Addressing Strategic Uncertainty with Incentives and Information

Marina Halac Elliot Lipnowski Daniel Rappoport*

We study the optimal design of incentives and information in multi-agent settings with externalities. A principal privately contracts with a set of agents who then simultaneously choose a binary action. There is a hidden state of nature that we call the fundamental state. The principal offers each agent a contingent individual allocation, and possibly gives agents information about the fundamental state and each other's contracts. Each agent's payoff depends on the profile of agents' actions, his allocation, and the fundamental state. We solve for the principals optimal incentive scheme that maximizes her expected payoff subject to inducing a desired action profile as the unique rationalizable outcome.

Our main result is a simplification of this multi-agent problem to a two-step procedure in which information is designed agent-by-agent: the principal chooses a fundamental-state-contingent distribution over agent rankings, and then, separately for each agent, the agent's information about the fundamental state and realized ranking. We highlight that such a ranking state together with the fundamental state—what we call the total state—is the right state variable for the principal's problem. Similar state variables appear in prior work on unique equilibrium implementation in supermodular games; most closely related, Halac, Lipnowski and Rappoport (2021a) and Morris, Oyama and Takahashi (2021). Our analysis elucidates that the total state captures agents' relevant uncertainty whenever their incentives are pinned down by their relative order in the sequence of deletion of dominated actions.

We illustrate our results by studying a team-effort problem, related to Winter (2004), Moriya and Yamashita (2020), and Halac, Lipnowski and Rappoport (2021a). Our two-step procedure permits an explicit characterization of the principal's solution, and we describe

*December 27, 2021. Halac: Yale University (marina.halac@yale.edu); Lipnowski: Columbia University (e.lipnowski@columbia.edu); Rappoport: University of Chicago (daniel.rappoport@chicagobooth.edu).

how this solution varies with the environment. We find that the principal may want to give agents no information, public information, or private information about the total state.

Our paper joins a growing literature on unique implementation, including work on incentive contracting and recent work on information design. In addition to the papers already cited, see, e.g., Segal (2003), Bernstein and Winter (2012), Chassang, Del Carpio and Kapon (2020), Halac, Kremer and Winter (2020, 2021b), and Camboni and Porcellacchia (2021) on incentive design; Hoshino (2019), Mathevet, Perego and Taneva (2020), Inostroza and Pavan (2021), and Li, Song and Zhao (2021) on information design; and prior work related to the latter strand such as Kajii and Morris (1997). Our main departure from the literature is that we study both of these tools jointly. We provide a general methodology that can be useful for a variety of applications in which strategic uncertainty may be addressed with incentives and information.

1. Model

A principal contracts with a set $N = \{1, \ldots, N\}$ of agents. There is a state of nature, or *fundamental state*, drawn from a finite set Ω according to a probability distribution $p_0 \in \Delta\Omega$ with full support.[1] The principal offers each agent $i \in N$ a private allocation $x_i \in X_i$, and possibly gives agents information about the fundamental state and each other's contracts. Agents then simultaneously choose a binary action, either 1 or 0.

Formally, a principal's incentive scheme is $\sigma = \langle q, \chi \rangle$, where $q \in \Delta\left[(\mathbb{N}^2)^N \times \Omega\right]$ is a prior with marginal distribution p_0 on Ω and $\chi = (\chi_i)_{i \in N}$ is an allocation rule, with $\chi_i : \text{supp}(\text{marg}_i\, q) \to X_i$. Let $T_i^q := \text{supp}(\text{marg}_i\, q)$ denote the support of the marginal of q along dimension i and $T^q := \prod_{i \in N} T_i^q$. The interpretation is that the principal privately informs each agent $i \in N$ of his *type* $t_i \in T_i^q$ and, through the allocation rule, of his contingent allocation from every profile of actions in every fundamental state.[2] Hence, an agent may

[1] Throughout, given a set Z, let ΔZ denote the set of finite-support probability distributions over Z.

[2] Given the finite type restriction, since the type itself is a strategically irrelevant label (Dekel, Fudenberg

face uncertainty about other agents' contracts and about the fundamental state, but is completely informed about his own contract. The choice of q, specifically the correlation between an agent's type and others' types and the fundamental state, determines how much an agent knows about others' contracts and the fundamental state.

An incentive scheme $\sigma = \langle q, \chi \rangle$ defines a Bayesian game between the agents. In this game, $\langle (T_i^q)_{i \in N}, \Omega, q \rangle$ is a common-prior type space; each agent simultaneously makes a type-contingent decision of whether to choose 1 or 0; and an agent i's payoff is a function $u_i : 2^N \times X_i \times \Omega \to \mathbb{R}$ of the set of agents who choose action 1, his allocation, and the fundamental state. The principal wishes to uniquely induce all agents to choose action 1 (with probability 1), with her payoff in such event given by $\sum_{i \in N} v_i(x_i, \omega)$ for $v_i : X_i \times \Omega \to \mathbb{R}$ bounded above. Say that an action is *rationalizable* for an agent type if it is interim correlated rationalizable for said type, and say that an incentive scheme $\sigma = \langle q, \chi \rangle$ is *unique implementation feasible (UIF)* if all agent types choosing 1 is the unique rationalizable outcome of the Bayesian game induced by σ. The principal solves

$$\sup_{\sigma \text{ is UIF}} V(\sigma), \tag{1}$$

where $V(\sigma)$ is her total expected payoff given scheme $\sigma = \langle q, \chi \rangle$ and all agents choosing 1:

$$V(\sigma) = \sum_{t \in T^\sigma,\ \omega \in \Omega} q(t, \omega) \sum_{i \in N} v_i \left(\chi_i(t_i), \omega \right).$$

We make a *dominant-allocation assumption* that says that for each agent $i \in N$, there exists $\overline{x}_i \in X_i$ such that choosing action 1 is dominant:

$$\min_{J \subseteq N \setminus \{i\}, \omega \in \Omega} \left[u_i(J \cup \{i\}, \overline{x}_i, \omega) - u_i(J, \overline{x}_i, \omega) \right] > 0.$$

This assumption implies that it is always possible for the principal to induce all agents choosing 1 as the unique rationalizable outcome. Our focus is on solving for optimal incentive

and Morris, 2007, Proposition 1), it is immaterial that types are labeled with natural number pairs.

schemes that achieve this goal.[3] It will follow from our results that the principal's problem in (1) does not generally admit a maximum, but en route to our characterization of her optimal value, we will construct approximately optimal incentive schemes.[4]

Remark 1. A special case of our model is the case of a supermodular game, in which $J \mapsto u_i(J \cup \{i\}, x_i, \omega) - u_i(J, x_i, \omega)$ is a weakly increasing map on $2^{N \setminus \{i\}}$ for every $i \in N$, $x_i \in X_i$, and $\omega \in \Omega$. In this case, the requirement that each type choosing action 1 be uniquely rationalizable is equivalent to the requirement that it be a unique Bayes-Nash equilibrium.

Remark 2. We have assumed that the set of feasible profiles of allocations is a product set $\prod_{i \in N} X_i$, and that the principal's objective (conditional on all agents choosing 1) takes an additively separable form. The tools we develop can be useful even without these separability conditions, provided that we appropriately generalize our dominant-allocation assumption.[5]

2. Solving for Optimal Schemes

We will find it convenient to express properties of an incentive scheme in terms of the order its type realizations induce on agents. Denote by Π the set of all permutations on N (i.e., all $\pi \in N^N$ with $\pi_i \neq \pi_j$ for all distinct $i, j \in N$), and consider incentive schemes $\sigma = \langle q, \chi \rangle$ such that every positive-probability type profile $t = (t_i^R, t_i^S)_{i \in N} \in T^q$ has $t_i^R \neq t_j^R$ for all distinct $i, j \in N$. Any such type profile t induces a *ranking state* $\pi(t) \in \Pi$ given by $\pi_i(t) = |\{j \in N : \ t_j^R \leq t_i^R\}|$. A key consequence of our analysis will be that the relevant state variable for the principal's problem consists of the ranking state $\pi \in \Pi$ together with the fundamental state $\omega \in \Omega$. We will refer to (π, ω) as the *total state*.

[3] Our work is thus complementary to the information design results of Morris, Oyama and Takahashi (2021), which focus on implementability. By combining our analysis with theirs, it may be possible to weaken our dominant-allocation assumption.

[4] That is, for any $\varepsilon > 0$, our proof constructs a UIF scheme σ_ε such that $V(\sigma_\varepsilon) > \sup_{\sigma \text{ is UIF}} V(\sigma) - \varepsilon$.

[5] Specifically, our analysis implies that the principal's program can still be reduced to a two-step procedure: first, choose a fundamental-state-contingent distribution over what we will call ranking states and assign an optimal principal value to any profile of agent beliefs about the fundamental and ranking states; second, design an information structure concerning the realized fundamental and ranking states. If the principal's value is independent of the fundamental state, the analysis of Morris (2020) (and the classic work cited therein), Ziegler (2020), or Arieli, Babichenko, Sandomirskiy and Tamuz (2021) can be applied.

Given a prior q, agent $i \in N$, and type $t_i \in T_i^q$, we have that t_i's belief $\mu_i^q(\cdot|t_i) \in \Delta(\Pi \times \Omega)$ about the total state is given by

$$\mu_i^q(\hat{\pi}, \hat{\omega}|t_i) := q_i\left(\{t_{-i}:\ \pi(t_i, t_{-i}) = \hat{\pi}\} \times \{\hat{\omega}\} \;\middle|\; t_i\right) \text{ for all } \hat{\pi} \in \Pi, \hat{\omega} \in \Omega,$$

where $q_i : T_i^q \to \Delta(T_{-i}^q \times \Omega)$ is given by $q_i(t_{-i}, \omega|t_i) := \frac{1}{\text{marg}_i\, q(t_i)} q(t_i, t_{-i}, \omega)$. The total state distribution $\mu^q \in \Delta(\Pi \times \Omega)$ is given by

$$\mu^q(\hat{\pi}, \hat{\omega}) := q\left(\{t:\ \pi(t) = \hat{\pi}\} \times \{\hat{\omega}\}\right) \text{ for all } \hat{\pi} \in \Pi, \hat{\omega} \in \Omega.$$

For any agent $i \in N$ and belief $\mu_i \in \Delta(\Pi \times \Omega)$ that he might hold, let us define his sufficient allocations $x_i \in X_i$ as those that induce the agent to choose action 1 under the hypothesis that all agents $j \in N \setminus \{i\}$ with rank $\pi_j < \pi_i$ choose action 1. Letting

$$I_i(x_i, \pi, \omega) := \min_{J \subseteq N \setminus \{i\}:\ J \supseteq \{j \in N:\ \pi_j < \pi_i\}} \left[u_i(J \cup \{i\}, x_i, \omega) - u_i(J, x_i, \omega)\right],$$

the agent's set of sufficient allocations is given by

$$\mathcal{X}_i^*(\mu_i) := \left\{x_i \in X_i : \sum_{\pi \in \Pi, \omega \in \Omega} \mu_i(\pi, \omega)\ I_i(x_i, \pi, \omega) > 0\right\}.$$

By our dominant-allocation assumption, this set is nonempty as it contains allocation $\overline{x}_i$.

Definition 1. *A* strict ranking scheme *is an incentive scheme* $\sigma = \langle q, \chi \rangle$ *such that:*

1. *Every positive-probability* $t \in T^q$ *has* $t_i^R \neq t_j^R$ *for all distinct* $i, j \in N$.

2. *Every* $i \in N$ *and* $t_i \in T_i^q$ *have* $\chi_i(t_i) \in \mathcal{X}_i^*(\mu_i^q(\cdot|t_i))$.

The next lemma shows that strict ranking schemes are useful because they ensure choosing 1 is uniquely rationalizable and, up to relabeling of types, constitute all such incentive schemes. See the Online Appendix for proofs of all of our results.

Lemma 1. *Every strict ranking scheme is UIF. Moreover, if an incentive scheme σ is UIF, there exists a strict ranking scheme σ^* with $V(\sigma^*) = V(\sigma)$.*

The proof is constructive: we relabel types so that the order in which agents have action 0 eliminated in an iterated deletion sequence exactly corresponds to the ranking state $\pi \in \Pi$.

Lemma 1 implies that to solve the principal's problem in (1), it is without loss to focus on strict ranking schemes. For any agent $i \in N$ and belief $\mu_i \in \Delta(\Pi \times \Omega)$ that he might hold, define the principal's interim value function by

$$v_i^*(\mu_i) := \sup_{x_i \in \mathcal{X}_i^*(\mu_i)} \sum_{\pi \in \Pi, \omega \in \Omega} \mu_i(\pi, \omega)\ v_i(x_i, \omega).$$

The principal's problem is then to choose a prior in order to maximize the expectation of $\sum_{i \in N} v_i^*(\mu_i)$. Our main result is a simplification of this problem to a two-step procedure in which information is designed agent-by-agent: first, the principal chooses a total state distribution $\mu \in \Delta(\Pi \times \Omega)$; second, separately for each agent, she chooses what information to provide to the agent about the realized total state (π, ω). Formally, for any agent $i \in N$ and distribution $\mu \in \Delta(\Pi \times \Omega)$, define

$$\widehat{v}_i^*(\mu) := \sup_{\tau_i \in \Delta\Delta(\Pi \times \Omega)} \int v_i^*(\mu_i)\ \mathrm{d}\tau_i(\mu_i) \ \text{ subject to } \int \mu_i\ \mathrm{d}\tau_i(\mu_i) = \mu, \tag{2}$$

which is the pointwise-lowest concave function above v_i^*. Denote the set of allowable total state distributions by $\mathcal{M}(p_0) = \{\mu \in \Delta(\Pi \times \Omega) : \ \mathrm{marg}_\Omega \mu = p_0\}$. We obtain:

Theorem 1. *The principal's optimal value satisfies*

$$\sup_{\sigma \text{ is UIF}} V(\sigma) = \sup_{\mu \in \mathcal{M}(p_0)} \sum_{i \in N} \widehat{v}_i^*(\mu).$$

The reduction in Theorem 1 is significant. Instead of optimizing over fundamental-state-contingent distributions over type profiles $q \in \Delta\left[(\mathbb{N}^2)^N \times \Omega\right]$, the principal simply chooses a fundamental-state-contingent distribution over rankings $\mu \in \Delta(\Pi \times \Omega)$. Then, agent-by-agent, the principal solves the single-agent information design problem in (2)—a well understood problem given the extensive literature on persuasion (see Kamenica, 2019).

The proof of Theorem 1 establishes that $\sup_{\sigma \text{ is UIF}} V(\sigma) \leq \sup_{\mu \in \mathcal{M}(p_0)} \sum_{i \in N} \widehat{v}_i^*(\mu)$ by using Lemma 1 and program (2), and shows that this inequality holds with equality by constructing a sequence of strict ranking schemes that approximates the payoff bound. The construction is the same as that in Halac, Lipnowski and Rappoport (2021a), but with types augmented to convey information about the total state.[6]

We close this section by noting that an optimum exists in many natural cases:

Definition 2. Say $(\mu, (\tau_i)_{i \in N})$ is *optimal* if $\mu \in \arg\max_{\tilde{\mu} \in \mathcal{M}(p_0)} \sum_{i \in N} \widehat{v}_i^*(\tilde{\mu})$ and τ_i is an optimum of program (2) defining $\widehat{v}_i^*(\mu)$ for every $i \in N$.

Fact 1. *If v_i^* is upper semicontinuous for each $i \in N$, some optimal $(\mu, (\tau_i)_{i \in N})$ exists.*

3. Team effort with transfers

We illustrate our results by studying a simple team-effort problem. Our two-step procedure permits an explicit characterization of the principal's solution, and we describe how this solution varies with the environment. In particular, we show that the principal may want to give agents no information, public information, or private information about the total state.

Consider a special case of our model in which a set $N = \{1, 2\}$ of agents privately choose whether to work (choose 1) or shirk (choose 0) on a joint project. The fundamental state ω is drawn uniformly from $\Omega = \{1, 2\}$ and determines agents' costs of effort, given by $c_i(\omega) > 0$ for $i \in N$. The project succeeds with probability P_k if k agents work and the rest shirk, and the allocation $x_i \in X_i = \mathbb{R}_+$ is a bonus that the principal pays agent i in the case of success. We thus write agent i's payoff as $u_i(J, x_i, \omega) = P_{|J|} x_i - c_i(\omega) \mathbf{1}_{i \in J}$. The principal's goal is to uniquely induce the agents to work at the least possible incentive cost, so $v_i(x_i, \omega) = -x_i$.

[6] In that paper's setting, this augmentation was not needed as providing no information was optimal.

We assume P is strictly increasing (i.e., $1 \geq P_2 > P_1 > P_0 \geq 0$) and strictly supermodular (i.e., $P_2 - P_1 > P_1 - P_0$), meaning that agents' efforts are productive and complementary. Since an agent's incentive to work is then always increasing in the other agent's effort, the agent's set of sufficient allocations takes a simple form. Specifically, denote by $\mu_i^\Pi \in \Delta\Pi$ and $\mu_i^\Omega \in \Delta\Omega$ the marginals of μ_i along Π and Ω respectively, and let $\pi^i \in \Pi$ be the ranking state in which agent i is ranked second. Defining the expected marginal product

$$\iota_i(\mu_i^\Pi) := \left[1 - \mu_i^\Pi(\pi^i)\right](P_1 - P_0) + \mu_i^\Pi(\pi^i)(P_2 - P_1),$$

and given that the agent's expected cost of effort is $c_i(\mu_i^\Omega) := \sum_{\omega \in \Omega} \mu_i^\Omega(\omega) c_i(\omega)$, direct computation yields $\mathcal{X}_i^*(\mu_i) = \left\{x_i \in X_i : x_i \iota_i(\mu_i^\Pi) > c_i(\mu_i^\Omega)\right\}$. Hence, $v_i^*(\mu_i) = -c_i(\mu_i^\Omega)/\iota_i(\mu_i^\Pi)$, and replacing the objective with its negative, the principal's problem can be written as

$$\inf_{\substack{\mu \in \mathcal{M}(p_0),\\ \tau_1, \tau_2 \in \Delta\Delta(\Pi \times \Omega)}} \sum_{i \in N} \int \frac{c_i(\mu_i^\Omega)}{\iota_i(\mu_i^\Pi)} \, \mathrm{d}\tau_i(\mu_i) \quad \text{subject to} \quad \int \mu_1 \, \mathrm{d}\tau_1(\mu_1) = \int \mu_2 \, \mathrm{d}\tau_2(\mu_2) = \mu. \tag{3}$$

We next present different examples that vary in how agents' effort costs depend on the fundamental state. We denote by $\tau_i^\Pi \in \Delta\Delta\Pi$ and $\tau_i^\Omega \in \Delta\Delta\Omega$ the distributions of the marginals of μ_i along Π and Ω respectively, and let $\varphi := (P_1 - P_0)/(P_2 - P_1) \in (0, 1)$.

An example with no information. Suppose agents' effort costs are constant.

Proposition 1. *Take $c_1(1) = c_1(2) =: c_H \geq c_L := c_2(1) = c_2(2)$. Then a feasible (μ, τ_1, τ_2) is optimal if and only if $\tau_1^\Pi\left(\mu^\Pi\right) = \tau_2^\Pi\left(\mu^\Pi\right) = 1$ and*

$$\mu^\Pi(\pi^1) = \begin{cases} \dfrac{\sqrt{c_H} - \varphi\sqrt{c_L}}{(1-\varphi)(\sqrt{c_H} + \sqrt{c_L})} & : \ \varphi\sqrt{c_H} < \sqrt{c_L} \\ 1 & : \ \textit{otherwise.} \end{cases}$$

In particular, in every optimum, neither agent learns anything about the ranking state; and some optimum exists in which neither agent learns anything about the fundamental state.

This result corresponds to a special case of the results in Halac, Lipnowski and Rappoport (2021a). When c_i is constant for each $i \in N$, the interim value functions v_i^* are all concave, so $\widehat{v}_i^* = v_i^*$ and providing no information to the agents about the realized ranking state is strictly optimal. Because the fundamental state is irrelevant, the principal is indifferent to providing information about it, as long as agents learn nothing about the ranking state. Our two-step procedure therefore reduces to a single optimization over $\mu \in \mathcal{M}(p_0)$ in this setting. The solution in Proposition 1 shows that the higher is agent 1's effort cost relative to agent 2's, the higher is the probability μ places on ranking state π^1 that ranks agent 1 second.

An example with public information. Suppose agents are ex-ante identical but their effort costs are perfectly negatively correlated: one has a high cost and the other a low cost, depending on the fundamental state.

Proposition 2. *Take $c_1(1) = c_2(2) =: c_H > c_L := c_2(1) = c_1(2)$. Then there is a unique optimal (μ, τ_1, τ_2). Each $i \in N$ has $\tau_i(\beta_1^* \otimes \delta_1) = \tau_i(\beta_2^* \otimes \delta_2) = 1/2$, where*

$$\beta_1^*(\pi^1) = \beta_2^*(\pi^2) = \begin{cases} \dfrac{\sqrt{c_H} - \varphi\sqrt{c_L}}{(1-\varphi)(\sqrt{c_H} + \sqrt{c_L})} & : \ \varphi\sqrt{c_H} < \sqrt{c_L} \\ 1 & : \ \textit{otherwise.} \end{cases}$$

In particular, in the unique optimum, agents learn the fundamental state and hold identical beliefs about the total state.

The proposition shows that providing public information is strictly optimal in this setting. In the unique optimum, each agent learns the fundamental state, and in turn learns something about the ranking state. Moreover, because each agent holds a unique belief in each fundamental state, it follows that agents perfectly learn each other's beliefs—that is, information must be public. The intuition is that the principal benefits from correlating agents' ranking state beliefs with their relative effort costs and thus, here, with the fundamental state. In fact, observe that ranking state beliefs are the same function of effort costs as in Proposition 1: because the fundamental state is publicly revealed, it is as if the principal optimizes the contracts separately over two deterministic environments.

An example with private information. Suppose the effort cost of only agent 1 varies with the fundamental state, and for simplicity let agent 2's constant effort cost be equal to agent 1's in one of the fundamental states.

Proposition 3. *Take $c_1(1) =: c_H > c_L := c_2(1) = c_2(2) = c_1(2)$. Then a feasible (μ, τ_1, τ_2) is optimal if and only if $\tau_1(\beta^{**}_\omega \otimes \delta_\omega) = 1/2$ for each $\omega \in \Omega$ and $\tau_2^\Pi\left(\int \beta^{**}_\omega \; dp_0(\omega)\right) = 1$, where*

$$\left(\beta_1^{**}(\pi^1), \beta_2^{**}(\pi^1)\right) = \begin{cases} \left(\frac{(2+\varphi)\sqrt{c_H} - 3\varphi\sqrt{c_L}}{(1-\varphi)(3\sqrt{c_L}+\sqrt{c_H})}, \frac{(2-\varphi)\sqrt{c_L} - \varphi\sqrt{c_H}}{(1-\varphi)(3\sqrt{c_L}+\sqrt{c_H})}\right) & : \frac{\sqrt{c_H}}{\sqrt{c_L}} \leq \frac{3}{1+2\varphi} \\ (1, 1/3) & : \text{otherwise.} \end{cases}$$

In particular, in every optimum, agent 1 has strictly more information than agent 2 about both the ranking state and the fundamental state.

The proposition shows that providing private information is strictly optimal in this setting. Agent 1 (whose effort cost varies with the fundamental state) learns the fundamental state and in turn something about the realized ranking state. In contrast, agent 2 (whose effort cost is constant) is given no information about the ranking state, and therefore is given strictly less information about the fundamental state than agent 1.

More examples. In the examples above, the principal optimally gives an agent either no information or full information about the fundamental state. We can show however that this is not a general property. For example, suppose agents' effort costs are perfectly positively correlated: $c_1(1) = c_2(1) =: c_H > c_L := c_2(2) = c_1(2)$. Then because the principal would want to correlate each agent's ranking state belief with the fundamental state in the same direction, it turns out that giving an agent partial information about the fundamental state is strictly optimal.

A natural extension of our team-effort problem is to let the probability of project success depend on the fundamental state. For $\omega \in \{1, 2\}$, suppose the project succeeds with probability $P_k(\omega)$ if k agents work and the rest shirk. By similar logic as for Proposition 1, we can show that if P_2 and $(c_i)_{i \in N}$ are constant, then it is optimal to give agents no information about the realized ranking and fundamental states. More generally, providing public or private information may be optimal when both project success and effort costs depend on the fundamental state, and our methodology can be used to solve for the optimal joint design of incentives and information.

References

Arieli, Itai, Yakov Babichenko, Fedor Sandomirskiy, and Omer Tamuz, "Feasible joint posterior beliefs," *Journal of Political Economy*, 2021, *129* (9), 2546–2594.

Bernstein, Shai and Eyal Winter, "Contracting with Heterogeneous Externalities," *American Economic Journal: Microeconomics*, 2012, *4*, 50–76.

Camboni, Matteo and Michael Porcellacchia, "Monitoring Team Members: Information Waste and the Self-Promotion Trap," 2021.

Chassang, Sylvain, Lucia Del Carpio, and Samuel Kapon, "Making the Most of Limited Government Capacity: Theory and Experiment," 2020.

Dekel, Eddie, Drew Fudenberg, and Stephen Morris, "Interim correlated rationalizability," *Theoretical Economics*, 2007, *2* (1), 15–40.

Halac, Marina, Elliot Lipnowski, and Daniel Rappoport, "Rank Uncertainty in Organizations," *American Economic Review*, 2021, *111* (3), 757–86.

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

_ , _ , and _ , “Monitoring Teams,” 2021.

Hoshino, Tetsuya, “Multi-Agent Persuasion: Leveraging Strategic Uncertainty,” 2019.

Inostroza, Nicolas and Alessandro Pavan, “Persuasion in Global Games with Application to Stress Testing,” 2021.

Kajii, Atsushi and Stephen Morris, “The Robustness to Equilibria to In- complete Information,” *Econometrica*, 1997, *65* (6), 1283–1309.

Kamenica, Emir, “Bayesian persuasion and information design,” *Annual Review of Economics*, 2019, *11*, 249–272.

Li, Fei, Yangbo Song, and Mofei Zhao, “Global Manipulation by Local Obfuscation,” 2021.

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

Moriya, Fumitoshi and Takuro Yamashita, “AsymmetricInformation Allocation to Avoid Coordination Failure,” *Journal of Economics & Management Strategy*, 2020, *29* (1), 173–186.

Morris, Stephen, “No Trade and Feasible Joint Posterior Beliefs,” 2020.

_ , Daisuke Oyama, and Satoru Takahashi, “Adversarial Information Design in Binary-Action Supermodular Games,” 2021.

Segal, Ilya R., “Coordination and Discrimination in Contracting with Externalities: Divide and Conquer?,” *Journal of Economic Theory*, 2003, *113*, 147–81.

Winter, Eyal, “Incentives and Discrimination,” *American Economic Review*, 2004, *94*, 764–773.

Ziegler, Gabriel, “Adversarial Bilateral Information Design,” 2020.

Online Appendix for

"Addressing Strategic Uncertainty with Incentives and Information"

by Marina Halac, Elliot Lipnowski, Daniel Rappoport

A. Proofs for Section 2

Proof of Lemma 1. Let us first recall our rationalizability notion. Given an incentive scheme $\sigma = \langle q, \chi \rangle$ we define the sets $\{T_i^\sigma(\kappa)\}_{i \in N, \kappa \in \mathbb{Z}_+}$ as follows. Let $T_i^\sigma(0) := \emptyset$, and then, recursively for $\kappa \in \mathbb{N}$, let $T_i^\sigma(\kappa)$ be the set of all $t_i \in T_i^q$ such that every $\eta \in \Delta\left(2^{N \setminus \{i\}} \times T_{-i}^q \times \Omega\right)$ with $\operatorname{marg}_{T_{-i}^q \times \Omega} \eta = q_i(\cdot|t_i)$ and $\{j \in N \setminus \{i\} : \ t_j \in T_i^\sigma(\kappa - 1)\} \subseteq J$, $\forall (J, t_{-i}, \omega) \in \operatorname{supp}(\eta)$ has

$$\sum_{J \subseteq N \setminus \{i\}, \ t_{-i} \in T_{-i}^q, \ \omega \in \Omega} \eta(J, t_{-i}, \omega) \left[u_i\left(J \cup \{i\}, \chi_i(t_i), \omega\right) - u_i\left(J, \chi_i(t_i), \omega\right)\right] > 0.$$

By definition of interim correlated rationalizability (Dekel et al., 2007), incentive scheme σ is UIF if and only if $\bigcup_{\kappa=0}^\infty T_i^\sigma(\kappa) = T_i^q$ for every $i \in N$.

Now, in what follows, say a type profile t has no ties if $t_i^R \neq t_j^R$ for all distinct $i, j \in N$.

To prove the first assertion, suppose $\sigma = \langle q, \chi \rangle$ is a strict ranking scheme. Let us prove by induction on $\kappa \in \mathbb{Z}_+$ that, if $i \in N$ and $t_i \in T_i^q$ have $t_i^R = \kappa$, then $t_i \in T_i^\sigma(\kappa)$—from which it will follow directly that σ is UIF. The claim holds vacuously for $\kappa = 0$, so take $\kappa \in \mathbb{N}$ and $i \in N$, and assume the claim holds for all $i' \in N$ and all $\kappa' \in \{0, \ldots, \kappa - 1\}$. Next observe that $\chi_i(t_i) \in \mathcal{X}_i^*(\mu_i^q(\kappa))$ because σ is a strict ranking scheme; and the inductive hypothesis implies $t_{-i} \in T_{-i}^\sigma(\kappa - 1)$ for every $t_{-i} \in T_{-i}^q$ such that (t_i, t_{-i}) has no ties and $\pi_j(t) < \pi_i(t)$. Hence, by definition, $t_i \in T_i^\sigma(\kappa)$ as desired.

To prove the second assertion, suppose $\sigma = \langle q, \chi \rangle$ is an arbitrary UIF incentive scheme.

For each $i \in N$, define the map $k_i^\sigma : T_i^q \to \mathbb{N}$ by letting $k_i^\sigma(t_i) := \min\{\kappa \in \mathbb{N} : \ t_i \in T_i^\sigma(\kappa)\}$. It is easy to see some one-to-one function $\tilde{\lambda} : \bigcup_{i\in N} [\{i\} \times T_i^q] \to \mathbb{N}$ exists such that, for any $i, j \in N$ and $t_i \in T_i^q$, $t_j \in T_j^\sigma$ with $k_i^\sigma(t_i) > k_j^\sigma(t_j)$, we have $\tilde{\lambda}_i(t_i) > \tilde{\lambda}_j(t_j)$. Then, define $\lambda : \bigcup_{i\in N} [\{i\} \times T_i^q] \to \mathbb{N}^2$ by letting $\lambda_i(t_i) := (\tilde{\lambda}_i(t_i), 1)$.

Now, define the incentive scheme $\sigma^* := \langle q^*, \chi^* \rangle$ by letting

$$q^*(t^*, \omega) := q\left(\left(\lambda_i^{-1}(t_i^*)\right)_{i\in N},\ \omega\right)$$

for every $t^* \in (\mathbb{N}^2)^N$ and $\omega \in \Omega$, and letting $\chi_i^*(t_i^*) := \chi_i\left(\lambda_i^{-1}(t_i^*)\right)$ for every $i \in N$ and $t_i^* \in T_i^{q^*}$. That the modified scheme is UIF follows from the original scheme being UIF (Dekel et al., 2007, by Proposition 1) given that type t_i's hierarchy of beliefs over $X \times \Omega$ under σ are the same as type $\lambda_i(t_i)$'s under σ^*. Further, because σ^* generates the same distribution over $X \times \Omega$ as σ does, it follows directly that $V(\sigma^*) = V(\sigma)$. All that remains is to see σ^* is a strict ranking scheme. That q^* exhibits no ties is immediate from the construction. Moreover, given any $i, j \in N$, observe any $t_i^* \in T_i^{q^*}$ and $t_j^* \in T_j^{q^*}$ have $k_i^{\sigma^*}(t_i^*) > k_j^{\sigma^*}(t_j^*)$ if and only if $k_i^\sigma(\lambda_i^{-1}(t_i^*)) > k_j^\sigma(\lambda_j^{-1}(t_j^*))$, which in turn implies $t_i^{R*} > t_j^{R*}$. It therefore follows from $t_i^* \in T_i^{\sigma^*}\left(k_i^{\sigma^*}(t_i^*)\right)$ that $\chi_i^*(t_i^*) \in \mathcal{X}_i^*(\mu_i^{q^*}(t_i^*))$, and so σ^* is a strict ranking scheme. *Q.E.D.*

Proof of Theorem 1. We first show that $\sup_{\sigma \text{ is UIF}} V(\sigma) \leq \sup_{\mu\in\mathcal{M}(p_0)} \sum_{i\in N} \widehat{v}_i^*(\mu)$. Given Lemma 1, it suffices to show that the principal's value for a strict ranking scheme $\langle q, \chi \rangle$ is no greater than $\sup_{\mu\in\mathcal{M}(p_0)} \sum_{i\in N} \widehat{v}_i^*(\mu)$. Bayesian updating implies that a given agent i's belief is, on average, equal to the true distribution over total states:

$$\sum_{t\in T^q} q(t)\mu_i^q(\cdot|t_i) = \sum_{t_i\in T_i^q} q_i(t_i)\mu_i^q(\cdot|t_i) = \mu^q \in \mathcal{M}(p_0).$$

Hence, the belief distribution $\tau_i \in \Delta\Delta(\Pi \times \Omega)$ given by $\sum_{t_i\in T_i^q} q_i(t_i)\delta_{\mu_i^q(\cdot|t_i)}$ is feasible in the program defining $\widehat{v}_i^*(\mu^q)$. It follows that

$$\sum_{t \in T^q,\ \omega \in \Omega} q(t,\omega) v_i^*(\mu_i^q(\cdot|t_i)) \leq \widehat{v}_i^*(\mu^q),$$

and so summing over $i \in N$ yields $V(q) \leq \sum_{i \in N} \widehat{v}_i^*(\mu^q)$.

To show $\sup_{\sigma \text{ is UIF}} V(\sigma) \geq \sup_{\mu \in \mathcal{M}(p_0)} \sum_{i \in N} \widehat{v}_i^*(\mu)$, consider an arbitrary $\mu \in \mathcal{M}(p_0)$ and $\varepsilon > 0$. We will construct a strict ranking scheme $\sigma = \langle q, \chi \rangle$ such that $V(\sigma) \geq \sum_{i \in N} [\widehat{v}_i^*(\mu) - 3\varepsilon]$. To do so, observe $\widehat{v}_i^*$ is bounded above by some constant $L_i \in \mathbb{R}$ for each $i \in N$ because v_i^* is. In what follows, let $m \in \mathbb{N}$ be large enough that $m \geq |N|$ and $\frac{2|N|}{m}[L_i - v_i^*(\overline{x}_i, \omega)] \leq \varepsilon$ for each $i \in N$ and $\omega \in \Omega$.

Consider any $i \in N$. Some $\tau_i \in \Delta\Delta(\Pi \times \Omega)$ exists such that $\int \mu_i \, \mathrm{d}\tau_i(\mu_i) = \mu$ and $\int v_i^* \, \mathrm{d}\tau_i \geq \widehat{v}_i^*(\mu) - \varepsilon$. For each $\mu_i \in \text{supp}(\tau_i)$, the definition of v_i^* implies some $x_i^{\mu_i} \in \mathcal{X}_i^*(\mu_i)$ exists such that $\sum_{\pi \in \Pi, \omega \in \Omega} \mu_i(\pi, \omega)\ v_i(x_i^{\mu_i}, \omega) \geq v^*(\mu_i) - \varepsilon$. By the splitting lemma, some $\gamma_i : \Pi \times \Omega \to \Delta\mathbb{N}$ exists such that, when the prior distribution over $\Pi \times \Omega$ is μ and the results of Blackwell experiment γ_i are observed, the induced distribution of beliefs over $\Pi \times \Omega$ is τ_i. Letting $\overline{s}_i \in \mathbb{N}$ denote the number of positive-probability signals in $\mathbb{N}$ given prior μ and experiment γ_i, we can assume without loss that the positive-probability signals are exactly $\{1, \ldots, \overline{s}_i\}$. For each $s_i \in \{1, \ldots, \overline{s}_i\}$, let $x_i^{s_i}$ denote $x_i^{\mu_i}$, where μ_i is the belief induced by signal realization s_i from this experiment.

Now, we construct our incentive scheme $\sigma = \langle q, \chi \rangle$. Define the prior $q \in \Delta[(\mathbb{N}^2)^N \times \Omega]$ by letting, for each $t = (t_i^R, t_i^S)_{i \in N} \in (\mathbb{N}^2)^N$ and $\omega \in \Omega$,

$$q(t,\omega) := \begin{cases} \frac{1}{m}\mu(\pi,\omega) \prod_{i \in N} \gamma_i(t_i^S|\pi,\omega) & : \ \exists \ell \in \{0, \ldots, m-1\} \text{ with } t_i^R = \ell + \pi_i \text{ for all } i \in N, \\ 0 & : \ \text{otherwise}; \end{cases}$$

and the allocation rule $\chi = (\chi_i)_{i\in N}$ via

$$\chi_i(t_i^R, t_i^S) := \begin{cases} x_i^{t_i^S} & : \ t_i^S \leq \overline{s}_i \text{ and } N \leq t_i^R \leq m, \\ \overline{x}_i & : \ \text{otherwise.} \end{cases}$$

By construction, this scheme has no ties: $t_i^R \neq t_j^R$ for all distinct $i, j \in N$ and any supported type profile $t \in T^q$. Moreover, for each $i \in N$, a direct computation shows every type $t_i \in T_i^q$ with $|N| \leq t_i^R \leq m$ has belief $\mu_i^q(\cdot|t_i) = \mu_i^{t_i^S}$ and thus has $\chi_i(t_i) = x_i^{t_i^S} \in \mathcal{X}_i^*(\mu_i^q(\cdot|t_i))$. Because every other $t_i \in T_i^q$ has $\chi_i(t_i) = \overline{x}_i \in \bigcap_{\mu_i \in \Delta(\Pi\times\Omega)} \mathcal{X}_i^*(\mu_i)$, it follows that σ is a strict ranking scheme. Finally, let us bound (from below) the value of this scheme to the principal. To do so, consider any agent $i \in N$ and $s_i \in \{1, \ldots, \overline{s}_i\}$, and observe that σ generates belief $\mu_i^{s_i} \in \Delta(\Pi \times \Omega)$ for agent i with probability

$$\begin{aligned} \text{marg}_i q \left\{ t_i = (t_i^R, t_i^S) \in T_i^q : \ \mu_i^q(\cdot|t_i) = \mu_i^{t_i^S} \right\} \ &\geq \ \sum_{\pi\in\Pi, \ \omega\in\Omega} \sum_{\ell=0}^{m-1} \tfrac{1}{m} \mathbf{1}_{|N|\leq \ell+\pi_i \leq m} \mu(\pi,\omega)\gamma_i(s_i|\pi,\omega) \\ &\geq \ \left(1 - \tfrac{2|N|}{m}\right) \sum_{\pi\in\Pi, \ \omega\in\Omega} \mu(\pi,\omega)\gamma_i(s_i|\pi,\omega) \\ &\geq \ \left(1 - \tfrac{2|N|}{m}\right) \tau_i(\mu_i^{s_i}). \end{aligned}$$

Hence, the principal's payoff from this strict ranking scheme is

$$\begin{aligned} V(\sigma) \ &\geq \ \sum_{i\in N} \left\{ \tfrac{2|N|}{m} \left[\min_{\omega\in\Omega} v_i^*(\overline{x}_i) \right] + \left(1 - \tfrac{2|N|}{m}\right) \sum_{s_i=1}^{\overline{s}_i} \tau_i(\mu_i^{s_i}) \sum_{\omega\in\Omega} \text{marg}_\Omega \mu_i^{s_i}(\omega) \ v_i(x_i^{s_i}, \omega) \right\} \\ &\geq \ \sum_{i\in N} \left\{ \tfrac{2|N|}{m} L_i - \varepsilon + \left(1 - \tfrac{2|N|}{m}\right) \sum_{s_i=0}^{\overline{s}_i - 1} \tau_i(\mu_i^{s_i}) \left[v_i^*(\mu_i^{s_i}) - \varepsilon \right] \right\} \\ &\geq \ \sum_{i\in N} \left\{ \tfrac{2|N|}{m} \widehat{v}_i^*(\mu) - \varepsilon + \left(1 - \tfrac{2|N|}{m}\right) \left[\widehat{v}_i^*(\mu) - 2\varepsilon \right] \right\} \\ &\geq \ \sum_{i\in N} \left[\widehat{v}_i^*(\mu) - 3\varepsilon \right], \end{aligned}$$

as required. *Q.E.D.*

Proof of Fact 1. Let $\mathcal{P}$ denote the set of Borel probability measures on $\Delta(\Pi \times \Omega)$, a compact space when endowed with its weak* topology.

Take any $i \in N$. Because an upper semicontinuous function over a compact space attains a maximum, for any $\mu \in \Delta(\Pi \times \Omega)$, the program $\sup_{\tau_i \in \mathcal{P}:\ \int \mu_i \, \mathrm{d}\tau_i(\mu_i) = \mu} \int v_i^* \, \mathrm{d}\tau_i$—which relaxes the program defining $\widehat{v}_i^*(\mu)$ by allowing distributions with infinite support—admits an optimum. Moreover, by the upper semicontinuous version of Berge's theorem, this optimal value is an upper semicontinuous function of μ. Now, Carathéodory's theorem tells us some optimum to the aforementioned program has affinely independent (hence, of cardinality no more than $N! * |\Omega|$) support. It follows that the program defining $\widehat{v}_i^*(\mu)$ admits an optimum, and that $\widehat{v}_i^*$ is upper semicontinuous.

Finally, because $\sum_{i \in N} \widehat{v}_i^*$ is upper semicontinuous and $\mathcal{M}(p_0)$ is compact, the program $\sup_{\mu \in \mathcal{M}(p_0)} \sum_{i \in N} \widehat{v}_i^*(\mu)$ admits an optimum. *Q.E.D.*

B. Proofs for Section 3

Toward proving the results of Section 3, some preliminary claims will be useful.

Claim 1. *Suppose $i \in N$ and $\mu \in \Delta(\Pi \times \Omega)$. If τ_i is an optimal solution to*

$$\min_{\tau_i \in \Delta\Delta(\Pi \times \Omega)} \int \frac{c_i(\mu_i^\Omega)}{\iota_i(\mu_i^\Pi)} \, d\tau_i(\mu_i) \quad \text{subject to} \quad \int \mu_i \, d\tau_i(\mu_i) = \mu,$$

then no $\tilde{\omega}, \hat{\omega} \in \Omega$ with $c_i(\tilde{\omega}) = c_i(\hat{\omega})$ and distinct $\tilde{\beta}, \hat{\beta} \in \Delta\Pi$ have both $\tilde{\beta} \otimes \delta_{\tilde{\omega}}$ and $\hat{\beta} \otimes \delta_{\hat{\omega}}$ in the support of τ_i.

Proof. Suppose $\tilde{\omega}, \hat{\omega} \in \Omega$ with $c_i(\tilde{\omega}) = c_i(\hat{\omega}) =: \overline{c}_i$ and distinct $\tilde{\beta}, \hat{\beta} \in \Delta\Pi$ have both $\tilde{\beta} \otimes \delta_{\tilde{\omega}}$ and $\hat{\beta} \otimes \delta_{\hat{\omega}}$ in the support of τ_i. Then, some $\varepsilon \in (0, 1]$ and $\check{\tau}_i \in \Delta\Delta(\Pi \times \Omega)$ exists such that

$$\tau_i = (1 - \varepsilon)\check{\tau}_i + \tfrac{\varepsilon}{2}\delta_{\tilde{\beta} \otimes \delta_{\tilde{\omega}}} + \tfrac{\varepsilon}{2}\delta_{\hat{\beta} \otimes \delta_{\hat{\omega}}}.$$

The alternative belief distribution

$$\tau_i' = (1-\varepsilon)\check{\tau}_i + \varepsilon\delta_{\frac{1}{2}\left(\tilde{\beta}\otimes\delta_{\tilde{\omega}}+\hat{\beta}\otimes\delta_{\hat{\omega}}\right)}$$

is then feasible in the given program. Moreover, by strict convexity of $\frac{\bar{c}_i}{\iota_i(\beta)}$ in $\beta \in \Delta\Pi$, the latter attains a strictly lower loss, so that τ_i is not optimal. *Q.E.D.*

Claim 2. *Suppose $i \in N$ and $\beta_0 \in \Delta\Pi$. If τ_i is an optimal solution to the program*

$$\min_{\tau_i\in\Delta\Delta(\Pi\times\Omega)} \int \frac{c_i(\mu_i^\Omega)}{\iota_i(\mu_i^\Pi)}\ d\tau_i(\mu_i) \quad \text{subject to} \quad \int (\mu_i^\Pi, \mu_i^\Omega)\ d\tau_i(\mu_i) = (\beta_0, p_0), \tag{4}$$

then some alternative optimal $\tilde{\tau}_i$ exists such that

- *Each $\omega \in \Omega$ admits a unique $\tilde{\beta}_\omega \in \Delta\Pi$ such that $\tilde{\tau}_i(\tilde{\beta}_\omega \otimes \delta_\omega) = p_0(\omega)$;*
- *Any μ_i in the support of τ_i and any $\omega, \hat{\omega} \in \Omega$ in the support of μ_i^Ω have $\tilde{\beta}_\omega = \tilde{\beta}_{\hat{\omega}}$.*

Proof. Let $\tilde{\tau}_i := \int\int \delta_{\mu_i^\Pi\otimes\delta_\omega}\ \mathrm{d}\mu_i^\Omega(\omega)\ \mathrm{d}\tau_i(\mu_i) \in \Delta\Delta(\Pi\times\Omega)$.

Various features are immediate from the construction. First, the average marginal distributions under $\tilde{\tau}_i$ are the same as those under τ_i, making $\tilde{\tau}_i$ feasible in the program. Second, because the fraction $\frac{c_i(\mu_i^\Omega)}{\iota_i(\mu_i^\Pi)}$ is affine in μ_i^Ω when holding μ_i^Π fixed, we know $\tilde{\tau}_i$ yields the same value in program (4) as τ_i does, and so is optimal too. Third, every $\tilde{\mu}_i$ in the support of $\tilde{\tau}_i$ admits some $\tilde{\beta} \in \Delta\Pi$ and $\omega \in \Omega$ for which $\tilde{\mu}_i = \tilde{\beta} \otimes \delta_\omega$. Fourth, for any μ_i in the support of τ_i and any $\omega, \hat{\omega} \in \Omega$ in the support of μ_i^Ω, some $\tilde{\beta} \in \Delta\Pi$ has both $\tilde{\beta}\otimes\delta_\omega$ and $\tilde{\beta}\otimes\delta_{\hat{\omega}}$ in the support of $\tilde{\tau}_i$—indeed, $\tilde{\beta} = \mu_i^\Pi$ has this property.

The claim will then follow if we know that no $\omega \in \Omega$ and distinct $\tilde{\beta}, \hat{\beta} \in \Delta\Pi$ have both $\tilde{\beta}\otimes\delta_\omega$ and $\hat{\beta}\otimes\delta_\omega$ in the support of $\tilde{\tau}_i$. And indeed, this fact follows directly from Claim 1. *Q.E.D.*

Claim 3. *For any $c_H \geq c_L > 0$, the program*

$$\min_{(\beta^H,\beta^L)\in[0,1]^2} \left\{ \frac{c_H}{(1-\beta^H)(P_1-P_0)+\beta^H(P_2-P_1)} + \frac{c_L}{(1-\beta^L)(P_1-P_0)+\beta^L(P_2-P_1)} \right\} \quad \text{subject to} \quad \beta^H + \beta^L = 1$$

has a unique optimal solution (β^H, β^L). *It has*

$$\beta^H = \begin{cases} \frac{\sqrt{c_H}-\varphi\sqrt{c_L}}{(1-\varphi)(\sqrt{c_H}+\sqrt{c_L})} & : \ \varphi\sqrt{c_H} < \sqrt{c_L} \\ 1 & : \ \textit{otherwise.} \end{cases}$$

Moreover, if $c_H > c_L$*, then* $\beta^H > \frac{1}{2}$.

Proof. Substituting in $\beta^L = 1 - \beta^H$, we can view the program as an optimization over $\beta^H \in [0, 1]$. The loss is continuous in β^H so that an optimum exists, and it is strictly convex in β^H so that this optimum is unique. Direct computation shows that the given form of β^H satisfies the first-order condition, and hence is the optimum.

Finally, supposing $c_H > c_L$, let us show $\beta^H > \frac{1}{2}$. Indeed, in this case,

$$2(\sqrt{c_H} - \varphi\sqrt{c_L}) - (1-\varphi)(\sqrt{c_H} + \sqrt{c_L}) = (1+\varphi)(\sqrt{c_H} - \sqrt{c_L}) > 0,$$

so that $\beta^H \geq \min\left\{1, \frac{\sqrt{c_H}-\varphi\sqrt{c_L}}{(1-\varphi)(\sqrt{c_H}+\sqrt{c_L})}\right\} > \frac{1}{2}$. *Q.E.D.*

B.1. Toward Proposition 1

Proof of Proposition 1. Some optimal solution to program (3) exists by Fact 1. Moreover, by Claim 1, any optimal solution (μ, τ_1, τ_2) has $\tau_1^\Pi(\mu^\Pi) = \tau_2^\Pi(\mu^\Pi) = 1$.

Hence, all that remains to see is that the program

$$\min_{\beta \in \Delta\Pi} \sum_{i \in N} \frac{c_i}{\iota_i(\beta)}$$

is uniquely solved by setting

$$\beta(\pi^1) = \begin{cases} \frac{\sqrt{c_H}-\varphi\sqrt{c_L}}{(1-\varphi)(\sqrt{c_H}+\sqrt{c_L})} & : \ \varphi\sqrt{c_H} < \sqrt{c_L} \\ 1 & : \ \text{otherwise,} \end{cases}$$

which follows directly from Claim 3 (with $\beta(\pi^1)$ corresponding to β^H in that claim). *Q.E.D.*

B.2. Toward Proposition 2

Claim 4. *Suppose $c_1(1) = c_2(2) > c_2(1) = c_1(2)$. Let $i \in N$, let $\beta_0 \in \Delta\Pi$ be uniform, and suppose τ_i is a feasible solution to the program (4) from Claim 2's statement. Then, some feasible solution to program (3) exists that generates loss $2\int \frac{c_i(\mu_i^\Omega)}{\iota_i(\mu_i^\Pi)}\, d\tau_i(\mu_i)$.*

Proof. Let $\psi : \Pi \times \Omega \to \Pi \times \Omega$ be the involution that changes every coordinate.[7] Define $\Psi : \Delta(\Pi \times \Omega) \to \Delta(\Pi \times \Omega)$ by letting $\Psi(\tilde{\mu}) := \tilde{\mu} \circ \psi^{-1}$ for every $\tilde{\mu} \in \Delta(\Pi \times \Omega)$. Let j be such that $N = \{i, j\}$, and define $\tau_j := \tau_i \circ \Psi^{-1}$. It follows from $v_1^* = v_2^* \circ \Psi$ that

$$\sum_{k\in N} \int \frac{c_k(\mu_k^\Omega)}{\iota_k(\mu_j^\Pi)}\, \mathrm{d}\tau_k(\mu_k) = 2\int \frac{c_i(\mu_i^\Omega)}{\iota_i(\mu_i^\Pi)}\, \mathrm{d}\tau_i(\mu_i).$$

If some $\mu \in \Delta(\Pi \times \Omega)$ is such that (μ, τ_1, τ_2) is feasible in program (3), we will have a feasible triple with the desired property. To that end, define $\mu := \int \mu_i\, \mathrm{d}\tau_i(\mu_i)$, and note that $\int \mu_j\, \mathrm{d}\tau_j(\mu_j) = \Psi(\mu)$ by construction. It then suffices to observe that $\mu = \Psi(\mu)$. But this property follows from both marginals μ^Π, μ^Ω being uniform on their respective domains.[8] *Q.E.D.*

Claim 5. *Suppose $c_1(1) = c_2(2) =: c_H > c_L := c_2(1) = c_1(2)$. Let $i \in N$, let $\beta_0 \in \Delta\Pi$ be uniform, and suppose τ_i is an optimal solution to the program (4) from Claim 2's statement. If $\tau_i\{\mu_i \in \Delta(\Pi \times \Omega) : \ \mu_i^\Omega(\omega) = 1$ for some $\omega \in \Omega\} = 1$, then $\tau_i(\beta_1^* \otimes \delta_1) = \tau_i(\beta_2^* \otimes \delta_2) = \frac{1}{2}$, where*

$$\begin{aligned}\beta_1^*(\pi^1) = \beta_2^*(\pi^2) &= \begin{cases} \frac{\sqrt{c_H} - \varphi\sqrt{c_L}}{(1-\varphi)(\sqrt{c_H}+\sqrt{c_L})} & : \ \varphi\sqrt{c_H} < \sqrt{c_L} \\ 1 & : \ \textit{otherwise} \end{cases} \\ &> \tfrac{1}{2}.\end{aligned}$$

[7] So, if $N = \{i, j\} = \{i', j'\}$, then $\psi(\pi^i, i') = (\pi^j, j')$.

[8] Consider the 2×2 matrix whose (i', j') entry is $\mu(\pi^{i'}, j') - \frac{1}{4}$ for each $i', j' \in N$. Every row and every column of this matrix sums to zero, and so it is proportional to $\pm\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$.

Proof. Assume τ_i has the hypothesized properties. First, observe no $\omega \in \Omega$ and distinct $\tilde{\beta}, \hat{\beta} \in \Delta\Pi$ have both $\tilde{\beta} \otimes \delta_\omega$ and $\hat{\beta} \otimes \delta_\omega$ in the support of τ_i, by Claim 1. Hence, some $\beta_1, \beta_2 \in \Delta\Pi$ exist such that $\tau_i\{\beta_1 \otimes \delta_1, \beta_2 \otimes \delta_2\} = 1$. Optimality of τ_i for program (4) then tells us $(\beta_i(\pi^i), \beta_i(\pi^j))$ is an optimal solution to

$$\min_{(\beta^H,\beta^L)\in[0,1]^2} \left\{ \frac{c^H}{(1-\beta^H)(P_1-P_0)+\beta^H(P_2-P_1)} + \frac{c^L}{(1-\beta^L)(P_1-P_0)+\beta^L(P_2-P_1)} \right\} \text{ subject to } \beta^H + \beta^L = 1.$$

The claim then follows directly from Claim 3. *Q.E.D.*

Now, we prove Proposition 2.

Proof of Proposition 2. Let (μ, τ_1, τ_2) be any optimal solution to (3) (which exists by Fact 1).

Our first step is to construct an alternative optimum that satisfies a symmetry property. To construct such an optimum, recall the map $\Psi : \Delta(\Pi \times \Omega) \to \Delta(\Pi \times \Omega)$ defined in the proof of Claim 4. Symmetry of p_0 implies $\Psi(\mu) \in \mathcal{M}(p_0)$ because $\mu \in \mathcal{M}(p_0)$; because $\mathcal{M}(p_0)$ is convex, it therefore also contains $\hat{\mu} := \frac{1}{2}[\mu + \Psi(\mu)]$. For each $\{i, j\} = N$, define $\hat{\tau}_i := \frac{1}{2}[\tau_i + \tau_j \circ \Psi^{-1}]$.

Some properties of $(\hat{\mu}, \hat{\tau}_1, \hat{\tau}_2)$ are immediate from the construction. First, the mean of $\hat{\tau}_i$ is $\hat{\mu}$ for each $i \in N$, so that $(\hat{\mu}, \hat{\tau}_1, \hat{\tau}_2)$ is feasible in program (3). Second, $\hat{\tau}_1 = \hat{\tau}_2 \circ \Psi^{-1}$. Third, that $v_1^* = v_2^* \circ \Psi$ implies $(\hat{\mu}, \hat{\tau}_1, \hat{\tau}_2)$ attains the same value as (μ, τ_1, τ_2) does in program (3), and so is optimal too.

Now, let $\beta_! \in \Delta\Pi$ be the uniform distribution and $i \in N$. Let us show, for $\beta_0 = \beta_!$ and $i \in N$, that $\hat{\tau}_i$ solves the program (4) defined in Claim 2's statement. Assume otherwise for a contradiction. So some $\check{\tau}_i \in \Delta\Delta(\Pi \times \Omega)$ has $\int(\mu_i^\Pi, \mu_i^\Omega)\, \mathrm{d}\check{\tau}_i(\mu_i) = (\beta_!, p_0)$ and $\int \frac{c_i(\mu_i^\Omega)}{\iota_i(\mu_i^\Pi)}\, \mathrm{d}\check{\tau}_i(\mu_i) < \int \frac{c_i(\mu_i^\Omega)}{\iota_i(\mu_i^\Pi)}\, \mathrm{d}\hat{\tau}_i(\mu_i)$. By Claim 4, some feasible solution to program (3) generates loss $2\int \frac{c_i(\mu_i^\Omega)}{\iota_i(\mu_i^\Pi)}\, \mathrm{d}\check{\tau}_i(\mu_i)$, contradicting the (previously established) optimality of $(\hat{\mu}, \hat{\tau}_1, \hat{\tau}_2)$ in program (3).

Having established $\hat{\tau}_i$ is optimal in program (4), for $\beta_0 = \beta_!$ and $i \in N$, let $\tilde{\tau}_i$ be as delivered by Claim 2. So $\tilde{\tau}_i$ is optimal in program (4), and

- Each $\omega \in \Omega$ admits a unique $\tilde{\beta}^i_\omega \in \Delta\Pi$ such that $\tilde{\tau}_i(\tilde{\beta}^i_\omega \otimes \delta_\omega) = p_0(\omega)$;

- Any μ_i in the support of $\hat{\tau}_i$ and any $\omega, \hat{\omega} \in \Omega$ in the support of μ_i^Ω have $\tilde{\beta}^i_\omega = \tilde{\beta}^i_{\hat{\omega}}$.

We can then apply Claim 5 to $\tilde{\tau}_i$, to learn $\tilde{\tau}_i$ is the uniform distribution over $\{\beta_1^* \otimes \delta_1,\ \beta_2^* \otimes \delta_2\}$. That $\beta_1^* \neq \beta_2^*$ (which holds because $\beta_1^*(\pi^1) = \beta_2^*(\pi^2) > \frac{1}{2}$) then implies (by the second bullet above) no μ_i in the support of $\hat{\tau}_i$ has μ_i^Ω putting positive probability on both values for the fundamental state.

Given the previous observation, for each $i \in N$, we can now apply Claim 5 to $\hat{\tau}_i$, to learn $\hat{\tau}_i$ is the uniform distribution over $\{\beta_1^* \otimes \delta_1,\ \beta_2^* \otimes \delta_2\}$ too. But then, by construction of $\hat{\tau}_i$, it would follow that $\tau_i \in \Delta\{\beta_1^* \otimes \delta_1,\ \beta_2^* \otimes \delta_2\}$ too. Finally, because $\int \mu_i^\Omega \, \mathrm{d}\tau_i(\mu_i) = p_0$, the only possibility for τ_i is that it is uniform as well. Because the pair (τ_1, τ_2) determines the total state distribution, the proposition follows. *Q.E.D.*

B.3. Toward Proposition 3

Claim 6. *Suppose c_2 is constant. If (μ, τ_1, τ_2) is optimal in program (3), then some alternative optimal $(\tilde{\mu}, \tilde{\tau}_1, \tilde{\tau}_2)$ exists such that*

- *The distribution $\tilde{\tau}_2$ is degenerate;*

- *Each $\omega \in \Omega$ admits a unique $\tilde{\beta}_\omega \in \Delta\Pi$ such that $\tilde{\tau}_1(\tilde{\beta}_\omega \otimes \delta_\omega) = p_0(\omega)$;*

- *Any μ_1 in the support of τ_1 and any $\omega, \hat{\omega} \in \Omega$ in the support of μ_1^Ω have $\tilde{\beta}_\omega = \tilde{\beta}_{\hat{\omega}}$.*

Proof. Let $\tilde{\tau}_1$ be as delivered by Claim 2 for $i = 1$ and $\beta_0 := \mu^\Pi$. Then, let $\tilde{\tau}_1 := \int \mu_1 \, \mathrm{d}\tilde{\tau}_1(\mu_1)$ and $\tilde{\tau}_2 := \delta_{\tilde{\mu}}$. By construction, $(\tilde{\mu}, \tilde{\tau}_1, \tilde{\tau}_2)$ is feasible in program (3), so all that remains is to see $(\tilde{\mu}, \tilde{\tau}_1, \tilde{\tau}_2)$ attains a weakly lower loss than (μ, τ_1, τ_2) does.

Let us observe $\int \frac{c_i(\mu_i^\Omega)}{\iota_i(\mu_i^\Pi)} \, \mathrm{d}\tilde{\tau}_i(\mu_i) \leq \int \frac{c_i(\mu_i^\Omega)}{\iota_i(\mu_i^\Pi)} \, \mathrm{d}\tau_i(\mu_i)$ for each agent $i \in N$. For $i = 1$, the inequality follows from optimality of $\tilde{\tau}_1$ in program (4) from Claim 2's statement. For $i = 2$, the inequality follows from $\tilde{\tau}^{\Pi}$ being degenerate, the identity $\tilde{\mu}^{\Pi} = \mu^{\Pi}$, and the integrand $\frac{c_2(\mu_2^{\Omega})}{\iota_2(\mu_2^{\Pi})} = \frac{c_2}{\iota_2(\mu_2^{\Pi})}$ being a convex function of the marginal μ_2^{Π}. *Q.E.D.*

Claim 7. *Suppose c_2 is constant and a unique $\vec{\beta} \in (\Delta\Pi)^{\Omega}$ minimizes*

$$\int \frac{c_1(\omega)}{\iota_1(\beta_\omega)} \; dp_0(\omega) + \frac{c_2}{\iota_2\big(\int \beta_\omega \; dp_0(\omega)\big)},$$

and $\beta_\omega \neq \beta_{\hat{\omega}}$ for all distinct $\omega, \hat{\omega} \in \Omega$, then every optimal solution (μ, τ_1, τ_2) to program (3) has

- $\tau_1(\beta_\omega \otimes \delta_\omega) = p_0(\omega)$ *for every* $\omega \in \Omega$*;*
- $\tau_2^{\Pi}\left(\int \beta_\omega \; dp_0(\omega)\right) = 1$*;*
- τ_1^{Π} *is a strict mean-preserving spread of* τ_2^{Π}*, and* τ_1^{Ω} *is a strict mean-preserving spread of* τ_2^{Ω}*.*

Proof. The third point follows immediately from the first two given that the entries of $\vec{\beta}$ are distinct: the first point implies τ_1^{Ω} is maximally informative and τ_1^{Π} is strictly informative, while the second point implies τ_2^{Π} is uninformative and τ_2^{Ω} is not maximally informative. Moreover, the second point follows directly from the first because the entries of $\vec{\beta}$ are all distinct, given Claim 1. So we turn to showing every optimal (μ, τ_1, τ_2) for program (3) satisfies the first point.

Consider first any optimal $(\hat{\mu}, \hat{\tau}_1, \hat{\tau}_2)$ for program (3) with the property that $\hat{\tau}_1$ reveals the fundamental state—that is, such that every belief in the support of $\hat{\tau}_1$ takes the form $\hat{\beta} \otimes \delta_{\hat{\omega}}$ for some $\hat{\beta} \in \Delta\Pi$ and $\hat{\omega} \in \Omega$. By Claim 1, no $\hat{\omega} \in \Omega$ and distinct $\beta, \hat{\beta} \in \Delta\Pi$ can exist such that $\beta \otimes \delta_{\hat{\omega}}$ and $\hat{\beta} \otimes \delta_{\hat{\omega}}$ are both in the support of $\hat{\tau}_1$. Said differently, every $\hat{\omega} \in \Omega$ admits a unique $\hat{\mu}_1$ in the support of $\hat{\tau}_1$ with $\hat{\mu}_1^{\Omega}(\hat{\omega}) > 0$. The uniqueness property of $\vec{\beta}$ then directly implies that $\hat{\tau}_1(\beta_{\hat{\omega}} \otimes \delta_{\hat{\omega}}) = p_0(\hat{\omega})$ for every $\hat{\omega} \in \Omega$.

In light of the above paragraph, it suffices to show, for any optimal (μ, τ_1, τ_2) for program (3), that τ_1 reveals the fundamental state. To that end, apply Claim 6: some optimal solution

$(\tilde{\mu}_1, \tilde{\tau}_1, \tilde{\tau}_2)$ to program (3) exists such that:

- The distribution $\tilde{\tau}_2^{\Pi}$ is degenerate;
- Each $\omega \in \Omega$ admits a unique $\tilde{\beta}_\omega \in \Delta\Pi$ such that $\tilde{\tau}_1(\tilde{\beta}_\omega \otimes \delta_\omega) = p_0(\omega)$;
- Any μ_1 in the support of τ_1 and any $\omega, \hat{\omega} \in \Omega$ in the support of μ_1^{Ω} have $\tilde{\beta}_\omega = \tilde{\beta}_{\hat{\omega}}$.

Now, the uniqueness property of $\vec{\beta}$, together with optimality of $(\tilde{\mu}, \tilde{\tau}_1, \tilde{\tau}_2)$, implies $(\tilde{\beta}_\omega)_{\omega\in\Omega} = \vec{\beta}$. Hence, because the entries of $\vec{\beta}$ are distinct, it follows that every μ_1 in the support of τ_1 admits some $\omega \in \Omega$ such that $\mu_1^{\Omega}(\omega) = 1$. Said differently, τ_1 reveals the fundamental state, as required. *Q.E.D.*

Claim 8. *Take $c_1(1) =: c_H > c_L := c_2(1) = c_2(2) = c_1(2)$. The program*

$$\min_{\vec{\beta}\in(\Delta\Pi)^{\Omega}} \int \frac{c_1(\omega)}{\iota_1(\beta_\omega)}\ dp_0(\omega) + \frac{c_2}{\iota_2\left(\int \beta_\omega\ dp_0(\omega)\right)}$$

has a unique optimal solution $(\beta_1^{**}, \beta_2^{**})$. It has

$$\left(\beta_1^{**}(\pi^1), \beta_2^{**}(\pi^1)\right) = \begin{cases} \left(\frac{(2+\varphi)\sqrt{c_H}-3\varphi\sqrt{c_L}}{(1-\varphi)(3\sqrt{c_L}+\sqrt{c_H})}, \frac{(2-\varphi)\sqrt{c_L}-\varphi\sqrt{c_H}}{(1-\varphi)(3\sqrt{c_L}+\sqrt{c_H})}\right) & : \frac{\sqrt{c_H}}{\sqrt{c_L}} \leq \frac{3}{1+2\varphi} \\ (1, 1/3) & : \text{otherwise.} \end{cases}$$

In particular, $\beta_1^{**} \neq \beta_2^{**}$.

Proof. Substituting in $\beta_\omega(\pi^2) = 1 - \beta_\omega(\pi^1)$ for each $\omega \in \Omega$, we can view the program as an optimization over $(\beta_1(\pi^1), \beta_2(\pi^1)) \in [0,1]^2$. The loss is continuous so that an optimum exists, and it is strictly convex so that this optimum is unique. Direct computation shows that the given form of $(\beta_1^{**}(\pi^1), \beta_2^{**}(\pi^1))$ satisfies the first-order condition, and hence is the optimum.

Finally, let us verify that $\beta_1^{**} \neq \beta_2^{**}$. Given the form of the solution, we need only check that the numerators differ in the case that $\frac{\sqrt{c_H}}{\sqrt{c_L}} \leq \frac{3}{1+2\varphi}$. And indeed,

$$[(2+\varphi)\sqrt{c_H} - 3\varphi\sqrt{c_L}\ ] - [(2-\varphi)\sqrt{c_L} - \varphi\sqrt{c_H}\ ] = 2(1+\varphi)\left(\sqrt{c_H} - \sqrt{c_L}\right) > 0.$$

Q.E.D.

Now, we prove Proposition 3.

Proof of Proposition 3. Some optimal solution to program (3) exists by Fact 1. Moreover, any two triples that satisfy the conditions of the proposition's statement—which yield the same total state distribution, provide the same information to agent 1 about the total state, and provide the same information to agent 2 about the ranking state—generate the exact same loss (and so are either both optimal or both suboptimal). Hence, given Claim 7, we need only see that $(\beta^{**}_\omega)_{\omega\in\Omega}$ is the unique solution to the program

$$\min_{\vec{\beta}\in(\Delta\Pi)^\Omega} \int \tfrac{c_1(\omega)}{\iota_1(\beta_\omega)} \, \mathrm{d}p_0(\omega) + \tfrac{c_2}{\iota_2\left(\int \beta_\omega \, \mathrm{d}p_0(\omega)\right)},$$

and that $\beta^{**}_1 \neq \beta^{**}_2$—exactly what Claim 8 proves. *Q.E.D.*