ADDRESSING STRATEGIC UNCERTAINTY WITH INCENTIVES AND INFORMATION

by

Marina Halac, Elliot Lipnowski, and Daniel Rappoport

COWLES FOUNDATION PAPER NO. 1789

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

2022

http://cowles.yale.edu/
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 and information. Each agent’s payoff depends on the profile of agents’ actions, his allocation, and the fundamental state. We solve for the principal’s 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, Oyama and Takahashi (2020); Morris, Oyama, and Takahashi (2020); and Halac, Lipnowski, and Rappoport (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 (2021). Our two-step procedure permits an explicit characterization of the principal’s solution, and we describe 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 contracts and on information design under adversarial equilibrium selection. In addition to the papers just cited, see Segal (2003); Bernstein and Winter (2012); Chassang, Del Carpio, and Kapon (2020); Halac, Kremer, and Winter (2020, 2022); and Camboni and Porcellacchia (2021) on incentive design; Hoshino (2021), Mathevet, Perego, and Taneva (2020); Inostroza and Pavan (2022); and Li, Song, and Zhao (2021) on information design; and work related to the latter strand such as Kajii and Morris (1997). Our paper studies both of these tools jointly. We provide a general methodology that can be useful for various applications in which strategic uncertainty may be addressed with incentives and information.

I. 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 $p_0 ∈ ΔΩ$ with full support.\footnote{Throughout, given a set $Z$, let $ΔZ$ denote the set of finite-support probability distributions over $Z$.} The principal offers each agent $i ∈ N$ a private allocation $x_i ∈ X_i$, 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 $σ = ⟨q, χ⟩$, where $q ∈ Δ(ℕ^2)^N × Ω$ is a prior with marginal distribution $p_0$ on $Ω$ and $χ = (χ_i)_{i ∈ N}$ is an allocation rule, with $χ_i: \text{supp}(\text{marg}_i(q)) → X_i$. Let $T_i^q := \text{supp}(\text{marg}_i(q))$ denote the support of the marginal of $q$ along dimension $i$ and $T^q := \prod_{i ∈ N} T_i^q$. The interpretation is that the principal privately informs each agent $i ∈ N$ of his type $t_i ∈ T_i^q$ and, through the allocation rule, of his prescribed allocation.\footnote{Given the finite type restriction, since the type itself is a strategically irrelevant label (Dekel, Fundenberg, and Morris 2007, Proposition 1), it is immaterial that types are labeled with natural number pairs.}

Hence, an agent may face uncertainty about other agents’ contracts and about the fundamental state, but is completely informed about his own contract.\footnote{Note, however, that the space of allocations can be rich. For example, if $X_i$ stipulates fundamental-state-contingent payments, then the assumption that agents observe their own contracts is consistent with them facing uncertainty about their own payments. Even so, this assumption can entail a payoff loss for the principal in fixed settings, as we demonstrate in Section V of Halac, Lipnowski, and Rappoport (2021).} 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 $σ = ⟨q, χ⟩$ defines a Bayesian game between the agents. In this game, $⟨(T_i^q)_{i ∈ N}, Ω, q⟩$ is a common-prior type space; agents simultaneously make type-contingent decisions of whether to choose 1 or 0; and agent $i$’s payoff is a function $u_i: 2^N × X_i × Ω → ℝ$ of the set of agents who choose 1, his allocation, and the fundamental state. The principal wishes to uniquely induce all agents to choose 1 (with probability 1), with her payoff in that event given by $\sum_{i ∈ N} v_i(x_i, ω)$ for $v_i: X_i × Ω → ℝ$ bounded above. Say an action is rationalizable for an agent type if it is interim correlated rationalizable, and say an incentive scheme $σ = ⟨q, χ⟩$ 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

$$\sup_{σ \text{ is UIF}} V(σ),$$

where $V(σ)$ is her total expected payoff given scheme $σ = ⟨q, χ⟩$ and all agents choosing 1:

$$V(σ) = \sum_{i ∈ N, ω ∈ Ω} q(t_i, ω) \sum_{i ∈ N} v_i(χ_i(t_i), ω).$$

We make a dominant-allocation assumption that says that for each agent $i ∈ N$, there exists $x_i ∈ X_i$ such that choosing action 1 is dominant:

$$\min_{J ⊆ N \backslash \{i\}, ω ∈ Ω} \left[ u_i(J \cup \{i\}, x_i, ω) − u_i(J, x_i, ω) \right] > 0.$$

Under this assumption, the principal can always choose 1 uniquely rationalizable.\footnote{By combining our analysis with that in Morris, Oyama, and Takahashi (2020), it may be possible to weaken our dominant-allocation assumption.} Our focus is on solving for optimal incentive schemes that achieve this goal. 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.\footnote{That is, for any $ε > 0$, our proof constructs a UIF scheme $σ_ε$ such that $V(σ_ε) > \sup_{σ \text{ is UIF}} V(σ) − ε$.}

REMARK 1: A special case of our model is the case of a supermodular game, in which $J → u_i(J \cup \{i\}, x_i, ω) − u_i(J, x_i, ω)$ is a weakly increasing map on $2^N \backslash \{i\}$ for every $i ∈ N$, $x_i ∈ X_i$, and $ω ∈ Ω$. 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 $\prod_{i ∈ N} X_i$, and that the principal’s objective (conditional on all agents choosing 1) is additively separable. Our tools can be useful even without these separability conditions, if we suitably generalize our dominant-allocation...
assumption[4] In contrast, relaxing our assumptions that actions are binary and agent preferences take a private-value form seems more challenging.

II. 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 $\Pi$ 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 $t = (t_i^\pi, t_j^\pi)_{i \in N} \in T^q$ has $t_i^\pi \neq t_j^\pi$ for all distinct $i, j \in N$. Any such type profile $t$ induces a ranking state $\pi(t) \in \Pi$ given by $\pi(t) = |\{ j \in N : t_i^\pi \leq t_j^\pi \}|$. 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 $(\pi, \omega)$ as the total state.

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

$$
\mu_i^q(\hat{\pi}, \hat{\omega} | t_i) := q_i(\{ t_i: \pi(t_i, t_i) = \hat{\pi} \} \times \{ \hat{\omega} \} | t_i)
$$

for all $\hat{\pi} \in \Pi, \hat{\omega} \in \Omega$,

where $q_i: T_i^q \rightarrow \Delta(T_i^q \times \Omega)$ is given by

$$
q_i(t_i, \omega | t_i) := \frac{1}{\text{marg},q(t_i)} q_i(t_i, t_i, \omega).
$$

The total state distribution $\mu_i^q \in \Delta(\Pi \times \Omega)$ is given by

$$
\mu_i^q(\hat{\pi}, \hat{\omega}) := q_i(\{ t: \pi(t) = \hat{\pi} \} \times \{ \hat{\omega} \})
$$

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

$$
I_i(x_i, \pi, \omega)
$$

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

the agent’s set of sufficient allocations is given by

$$
X_i^*(\mu_i)
$$

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

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

**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 $\chi_i(t_i) \in X_i^*(\mu_i^q(\cdot | t_i))$.

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 $\sigma$ is UIF, there exists a strict ranking scheme $\sigma^*$ 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 coincides with 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

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

---

6 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 et al. (2021) can be applied.
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 \((\pi, \omega)\). Formally, for any agent \( i \in N \) and distribution \( \mu \in \Delta (\Pi \times \Omega) \), define

\[
\hat{v}_i^* (\mu) := \sup_{\tau_i \in \Delta \Delta (\Pi \times \Omega)} \int v_i^*(\mu_i) \, d\tau_i (\mu_i)
\]

subject to

\[
\int \mu_i \, d\tau_i (\mu_i) = \mu,
\]

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

**THEOREM 1:** The principal’s optimal value satisfies

\[
\sup_{\sigma} V(\sigma) = \sup_{\mu \in \mathcal{M}(p_0), i \in N} \sum_{i \in N} \hat{v}_i^* (\mu).
\]

The reduction in Theorem 1 is significant. Instead of optimizing over fundamental-state-contingent distributions over type profiles \( \sigma \in \Delta \left( [\mathbb{N}^2] \times \Omega \right) \), 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 \( \sup_{\sigma \text{ is UIF}} V(\sigma) \leq \sup_{\mu \in \mathcal{M}(p_0)} \sum_{i \in N} \hat{v}_i^* (\mu) \) 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 (2021), but with types augmented to convey information about the total state

\[7\] In that paper’s setting, this augmentation was not needed as providing no information was optimal.

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

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

**FACT 1:** If \( v_i^* \) is upper semicontinuous for each \( i \in N \), then some optimal \( (\mu_i, (\tau_i)_{i \in N}) \) exists.

### III. 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 \( \omega \) 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 \( x_j \in X_j = \mathbb{R}_+ \) is a bonus that the principal pays agent \( j \) in the case of success. We thus write agent \( j \)’s payoff as \( u_i(\pi, x_j, \omega) = P_k x_j - c_j(\omega) I_{i \in J} \). The principal’s goal is to uniquely induce the agents to work at the least possible incentive cost, so \( v_j(x_j, \omega) = -x_j \).

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 \( \mu_i \) along \( \Pi \) and \( \Omega \) respectively, and let \( \pi^J \in \Pi \) be the ranking state in which agent \( i \) is ranked second. Defining the expected marginal product

\[
\nu_i(\mu_i^\Pi) := \left[ 1 - \mu_i^\Pi(\pi^J) \right] (P_1 - P_0) + \mu_i^\Pi(\pi^J) (P_2 - P_1),
\]
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 \( X_i^\Omega(\mu_i) = \{ x_i \in X_i : x_i t_i(\mu_i^\Omega) > c_i(\mu_i^\Omega) \} \). Hence, \( v_i^\ast(\mu_i) = -c_i(\mu_i^\Omega) / t_i(\mu_i^\Omega) \), and replacing the objective with its negative, the principal’s problem can be written as

\[
\inf_{\mu \in \mathcal{M}(p_0)} \sum_{i \in N} \int \frac{c_i(\mu_i^\Omega)}{t_i(\mu_i^\Omega)} d\tau_i(\mu_i) \tag{3}
\]

subject to

\[
\int \mu_1 d\tau_1(\mu_1) = \int \mu_2 d\tau_2(\mu_2) = \mu.
\]

We next present different examples that vary in how agents’ effort costs depend on the fundamental state. We denote by \( \tau_i^\Pi \in \Delta \Pi \) and \( \tau_i^\Omega \in \Delta \Omega \) the distributions of the marginals of \( \mu_i \) along \( \Pi \) and \( \Omega \) 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 \( c_1(1) = c_1(2) = c_H \geq c_1(1) = c_2(1) = c_L \). Then a feasible \((\mu, \tau_1, \tau_2)\) is optimal if and only if \( \tau_1^\Pi(\mu_i^\Omega) = \tau_2^\Pi(\mu_i^\Omega) \) and

\[
\mu(\tau^\Pi) = 1
\]

\[
\begin{cases}
\sqrt{c_H} - \varphi \sqrt{c_L} : \varphi \sqrt{c_H} < \sqrt{c_L} \\
1 - \varphi (\sqrt{c_H} + \sqrt{c_L}) : otherwise.
\end{cases}
\]

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 (2021). When \( c_i \) is constant for each \( i \in N \), the interim value functions \( v_i^\ast \) are all concave, so \( \tilde{v}_i^\ast = v_i^\ast \) 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 agent 1’s effort cost is relative to agent 2’s, the higher the probability \( \mu \) places on ranking state \( \tau^\Omega \) 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 \( c_1(1) = c_2(2) = c_H > c_L := c_1(1) = c_2(2) \). Then there is a unique optimal \((\mu, \tau_1, \tau_2)\). Each \( i \in N \) has \( \tau_i(\beta^\ast_1 \otimes \delta_1) = \tau_i(\beta^\ast_2 \otimes \delta_2) = 1/2, where

\[
\beta^\ast_1(\tau^\Pi) = \beta^\ast_2(\tau^\Pi) = \begin{cases}
\varphi \sqrt{c_H} : & \varphi \sqrt{c_H} < \sqrt{c_L} \\
1 - \varphi (\sqrt{c_H} + \sqrt{c_L}) : & otherwise.
\end{cases}
\]

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
\((\mu, \tau_1, \tau_2)\) is optimal if and only if 
\[ \tau_1 \left( \beta_w^* \otimes \delta_\omega \right) = 1/2 \] for each \( \omega \in \Omega \) and 
\[ \tau_2 \left( \int \beta_w^* \, dp_0(\omega) \right) = 1, \] where

\[ \left( \beta_1^*(\pi^1), \beta_2^*(\pi^1) \right) = \left( \frac{(2 + \varphi) \sqrt{c_H} - 3 \varphi \sqrt{c_L}}{(1 - \varphi)(3 \sqrt{c_L} + \sqrt{c_H})}, \frac{(2 - \varphi) \sqrt{c_L} - \varphi \sqrt{c_H}}{(1 - \varphi)(3 \sqrt{c_L} + \sqrt{c_H})} \right) : \sqrt{c_L} \leq \frac{3}{1 + 2\varphi} \]

\[ (1, 1/3) : \text{otherwise.} \]

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, consider perfectly positively correlated agent effort costs: \( c_1(1) = c_2(1) > c_2(2) = c_1(2) \). Because the principal would like 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 \( \omega \in \Omega \), suppose the project succeeds with probability \( P_\omega(\omega) \) if exactly \( k \) agents work. By similar logic as in Proposition 1, we can show that if \( P_2 \) (and \( c_i \)) \( i \in \mathbb{N} \) are constant, then it is optimal to give agents no information about the realized total state. 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


