* \in \arg\max_{\tau \in \Delta(S)} \{E_\tau v_0(p) \text{ s.t. } E_\tau p = p_0\}, \qquad \tau_1^* \in \arg\max_{\tau \in \Delta(S)} \{E_\tau v_1(p) \text{ s.t. } E_\tau p = p_0\},$$

we have $\operatorname{supp} \tau_1^* \cap \operatorname{conv}(\operatorname{supp} \tau_0^*) \subseteq \operatorname{ext}(\operatorname{conv}(\operatorname{supp} \tau_0^*))$.

Proof. The proof proceeds identically to the proof of Proposition 4 in the main text, relying on Lemma 2 in Yoder (2021) instead of Lemma 3 in the main text, except that in order to show equation (15), we must also show that v_Δ is finite on $\operatorname{supp} \tau_0^*$ by inserting the following passage immediately after equation (14):

By Lemma 2 in Yoder (2021), $\operatorname{conv}(\operatorname{supp} \tau_0^*) \subseteq \operatorname{dom} V_0$ and V_0 coincides with v_0 on $\operatorname{supp} \tau_0^*$; hence, $\operatorname{supp} \tau_0^* \subseteq \operatorname{dom} v_0$. Then since $v_0 = v_1 + v_\Delta$ and neither v_1 nor v_Δ map to ∞, it must be that $\operatorname{supp} \tau_0^* \subseteq \operatorname{dom} v_\Delta$. □

Lemma S.5 (Characterization Without Semicontinuity in Extended Real-Valued Binary-State Problems)**.** *Suppose $p_0 \in (0, 1)$, and that $v : [0, 1] \to \mathbb{R} \cup \{-\infty\}$ is finite on $(0, 1)$; let $\bar{v}$ be the upper semicontinuous hull of v, i.e., $\bar{v}(p) \equiv \max\{z | (p, z) \in \operatorname{cl}(\operatorname{hypo}(v))\}$; and let $\bar{V}$ be the concavification of $\bar{v}$. If $\tau^* \in \arg\max_{\tau \in \Delta([0,1])}\{E_\tau v(p) \text{ s.t. } E_\tau p = p_0\}$, then $\tau^* \in \arg\max_{\tau \in \Delta([0,1])}\{E_\tau \bar{v}(p) \text{ s.t. } E_\tau p = p_0\}$ and $E_{\tau^*} v(p) = \bar{V}(p_0)$.*

Proof. Follows identically to the proof of Lemma S.3, replacing $\mathbb{R}$ with $\overline{\mathbb{R}}$ when considering case (ii), relying on Yoder (2021) Proposition 2 instead of Kamenica and Gentzkow (2011) Online Appendix Propositions 3 and 4. □

Lemma S.6. *Let $S \subseteq \mathbb{R}^n$, and for $v, u : S \to \mathbb{R} \cup \{-\infty\}$, let $\overline{v}$ and $\overline{v + u}$ be the upper semicontinuous hulls of v and $v + u$, respectively. If u is continuous, then $\overline{v + u} = \overline{v} + u$.*

Proof. The proof proceeds analogously to that of Lemma S.4. By definition, $\overline{v}$ (resp., $\overline{v + u}$) is the smallest upper semicontinuous function that majorizes v (resp., $v + u$). Since u is continuous, $\overline{v} + u$ is upper semicontinuous; it follows that $\overline{v} + u$ majorizes $\overline{v + u}$. Hence, for each $p \in S$, we have

$$v(p) + u(p) \leq \overline{v + u}(p) \leq \overline{v}(p) + u(p). \tag{S.7}$$

Then consider

$$v_*(x) \equiv \begin{cases} \overline{v+u}(x) - u(x), & x \in \text{dom}(u) \\ \overline{v}(x), & x \notin \text{dom}(u). \end{cases}$$

It follows from the first inequality in (S.7) that v_* majorizes v, since $\overline{v}(x) \geq v(x)$ for each $x \in S$, and from the second inequality in (S.7) that $\overline{v}$ majorizes v_*.

Moreover, v_* is upper semicontinuous: Since u is continuous, $\overline{v+u}(x) - u(x)$ is upper semicontinuous on $\text{dom}(u)$, and $\text{dom}(u) = u^{-1}((-\infty, \infty))$ is open in S. Then for all $y \in \text{dom}(u)$ and $y_n \to y$, $\limsup_{n\to\infty} v_*(y_n) = \limsup_{n\to\infty} \overline{v+u}(y_n) - u(y_n) \leq (\limsup_{n\to\infty} \overline{v+u}(y_n)) + (\limsup_{n\to\infty} -u(y_n)) \leq \overline{v+u}(y) - u(y) = v_*(y)$. And for all $y \notin \text{dom}(u)$ and $y_n \to y$, $\limsup_{n\to\infty} v_*(y_n) \leq \limsup_{n\to\infty} \overline{v}(y_n) \leq \overline{v}(y) = v_*(y)$, since $\overline{v}$ is upper semicontinuous and majorizes v_*. Upper semicontinuity of v_* follows from, e.g., Aliprantis and Border (2013) Lemma 2.42.

Then we must have $\overline{v} = v_*$, since by definition, $\overline{v}$ is the smallest upper semicontinuous function that majorizes v. It follows that $\overline{v} + u = v_* + u = \overline{v+u}$, as desired. □

Proposition S.6 (Additive Convexity and Monotonicity in Extended Real-Valued Problems)**.** *Suppose that $v_0, v_1 : [0,1] \to \mathbb{R} \cup \{-\infty\}$ are finite on $(0,1)$, and that v_0 is additively more concave than v_1. Then for any $p_0 \in (0,1)$, and any solutions to the persuasion problems*

$$\tau_0^* \in \arg \max_{\tau \in \Delta([0,1])} \{E_\tau v_0(p) \text{ s.t. } E_\tau p = p_0\}, \quad \tau_1^* \in \arg \max_{\tau \in \Delta([0,1])} \{E_\tau v_1(p) \text{ s.t. } E_\tau p = p_0\},$$

τ_1^* is weakly Blackwell-more informative than (i.e., a mean-preserving spread of) τ_0^*.

Proof. Follows identically to the proof of Proposition 1 in the main text, relying on Proposition S.5 instead of Proposition 4, and Lemmas S.5 and S.6 instead of Lemmas S.3 and S.4. □

Proposition S.7 (A Necessary and Sufficient Secant Line Condition for Extended Real-Valued Problems)**.** *Suppose that $v : [0,1] \to \mathbb{R} \cup \{-\infty\}$ is an upper semicontinuous function which is finite on $(0,1)$, and that $p_0 \in (0,1)$. An experiment τ^* which produces the pair of results* $\text{supp}\, \tau^* = \{\underline{p}, \overline{p}\}$ *solves the persuasion problem* $\max_\tau\{E_\tau v(p) \text{ s.t. } E_\tau p = p_0\}$ *if and only if the secant line to v through $(\underline{p}, v(\underline{p}))$ and $(\overline{p}, v(\overline{p}))$ lies on or above the graph of v:*

$$v(\underline{p}) + \frac{v(\overline{p}) - v(\underline{p})}{\overline{p} - \underline{p}}(p - \underline{p}) \geq v(p) \text{ for each } p \in [0,1]. \tag{S.8}$$

Proof. Follows identically to the proof of Proposition 2 in the main text, relying on Lemma 2 in Yoder (2021) instead of Lemma 3. □

S.4.2 Proof of Proposition S.4

Observe that by modifying their proofs to rely on Proposition 1 instead of Proposition S.6, Proposition 2 instead of Proposition S.7, and Kamenica and Gentzkow (2011) Online Appendix Propositions 3 and 4 instead of Yoder (2021) Proposition 2, each of the lemmas in Appendices A and B.2 of the main text can be extended to the setting of Proposition S.4. With those lemmas thus extended, (i) and (ii) follow identically to the proofs of Theorems 1 and 2 in the main text. Furthermore, (iv) follows identically to the proof of Theorem 4 in the main text, relying on Proposition S.6 instead of Proposition 1.

Finally, (iii) follows identically to the proof of Theorem 3 in the main text, relying on Proposition S.7 instead of Proposition 2, with three exceptions. First, so that ψ_χ is finite at the endpoints of the unit interval, replace its definition as follows:

$$\psi_\chi(p) \equiv \begin{cases} s_\lambda(p) + \theta_\lambda c(p), & p \in [\underline{y}_\lambda, \overline{y}_\lambda]; \\ s_i(p) + \theta_i c(p), & p \in [\underline{y}_i, \underline{y}_{i-1}) \cup (\overline{y}_{i-1}, \overline{y}_i], \forall i \in \{\lambda+1, \dots, N\}; \\ \min\{s_N(p) + \theta_N c(p), 0\}, & p \in [0, \underline{y}_N) \cup (\overline{y}_N, 1]. \end{cases}$$

Second, to account for the new definition of ψ_χ, replace (26) as follows:

$$\psi_\chi(p) - \theta_i c(p) \leq s_j(p) + (\theta_j - \theta_i)c(p) \text{ (with equality when } j < N).$$

Finally, replace the continuity argument after (22) with the following passages showing that ψ_χ is upper semicontinuous and finite-valued:

ψ_χ is upper semicontinuous: First note that $\{s_i\}_{i=\lambda}^N$ are linear, hence continuous. By definition we have $s_i(\underline{y}_{i-1}) + \theta_i c(\underline{y}_{i-1}) = s_{i-1}(\underline{y}_{i-1}) + \theta_{i-1} c(\underline{y}_{i-1})$ and $s_i(\overline{y}_{i-1}) + \theta_i c(\overline{y}_{i-1}) = s_{i-1}(\overline{y}_{i-1}) + \theta_{i-1} c(\overline{y}_{i-1})$ for each $i \in \{\lambda+1, \dots, N\}$; since H (and thus c) is continuous, it follows that ψ_χ is continuous on $[\underline{y}_N, \overline{y}_N]$. Since $s_N(p) + \theta_N c(p)$ is continuous, so is $\min\{s_N(p) + \theta_N c(p), 0\}$; it follows that ψ_χ is continuous on $[0, \underline{y}_N) \cup (\overline{y}_N, 1]$ as well. Upper semicontinuity then follows from the fact that $\lim_{p \uparrow \underline{y}_N} \psi_\chi(p) = \lim_{p \uparrow \underline{y}_N} \min\{s_N(p) + \theta_N c(p), 0\} = \min\{s_N(\underline{y}_N) + \theta_N c(\underline{y}_N), 0\} \leq s_N(\underline{y}_N) + \theta_N c(\underline{y}_N) = \psi_\chi(\underline{y}_N)$ and $\lim_{p \downarrow \overline{y}_N} \psi_\chi(p) = \lim_{p \downarrow \overline{y}_N} \min\{s_N(p) + \theta_N c(p), 0\} = \min\{s_N(\overline{y}_N) + \theta_N c(\overline{y}_N), 0\} \leq s_N(\overline{y}_N) + \theta_N c(\overline{y}_N) = \psi_\chi(\overline{y}_N)$.

ψ_χ takes finite values, and hence is a results-based contract: Since H is continuously differentiable on $(0,1)$ and $p_0 \in (0,1)$, s_λ takes finite values. Moreover, since χ is feasible, for each $i \in \{\lambda+1, \dots, N\}$, if s_{i-1} takes finite values, so does s_i. It follows by induction that s_i takes finite values for each $i \in \{\lambda, \dots, N\}$. Furthermore, since χ is feasible, we have $\{\underline{y}_N, \overline{y}_N\} \subset \operatorname{dom} H$; since $(0,1) \subseteq \operatorname{dom} H$ by assumption, it must be that $[\underline{y}_N, \overline{y}_N] \subset \operatorname{dom} H$. Then ψ_χ is finite on $[\underline{y}_N, \overline{y}_N]$. Since $c(p) > -\infty$ for each $p \in [0,1]$, $\min\{0, s_N(p) + \theta_N c(p)\}$

is finite for each $p \in [0,1]$; it follows that ψ_χ is finite on $[0, \underline{y}_N) \cup (\overline{y}_N, 1]$ as well.

References

ALIPRANTIS, C. D. AND K. C. BORDER (2013): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer-Verlag Berlin and Heidelberg GmbH & Company KG.

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

JULLIEN, B. (2000): "Participation Constraints in Adverse Selection Models," *Journal of Economic Theory*, 93, 1–47.

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

MACHINA, M. AND J. PRATT (1997): "Increasing Risk: Some Direct Constructions," *Journal of Risk and Uncertainty*, 14, 103–127.

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

POMATTO, L., P. STRACK, AND O. TAMUZ (2018): "The Cost of Information," *arXiv preprint arXiv:1812.04211*.

WALD, A. (1945): "Sequential Tests of Statistical Hypotheses," *The Annals of Mathematical Statistics*, 16, 117–186.

YODER, N. (2021): "Extended Real-Valued Information Design," Working paper.