Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 56-2 (2023) 1754–1759

Signaling-based Robust Incentive Designs with Randomized Monitoring ⋆

Sina Sanjari * **Subhonmesh Bose** ** **Tamer Başar** ***

* *Department of Electrical and Computer Engineering at University of Illinois Urbana-Champaign, Illinois, USA. (e-mail: sanjari@illinois.edu).*

** *Department of Electrical and Computer Engineering at University of Illinois Urbana-Champaign, Illinois, USA. (e-mail: boses@illinois.edu)*

*** *Coordinated Science Laboratory at University of Illinois Urbana-Champaign, Illinois, USA (e-mail: basar1@illinois.edu)*

Abstract: Incentive design problems entail hierarchical decision-making where a leader crafts a strategy to induce a desired response from a follower. Such dynamic games with decentralized information structures have been well-studied under three assumptions–the leader must have access to the follower's observations, actions, and the objective function. Lack of knowledge on any of these can potentially lead to performance loss for the leader. In this paper, we first study a setup where the leader observes the follower's action through a random monitoring channel and learns about the follower's observation through a follower-designed signal. In this setup, we establish the existence of a signaling-based incentive equilibrium strategy for the leader that induces honest reporting and desired control response from the follower. Then, we study a setting, where the follower's costs are parametric, but the parameters are not known to the leader. We construct an incentive strategy that reduces the sensitivity of the leader's performance to uncertainty in the parameter, close to an initial estimate. More generally, for the case when the leader's knowledge about the follower's cost and distributions of cost-relevant random variables is inaccurate, we establish the existence of a robust incentive equilibrium strategy that bounds the performance loss from the inaccuracy in the model.

Keywords: Differential and dynamic games, multi-agent systems, robustness analysis, data-driven decision making, hierarchical multilevel and multilayer control.

1. INTRODUCTION

Incentive design problems entail hierarchical decision-making between at least two agents, a leader and a follower with different objectives. These are dynamic games where the leader starts by announcing a strategy, then the follower acts, following which the leader implements her announced strategy, based on the action taken by the follower. Stochastic incentive design problems add an element of richness to this setup and allow costs and observations of both players to depend on an uncertain state of nature and/or observations that may be public or private. The goal of incentive design is to find leader's strategies that induce a desired response from the follower. Such problems have found applications in various domains, e.g., in the design of tax codes, environmental regulations, and demand response.

Stochastic incentive design problems can be cast as stochastic Stackelberg games with dynamic information. Direct equilibrium characterization in such games is often challenging; rather, an indirect approach to equilibrium characterization becomes tractable. Specifically, one starts by solving a decentralized control problem among the leader and the follower with a static information structure that aims to optimize the leader's performance. Then, the resulting leader-optimal control strategies are utilized to design an incentive strategy that leverages the dynamic information available on the follower's action. This design is such that the follower's response to the leader's incentive strategy coincides with what the leader wants the follower to play. In other words, the leader's performance with such an incentive strategy (along with the follower's desired response) becomes leader-optimal, thus leading to a Stackelberg equilibrium of the incentive design game. The incentive strategies considered in Başar (1984), among others, are affine, and hence smooth, in the control action of the follower. From a strategy design standpoint, such "soft" smooth incentive strategies are preferred to discontinuous "threat" strategies, where the leader's reaction to the follower's action can be large even for a small deviation on the follower's part from his leader-optimal response.

The study of incentive design problems has a long history; see e.g., Başar (1984); Cansever and Başar (1985a); Başar (1983); Başar (1989); Başar and Olsder (1999); Ho et al. (1982); Zheng and Başar (1982); Zheng et al. (1984). In such problems, the leader attempts to influence the followers' decisions through the design of an

⋆ This research was supported in part by grant FA9550-19-1-0353 from AFOSR, and in part by the National Science Foundation under the grant CPS-2038403.

10.1016/j.ifacol.2023.10.1885

incentive strategy that leverages the follower's monitored actions/observations with accurate knowledge of the follower's cost structure. In practice, the leader may not have access to the follower's actions and observations. She might also lack accurate knowledge of the follower's cost structure. In this paper, we address incentive design under such lack of accurate information about the follower.

In Sections 2 and 3, we consider a randomized channel for monitoring observations and actions of the follower. Without the ability to perfectly monitor, one might surmise that the leader can never incentivize the follower to act in the leader's favor, at least with smooth incentive strategies. However, we show that this is not always the case. We allow the follower to design a signaling mechanism containing information regarding the follower's private observation, and send it to the leader. This signal is utilized by the leader to design a signaling-based incentive strategy. The follower may benefit from deceiving the leader, depending on her cost. This element of belief-shaping via signaling is not new in the literature on decentralized control and games, e.g., see Kamenica and Gentzkow (2011); Groves (1973); Fudenberg and Tirole (1991); Groves and Loeb (1979), and Dasgupta et al. (1979). In this work, we introduce signaling into the incentive design literature. In particular, in Theorem 1, under continuity and convexity of the follower's cost, we show that there exists an incentive equilibrium strategy for the leader that leads to extraction of an honest mechanism from the follower, in turn leading to revelation of the follower's private observation, and a leader-aligned response from the follower.

In Sections 4 and 5, we study robustness of incentive design to incorrect follower models. That is, we assume that the leader is no longer privy to the follower's true cost structure and/or the distributions of various cost-relevant random variables. We seek strategies that perform well, from the leader's vantage point, even when they are designed using possibly incorrect follower models. When the assumed models are "close" to the true models, then a robust incentive strategy leads to a performance that is close to that under the correct model. For such problems, in Theorems 2 and 3, we establish the existence of a robust signaling-based incentive equilibrium strategy under sufficient conditions on the convergence of a sequence of incorrect models. Our study of such problems is inspired by the literature on robust control design in stochastic control and game theory, e.g., in Başar and Bernhard (2008); Khalil et al. (1996); Kara and Yüksel (2020, 2019); Wiesemann et al. (2013); Hansen and Sargent (2001); Yüksel and Linder (2012).

Due to space limitations, the proofs of the results are not included here; they are available in an extended arXiv version of the paper.

2. STOCHASTIC STACKELBERG GAME $\mathcal{P}$

We study a single-stage Stackelberg game with dynamic information structure (IS) between leader L and follower F. L first announces a strategy at the start of the game. Then, F acts by taking an action and sending a signal to L regarding her private information. The realized costs of L and F depend on L's announced strategy as well as the actions and the signaling mechanism selected by F. In this section, we formally define this game and describe relevant equilibrium/optimality notions that we study in the sequel.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the underlying probability space describing the system's distinguishable events. Let $\mathbb{Y}^L$ be a subset of a finite-dimensional Euclidean space, endowed with its Borel σ-field $\mathcal{Y}^L$ that describes the possible private observations of L. Let $(\mathbb{Y}^F, \mathcal{Y}^F)$ describe the same for F. Also, let $\mathbb{U}^L$ be a subset of a finite-dimensional Euclidean space, that together with the Borel σ-field $\mathcal{U}^L$, describes the space of control actions u^L for L. Similarly, define $(\mathbb{U}^F, \mathcal{U}^F)$ and u^F for F. Each player selects a control action via an admissible strategy–a measurable map of her available information. F is privy to her own observation, i.e., the information is $I^F = \{y^F\}$ for F for an exogenous random variable y^F. Let Γ^F denote her set of admissible strategies as a set of all measurable functions γ^F from $(\mathbb{Y}^F, \mathcal{Y}^F)$ to $(\mathbb{U}^F, \mathcal{U}^F)$. F also designs a signal s that L observes. In other words, F designs a stochastic kernel $\pi(\cdot|y^F)$ that induces s. We assume that s takes values in $\mathbb{Y}^F$. Let Π denote the space of signaling strategies π. For L, we consider the following two ISs.

- Randomized observation- and control-sharing: L observes her private information y^L and the signal s sent by F; however, L only observes y^F and u^F via randomized channels. The status of these channels is determined by two binary random variables z^y and z^u, i.e., if $z^y = 1$ ($z^u = 1$), L observes y^F (u^F), otherwise, L does not observe y^F (u^F). Let I^L_{RCS} be L's information set under this IS.
- Randomized observation-sharing: L observes her private information y^L and the signal s sent by F. However, L only observes y^F via a randomized channel z^y. We let L's corresponding information set be given by I^L_{ROS}.

Let Γ^L_{RCS} and Γ^L_{ROS} denote the sets of admissible strategies for L under the corresponding ISs. Let ω_0 be an Ω_0-valued random variable that defines the common exogenous uncertainty that affects both players' observations and/or costs. Each player seeks to minimize her expected cost, given by

$$J^L(\gamma^L, \gamma^F, \pi) = \mathbb{E}^{\gamma^L, \gamma^F, \pi}[c^L(\omega_0, u^L, u^F)], \tag{1}$$

$$J^F(\gamma^L, \gamma^F, \pi) = \mathbb{E}^{\gamma^L, \gamma^F, \pi}\left[c^F(\omega_0, u^L, u^F, s)\right], \tag{2}$$

for Borel-measurable functions $c^L : \Omega_0 \times \mathbb{U}^L \times \mathbb{U}^F \to \mathbb{R}_+$ and $c^F : \Omega_0 \times \mathbb{U}^L \times \mathbb{U}^F \times \mathbb{Y}^F \to \mathbb{R}_+$. Here, we use the notation $\mathbb{E}^{\gamma^L, \gamma^F, \pi}$ to denote the expectation with respect to $\mathbb{P}$ when actions of players and the signal s are induced by γ^L, γ^F and π, respectively. Next, we define the notion of an approximate Stackelberg equilibrium.

Definition 1 (ϵ-Stackelberg Equilibrium (SE)). *Given $\epsilon \geq 0$, $(\gamma^{L\star}, \gamma^{F\star}, \pi^\star)$ with L's IS I^L_{RCS}, constitutes an ϵ-SE, if*

$$J^L(\gamma^{L\star}, \gamma^{F\star}, \pi^\star) \leq \inf_{\gamma^L \in \Gamma^{RCS}} \sup_{(\gamma^F, \pi) \in R(\gamma^L)} J^L(\gamma^L, \gamma^F, \pi) + \epsilon,$$

where $R(\gamma^L)$ is defined as

$$R(\gamma^L) := \left\{ (\hat{\gamma}^F, \hat{\pi}) \in \Gamma^F \times \Pi \;\middle|\; J^F(\gamma^L, \hat{\gamma}^F, \hat{\pi}) = \inf_{(\gamma^F, \pi) \in \Gamma^F \times \Pi} J^F(\gamma^L, \gamma^F, \pi) \right\}. \tag{3}$$

This equilibrium concept corresponds to the pessimistic SE, where L selects a strategy that approximately optimizes her cost, accounting for the worst possible (for L) choice of F among those responses that optimize her cost.

Next, we introduce a notion of optimality for a strategy profile from L's vantage point.

Our approach to equilibrium characterization is indirect–we first calculate how L would ideally like to act and how she wants F to act. That is, we solve for an approximate leader-optimal strategy, defined next, from a decentralized control problem with static information structure I^L_{ROS}. Then, we devise an incentive strategy with the available dynamic information that utilizes the leader-optimal strategy. This incentive strategy is such that it induces the same response from F that L computes for F within her leader-optimal strategy, and it achieves the optimal performance.

Definition 2 (ϵ-Leader-Optimality). *Given $\epsilon \geq 0$, $(\gamma^{L\star}, \gamma^{F\star}, \pi^\star)$ with L's IS I^L_{ROS} constitutes an ϵ-leader-optimal solution, if*

$$J^L(\gamma^{L\star}, \gamma^{F\star}, \pi^\star) \leq \inf_{\gamma^L, \gamma^F, \pi \in \Gamma^{ROS} \times \Gamma^F \times \Pi} J^L(\gamma^L, \gamma^F, \pi) + \epsilon.$$

We call $\gamma^{L\star}$ with L's information structure I^L_{RCS} an ϵ-incentive equilibrium (IE), if there exists $(\gamma^{F\star}, \pi^\star)$ such that $(\gamma^{L\star}, \gamma^{F\star}, \pi^\star)$ is ϵ-leader-optimal and constitutes an ϵ-SE.

3. SIGNALING-BASED INCENTIVE DESIGN

Existence of an IE strategy has been established under perfect monitoring in Başar (1984), i.e., L observes y^F and u^F. Now, we turn our attention to the setting where the monitoring channel is random, modeled via the IS I^L_{RCS}. We show that a signaling-based IE strategy for L exists that leads to the revelation of y^F by F via F's honest signaling mechanism and induces the desired behavior in F. Throughout this section, we assume that L knows F's cost and also the joint distribution of exogenous random variables. Our result requires the following set of assumptions.

Assumption 1.

(i) $c^F(\omega_0, \cdot, \cdot, \cdot)$ *is jointly strictly convex for every* ω_0
(ii) $c^F(\omega_0, \cdot, \cdot, \cdot)$ *is continuously differentiable for every* ω_0
(iii) $\mathbb{P}\{z^y = z^u = 1\} > 0$
(iv) $c^F(\omega_0, \cdot, \cdot, s)$ *is radially unbounded for every* ω_0 *and* s *i.e.,* $c^F(\omega_0, u^L, u^F, s) \to \infty$ *if* $||u^L|| + ||u^F|| \to \infty$ *for every* ω_0 *and* s.

Theorem 1. *Consider $\mathcal{P}$ with I^L_{RCS} as L's IS. Let Assumption 1 hold. Let $\epsilon \geq 0$ and $(\gamma^{L\star}, \gamma^{F\star}, \pi^\star)$ constitute an ϵ-leader-optimal strategy profile with L's IS I^L_{ROS}, for which $\gamma^{L\star}$ is affine in s, and*

$$\mathbb{E}\left[c^F_{u^L}\left(\omega_0, u^{L\star}, u^{F\star}, s^\star\right)\middle| y^F\right] \neq 0 \quad \mathbb{P}\text{-}a.s. \tag{4}$$

where $s^\star \sim \pi^\star(\cdot|y^F) := \delta_{y^F}(\cdot)$, $u^{L\star} := \gamma^{L\star}(y^L, y^F)$, $u^{F\star} := \gamma^{F\star}(y^F)$, and $c^F_{u^L}$ denotes the partial derivative of c^F with respect to u^L. Then, there exists $\widetilde{\gamma}^{L\star} \in \Gamma^L_{RCS}$ for L, given by

$$\begin{aligned}\widetilde{\gamma}^{L\star}(y^L, y^F, u^F, s) &= z^y \gamma^{L\star}(y^L, y^F) + (1 - z^y)\gamma^{L\star}(y^L, s) \\ &\quad + z^y z^u Q^1(y^F, y^L)[u^F - \gamma^{F\star}(y^F)] \\ &\quad + z^y Q^2(y^F, y^L)[s - y^F],\end{aligned} \tag{5}$$

which constitutes an ϵ-IE strategy, for some Borel measurable functions Q^1 and Q^2.

We now provide insights into the result. First, the approximate IE strategy in (5) is affine in u^F and s (hence, continuous in u^F and s). It consists of three parts–the first two expressions in (5) correspond to an approximate leader-optimal solution which will be realized under the honest mechanism $\pi^\star$. The second and third parts can be viewed as a penalty for F deviating from L's desired control action and signal, respectively, where Q^1 and Q^2 correspond to the magnitude of these penalties for action and signal deviation, respectively. Owing to random monitoring, L does not always have access to u^F and y^F, and hence, these penalties can only be levied if z^y and z^u are 1. Our proof of this result is similar in spirit to that in Başar (1984), but differs from it in that L has access to u^F and y^F only via random monitoring channels. The incentive design in the signal plays a pivotal role in our design. Essentially, this incentivization allows L to elicit an honest observation-reporting behavior from F, leveraging probabilistic monitoring of y^F. Without the incentive to relay y_F truthfully, F may misreport, and that in turn can lead L to incur a possibly significant performance loss. In (5), L utilizes the uncertainty of F regarding the monitoring channel to induce F's revelation of y^F.

It is vital that F does not observe z^y and z^u although F knows their joint distributions. If F has access to the realizations of z^y and/or z^u, then L, in general, cannot find incentive strategies that attain the leader-optimal performance. In short, when $z^y = 0$ and/or $z^u = 0$, the terms containing Q^1 and Q^2 in the cost of F disappear. As a result, L loses her power to shape F's response using the dynamic information. In the same vein, we require Assumption 1 (iii) that keeps F guessing about the status of the monitoring channels. We remark that our incentive strategy in Theorem 1 requires L to accurately know c^F in order to correctly compute Q^1 and Q^2.

In the following, we provide an example of a quadratic Gaussian (QG) game where Assumption 1 holds and Theorem 1 applies. Consider a QG game with $y^L = \omega_0 + w^L$ and $y^F = \omega_0 + w^F$, where ω_0, w^L, w^F are independent standard normal random variables. Let the costs of the players be given by

$$c^L(\omega_0, u^L, u^F) = r^L(u^L)^2 + q^L(u^L + u^F + \omega_0)^2, \tag{6}$$

$$c^F(\omega_0, u^L, u^F) = r^F(u^F)^2 + q^F(u^F + u^L + \omega_0)^2, \tag{7}$$

where $r^F, r^L, q^F, q^L > 0$. We first compute $\gamma^{L\star}, \gamma^{F\star}$ and $\pi^\star$ such that $(\gamma^{L\star}, \gamma^{F\star}, \pi^\star)$ constitutes a leader-optimal solution. Let $\pi^\star(\cdot|y^F) = \delta_{y^F}(\cdot)$. Since, c^L is strictly convex, the unique optimal solution $\gamma^{L\star}, \gamma^{F\star}$ is linear and satisfies the following stationarity conditions: $\mathbb{P}$-a.s.,

$$\mathbb{E}\left[c^L_{u^L}(\omega_0, u^{L\star}, u^{F\star})\middle| y^L, y^F\right] = 0 \tag{8}$$

$$\mathbb{E}\left[c^L_{u^F}(\omega_0, u^{L\star}, u^{F\star})\middle| y^F\right] = 0, \tag{9}$$

where $u^{L\star} = \gamma^{L\star}(y^L, y^F) = \alpha^L y^L + \alpha^F y^F$ and $u^{F\star} = \gamma^{F\star}(y^F) = \beta^F y^F$, with

$$\alpha^L = -\frac{q^L}{3(r^L + q^L)}, \beta^F = -\frac{1}{2}, \alpha^F = \frac{q^L}{6(r^L + q^L)}. \tag{10}$$

Let $p = \mathbb{P}\{z^y = z^u = 1\}$ and $r = \mathbb{P}\{z^y = 1\}$. Since, c^F is strictly convex, the unique best response of F satisfies the following stationarity conditions $\mathbb{P}$-a.s.:

$$\mathbb{E}\Bigg[c^F_{u^F}(\omega_0, u^{L\star}, u^F) + pQ^1 c^F_{u^L}(\omega_0, u^L, u^{F\star}) + \big(pQ^2 + (1-r)\alpha^F\big)\, c^F_{u^L}(\omega_0, u^{L\star}, u^F)\Bigg| y^F\Bigg] = 0. \quad (11)$$

Hence, we can select Q^1 and Q^2 as follows:

$$Q^1 = -\frac{2r^F\beta^F}{p(\beta^F + \alpha^F + \frac{1}{2}(\alpha^L+1))},\; Q^2 = -\frac{1+(1-r)\alpha^F}{p}.$$

Both Q^1 and Q^2 are inversely proportional to p, the probability of monitoring. This has an intuitive meaning in that the lower the chance of monitoring, the more energy L must use in its strategy to appropriately incentivize F. The selection of Q^1 and Q^2 is not unique. In the next section, we show that Q^1 and Q^2 can be selected in a way that an IE strategy reduces the sensitivity of L's performance to perturbation on F's model.

4. DESIGNING INCENTIVE STRATEGY ROBUST TO FOLLOWER'S COST PARAMETERS

Our analysis in the previous section is premised on L knowing F's cost accurately–an assumption we relax and study incentive design that is robust to L's knowledge of F's cost. We follow in spirit the modeling framework of Cansever and Başar (1985b), but adapted to our setting with random monitoring. Specifically, we suppose that F's cost is parameterized by a parameter vector $\alpha \in \mathbb{R}^b$. Assume that L does not have access to α, but has a prior estimate $\alpha^\star$ referred to as α's nominal value. Here, we study the question whether L can design an incentive equilibrium that is not that *sensitive* to α (made precise later), so that L's lack of knowledge of α does not significantly impact her performance.

Let $\gamma^{F\star}_\alpha(\gamma^L)$ be F's unique best response strategy to L's choice of γ^L, when the cost parameter is α. We seek L's strategy γ^L for which

$$\frac{\partial^n \gamma^{F\star}_\alpha(\gamma^L)}{\partial \alpha^n}\bigg|_{\alpha^\star} = 0, \mathbb{P}-\text{a.s.},\; n = 1,\dots,N. \quad (12)$$

In effect, such a condition implies that F's response to L's strategy does not vary much in the neighborhood of $\alpha^\star$. Recall from the previous section that we utilized a leader-optimal strategy $\gamma^{L\star}$ to construct an IE $\tilde{\gamma}^{L\star}$. When applied to our example, we found that $\tilde{\gamma}^{L\star}$ was non-unique; one could choose Q_1 and Q_2 in a myriad of ways, each of which was an affine strategy in u^F that yielded leader-optimal performance. Our next result provides a construction of $\tilde{\gamma}^{L\star}$ for which $\gamma^{F\star}_\alpha(\tilde{\gamma}^{L\star})$ satisfies (12).

We require the following assumptions to state our result.

Assumption 2.

(i) *c^F is independent of s.*
(ii) *$c^F(\omega_0,\cdot,\cdot;\alpha)$ is jointly strictly convex for every ω_0 and $\alpha \in \mathbb{A}$.*
(iii) *$c^F(\omega_0,\cdot,\cdot;\alpha)$ is twice continuously differentiable for every ω_0 and $\alpha \in \mathbb{A}$.*
(iv) *$c^F(\omega_0,\cdot,\cdot;\alpha)$ is radially unbounded for every ω_0 and $\alpha \in \mathbb{A}$.*
(v) *$\mathbb{P}\{y^L \in \cdot | y^F\}$ has infinite support.*

Theorem 2. *Consider $\mathcal{P}$ with I^L_{RCS} as L's IS. Let Assumptions 1(iii) and 2 hold, together with the following three conditions.*

(i) *For any $\epsilon \geq 0$, let $(\gamma^{L\star}, \gamma^{F\star}, \pi^\star)$ constitute an ϵ-leader-optimal strategy profile with L's IS I^L_{ROS}, for which $\gamma^{L\star}$ is affine in s, and $\mathbb{P}$-a.s.,*

$$F(y^L, y^F) := \mathbb{E}\left[c^F_{u^L}\left(\omega_0, u^{L\star}, u^{F\star}; \alpha^\star\right)\Big| y^F, y^L\right] \neq 0. \quad (13)$$

(ii) *Given $\tilde{\gamma}^{L\star} \in \Gamma^L_{RCS}$ for L by (5), $\gamma^{F\star}_\alpha \in R_\alpha(\tilde{\gamma}^{L\star})$ is differentiable in α for all y^F.*
(iii) *There does not exist $k(y^F)$ such that*

$$F(y^L, y^F) = k(y^F) f_n(y^L, y^F), \qquad \forall n \in \mathbb{N} \quad (14)$$

where $c^F_{u^L,\alpha^n}$ ($c^F_{u^F,\alpha^n}$) denotes the partial derivative of c^F on u^L (u^F) and n-th order derivative on α, and

$$\begin{aligned} f_n&(y^L, y^F) \\ &:= F(y^L, y^F)\mathbb{E}\left[c^F_{u^F\alpha^n}\left(\omega_0, u^{L\star}, u^{F\star}; \alpha^\star\right)\Big| y^F\right] \\ &- G(y^F)\mathbb{E}\left[c^F_{u^L\alpha^n}\left(\omega_0, u^{L\star}, u^{F\star}; \alpha^\star\right)\Big| y^F\right], \end{aligned} \quad (15)$$

$$G(y^F) := \mathbb{E}\left[c^F_{u^F}\left(\omega_0, u^{L\star}, u^{F\star}; \alpha^\star\right)\Big| y^F\right]. \quad (16)$$

Then, $\tilde{\gamma}^{L\star}$ in (5) satisfies (12).

The art of this construction lies in the existence of Q^1 and Q^2 for which one can meaningfully set derivatives of F's best response strategy with respect to α to zero. For the set of assumptions we have made, F's response is indeed unique, when L chooses an affine strategy. The result implies that L can design an incentive strategy for which F's response varies "slowly" with α and yet yields the leader-optimal performance. Under appropriate assumptions of continuity, one then expects L's performance to vary slowly with α, thus mitigating the impact of L's lack of knowledge of α.

In the following, we provide an example QG game similar to that presented in the preceding section, where Assumption 2 holds and Theorem 2 applies. Let L's cost be given by (6) and F's cost be given by

$$c^F(\omega_0, u^L, u^F; \alpha) = r^F(u^F)^2 + q^F(\alpha - u^F - u^L - \omega_0)^2,$$

where α is unknown to L. However, L has access to the nominal value $\alpha^\star = 2$. Assume $q^F = r^F = 1/2$, $q^L = r^L = 1/3$, and ω_0 to be a unit-variance Gaussian with unit mean, while w_L and w_F are standard normal random variables. Again, we assume $y^L = \omega_0 + w^L$ and $y^F = \omega_0 + w^F$. Thus, we get $\mathbb{E}[\omega_0|y^F] = (y^F+1)/2$ and $\mathbb{E}[\omega_0|y^F, y^L] = (y^F+y^L+1)/3$. Along the lines of (8)–(10), for $\alpha = \alpha^\star$, the leader-optimal strategies are obtained as

$$\gamma^{L\star}(y^L, s) = -\frac{1}{6}y^L + \frac{1}{12}s + \frac{1}{12},\; \gamma^{F\star}_{\alpha^\star}(y^F) = -\frac{1}{2}y^F - \frac{1}{2}.$$

We can design Q^1 and Q^2 similar to that in the previous section; however, not all such designs make L's performance less sensitive to variations in α from its nominal value $\alpha^\star$. We start by selecting $Q^2 = (13+r)/12p$, where recall that $p = \mathbb{P}\{z^y = z^u = 1\}$ and $r = \mathbb{P}\{z^y = 1\}$. Following (11), we then infer that Q^1 must satisfy $\mathbb{P}$-a.s.,

$$\mathbb{E}[pQ^1(y^L, y^F)(2y^L - y^F - 25)|y^F] = -6(y^F + 5). \quad (17)$$

Any function of Q^1 of the form below satisfies (17):

$$Q^1(y^L, y^F) = \frac{-6(y^F+5)g_1(y^L,y^F)}{p\mathbb{E}[g_1(y^L,y^F)(2y^L-y^F-25)|y^F]}, \quad (18)$$

for some $g_1(y^L, y^F)$ such that the denominator is not zero $\mathbb{P}$-a.s. We find g_1 such that the first derivative of the best response strategy of F, $\gamma_\alpha^{F\star}$, with respect to α vanishes and hence, set

$$\mathbb{E}\left[g_1(y^L,y^F)\underbrace{(2y^L-7y^F-55)}_{:=f_1(y^L,y^F)}\middle| y^F\right] = 0 \quad \mathbb{P}\text{-a.s.} \quad (19)$$

where f_1 above is the same as defined in (15). Then,

$$g_1(y^L,y^F) = -108y^L + 152y^F - 69y^Fy^L + \frac{69}{2}(y^F)^2 + 48,$$

satisfies (19), and g_1 in (18) yields

$$Q^1(y^L,y^F) = \frac{-2(y^F+5)g_1(y^L,y^F)}{p(127(y^F)^2+623y^F-60)}. \quad (20)$$

Then, the best response of F for any α becomes

$$\begin{aligned}\gamma_\alpha^{F\star}(y^F) = \Big\{&\alpha(1+p\mathbb{E}[Q^1|y^F]) - p\mathbb{E}[Q^1\omega_0|y^F] - \mathbb{E}[\omega_0|y^F] \\ &- p\mathbb{E}[u^{L\star}Q^1|y^F] - \mathbb{E}[u^{L\star}|y^F] + p^2u^{F\star}\mathbb{E}[(Q^1)^2|y^F] \\ &+ pu^{F\star}\mathbb{E}[Q^1|y^F]\Big\}\Big/\Big\{\mathbb{E}[(2+p^2(Q^1)^2+2pQ^1|y^F]\Big\}.\end{aligned}$$

Following (19), $p\mathbb{E}[Q^1|y^F] = -1$ for any y^F, and hence, the impact of the uncertain parameter α vanishes from $\gamma_\alpha^{F\star}(y^F)$, removing the effect of L's lack of exact knowledge of α.

We now illustrate the importance of our sensitivity-reducing incentive design. With the α-insensitive F's response, the L's optimal cost is $J^{L\star} = 5/36 = 0.139$ for any α. Consider another candidate incentive strategy that L might design. For the same Q^2, suppose we select $g_1 \equiv 1$ and obtain $Q^1 = (y^F+5)/4p$. With $\alpha = 1.5 \neq \alpha^\star = 2$, we estimate $J_L = 0.164$ with 100K samples using sample average approximation. With the same samples, the estimate under the α-insensitive selection of Q^1 yields $J_L = 0.139$ that matches the optimal performance. In other words, the α-insensitive design compensates for L's flawed knowledge of α, while another candidate design leads to performance loss from this mistake.

5. INCENTIVE DESIGN WITH A CONVERGENT SEQUENCE OF FOLLOWER MODELS

We now turn our attention to incentive design when L may have inaccurate knowledge of F's model, by which we mean F's cost function c^F and the joint distribution $\mathcal{T}$ on ω_0, y^L, y^F. Consider a sequence of models $\mathcal{M}_n$ characterized by $\{c_n^F\}_n$ and $\{\mathcal{T}_n\}_n$ for $n \in \mathbb{N}$ such that they converge to the correct model c^F and $\mathcal{T}$, respectively, in some sense as $n \to \infty$; we define the notion of convergence precisely in the sequel. Let L design an IE strategy $\{\gamma_n^{L\star}\}_n$, computed based on the inaccurate model $\mathcal{M}_n$. We study how $\{\gamma_n^{L\star}\}_n$ fares on F, whose model is described by $\mathcal{M} = (c^F, \mathcal{T})$, and not $\mathcal{M}_n$. In essence, we ask: how robust is our incentive design strategy to incorrect models for F?

We make the following assumptions.

Assumption 3.

(i) c^F *is independent of* s.

(ii) $c_n^F(\omega_0,\cdot,\cdot)$ *and* $c^F(\omega_0,\cdot,\cdot)$ *are strictly convex and continuously differentiable for every* ω_0 *and* $n \in \mathbb{N}$.

(iii) c_n^F, c^F, c^L *are uniformly bounded by a constant* M *for* $M < \infty$ *for every* $n \in \mathbb{N}$.

(iv) $c_n^F(\omega_0, u^L, u^F)$ *converges to* $c^F(\omega_0, u^L, u^F)$ *uniformly in* $u^F \in \mathbb{U}^F$ *and* $u^L \in \mathbb{U}^L$ *as* $n \to \infty$.

(v) $\mathcal{T}_n$ *converges to* $\mathcal{T}$ *in the total variation metric as* $n \to \infty$ *for every* y^F*, i.e.,*

$$\|\mathcal{T}_n - \mathcal{T}\|_{\mathsf{TV}} := 2 \sup_{A\in\mathcal{B}(\Omega_0\times\mathbb{Y}^L\times\mathbb{Y}^F)} |\mathcal{T}_n(A) - \mathcal{T}(A)| \to 0$$

as $n \to \infty$*, where* $\mathcal{B}(\Omega_0\times\mathbb{Y}^L\times\mathbb{Y}^F)$ *is the Borel* σ*-field on* $\Omega_0\times\mathbb{Y}^L\times\mathbb{Y}^F$.

(vi) $\mathbb{U}^F$ *is compact.*

Under the assumption of strict joint convexity of c_n^F, L can design IEs $\tilde{\gamma}_n^{L\star}$'s for the incorrect models $\mathcal{M}_n$ that are affine in s and u^F, following the same recipe as presented in Theorem 1. For these affine incentive strategies, F's response to $\tilde{\gamma}_n^{L\star}$ becomes unique, following the affine structure of L's strategy and strict joint convexity of c_n^F in u^L and u^F. Call this response $R(\tilde{\gamma}_n^{L\star}; \mathcal{M}_n)$. Similarly, define $R(\tilde{\gamma}_n^{L\star}; \mathcal{M})$.

Theorem 3. *Consider* $\mathcal{P}$ *with* I_{RCS}^L *as* L*'s IS. Suppose Assumptions 1(iii), 2(i) and 3 hold. Let* $\epsilon \geq 0$ *and* $(\gamma_n^{L\star}, \gamma_n^{F\star}, \pi^\star)$ *constitute an* ϵ*-leader-optimal strategy with* I_{ROS}^L *as* L*'s IS under the incorrect model* $\mathcal{T}_n$*, for which* $\gamma_n^{L\star}$ *is affine in* s*, and*

$$\mathbb{E}^{\nu_n}\left[\nabla_{u^L} c_n^F\left(\omega_0, \gamma_n^{L\star}(y^L,y^F), \gamma_n^{F\star}(y^F)\right)\middle| y^F\right] \neq 0 \quad \mathbb{P}\text{-}a.s.$$

for any $n \in \mathbb{N}$ *where* ν_n *is the joint conditional distribution of* ω_0, y^L *given* y^F *under* $\mathcal{T}_n$*. Consider a sequence of strategies* $\{\tilde{\gamma}_n^{L\star}\}_n \subset \Gamma_{RCS}^L$ *for* L*, given by*

$$\begin{aligned}\tilde{\gamma}_n^{L\star}(y^L,y^F,u^F,s) &= z^y\gamma_n^{L\star}(y^L,y^F) + (1-z^y)\gamma_n^{L\star}(y^L,s) \\ &\quad + z^yz^uQ_n^1(y^F,y^L)[u^F - \gamma_n^{F\star}(y^F)] \\ &\quad + z^yQ^2(y^F,y^L)[s-y^F], \qquad (21)\end{aligned}$$

that constitutes an ϵ*-IE for* $\mathcal{P}$ *for the incorrect models* $\mathcal{M}_n$*. Assume that* $R(\tilde{\gamma}_n^{L\star};\mathcal{M})$ *and* $R(\tilde{\gamma}_n^{L\star};\mathcal{M}_n)$ *admit unique limit points as* $n \to \infty$*. Then, the two limit points are the same and* $\{\tilde{\gamma}_n^{L\star}\}$ *attains a performance* ϵ*-close to a performance of IE for* $\mathcal{P}$ *for the correct model* $\mathcal{M}$ *as* $n \to \infty$.

To prove Theorem 3, we first argue that the leader-optimal performance is continuous in the models. Then, the rest follows from the construction of the incentive strategy from the leader-optimal strategy, and the continuity of that map in F's models. We remark that our indirect equilibrium characterization technique thus plays a vital role in proving the requisite robustness results that are challenging to establish for general Stackelberg games.

Assumption 3(v) is only needed to guarantee that F's response strategies $R(\gamma^L;\mathcal{M})$ and $R(\gamma^L;\mathcal{M}_n)$ exist when L plays a strategy that is affine in u^F. The validity of our assumption that $R(\tilde{\gamma}_n^{L\star};\mathcal{M})$ and $R(\tilde{\gamma}_n^{L\star};\mathcal{M}_n)$ admit unique limit points as $n \to \infty$ requires further analysis; we plan to address it in future research. Moreover, recall that in Assumption 3(iv), we require that the sequence of incorrect distributions $\{\mathcal{T}_n\}$ converges to the true distribution $\mathcal{T}$ in total variation metric. Such convergence can be demanding in various data-driven applications; one would ideally impose weak convergence. While relaxation of this assumption is part of our ongoing efforts, we provide two concrete examples that meet the convergence criterion in total variation.

• Suppose that $y^L = g^L(\omega_0) + w^L$ and $y^F = g^F(\omega_0) + w^F$ for some Borel measurable functions g^L and g^F, where $w^L \sim \mu^L$, $w^F \sim \mu^F$, and $\omega_0 \sim \mathbb{P}^0$. Let w^L be independent of w^F. Suppose that g^L and g^F are known to L, μ^L, μ^F and $\mathbb{P}^0$ are only known approximately to L. If the densities μ_n^L and μ_n^F converge in distribution to μ^L and μ^F, respectively, then Scheffe's theorem guarantees that they converge in total variation as well. In turn, $\mathcal{T}_n$ converges to $\mathcal{T}$ in total variation, as required by Assumption 3.

• Suppose that $y^L = g^L(\omega_0 + w^L)$ and $y^F = g^F(\omega_0 + w^F)$ where $w^L \sim \mu^L$ and $w^F \sim \mu^F$. Let w^L be independent of w^F. Assume that g^L and g^F are unknown to L but μ^L, μ^F, and $\mathbb{P}^0$ are known to L. Suppose $\mathbb{Y}^L$ and $\mathbb{Y}^F$ are finite. If incorrect models g_n^L and g_n^F converge pointwise to g^L and g^F, then for any continuous bounded function f,

$$\begin{aligned}
&\left|\int f(\omega_0, y^L, y^F)d\mathcal{T}_n - \int f(\omega_0, y^L, y^F)d\mathcal{T}\right| \\
&= \bigg| \int f\left(\omega_0, g_n^L(\omega_0 + w^L), g_n^F(\omega_0 + w^F)\right) d\mathbb{P}^0 d\mu^L d\mu^F \\
&- \int f(\omega_0, g^L(\omega_0 + w^L), g^F(\omega_0 + w^F))d\mathbb{P}^0 d\mu^L d\mu^F \bigg|
\end{aligned}$$

which converges to zero by the dominated convergence theorem as $n \to \infty$. By considering a dense subset of convergence determining functions f of the set of all continuous and bounded functions, we infer that $\mathcal{T}_n$ converges weakly to $\mathcal{T}$. Since $\mathbb{Y}^L$ and $\mathbb{Y}^F$ are finite, they converge in total variation, verifying Assumption 3.

6. CONCLUSIONS

We have studied a class of incentive design problems between a leader (L) and a follower (F), where we allow L to observe F-relevant variables with noise. For such problems, for the case when L has access to the correct models, we have established the existence of a signaling-based incentive equilibrium strategy that induces an honest mechanism on F, leading to a desired response from L's standpoint. Further for the scenario where L does not have access to F's cost structure, we have established the existence of an incentive equilibrium strategy that reduces the sensitivity of her performance to the unknown parameter in F's cost. We have finally established robustness of incentive equilibrium strategies to incorrect models, that included F's cost structures and distributions of cost-relevant random variables. Finally, we have presented several examples to demonstrate our main results.

We are interested in pursuing a number of interesting research directions. First among these directions concerns learning in incentive design, where L can repeatedly interact with F and learn their cost structures. We also want to study multi-stage dynamic versions of incentive design problems (with possible state evolutions for both parties), and these players can be risk-sensitive decision-makers.

REFERENCES

Başar, T. (1989). Stochastic incentive problems with partial dynamic information and multiple levels of hierarchy. *European J. Political Economy (special issue on "Economic Design")*, 5, 203–217.

Başar, T. and Olsder, G. (1999). *Dynamic Noncooperative Game Theory*. SIAM Classics in Applied Mathematics, Philadelphia, PA.

Başar, T. (1983). Performance bounds for hierarchical systems under partial dynamic information. *Journal of Optimization Theory and Applications*, 39(1), 67–87.

Başar, T. (1984). Affine incentive schemes for stochastic systems with dynamic information. *SIAM Journal on Control and Optimization*, 22(2), 199–210.

Başar, T. and Bernhard, P. (2008). *H-infinity Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*. Springer Science & Business Media.

Cansever, D. and Başar, T. (1985a). On stochastic incentive control problems with partial dynamic information. *Systems and Control Letters*, 6(1), 69–75.

Cansever, D. and Başar, T. (1985b). Optimum/near-optimum incentive policies for stochastic decision problems involving parametric uncertainty. *Automatica*, 21(5), 575–584.

Dasgupta, P., Hammond, P., and Maskin, E. (1979). The implementation of social choice rules: Some general results on incentive compatibility. *The Review of Economic Studies*, 46(2), 185–216.

Fudenberg, D. and Tirole, J. (1991). *Game Theory*. MIT press.

Groves, T. (1973). Incentives in teams. *Econometrica: Journal of the Econometric Society*, 617–631.

Groves, T. and Loeb, M. (1979). Incentives in a divisionalized firm. *Management Science*, 25(3), 221–230.

Hansen, L.P. and Sargent, T.J. (2001). Robust control and model uncertainty. *American Economic Review*, 91(2), 60–66.

Ho, Y.C., Luh, P., and Olsder, G.J. (1982). A control-theoretic view on incentives. *Automatica*, 18(2), 167–179.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6), 2590–2615.

Kara, A.D. and Yüksel, S. (2019). Robustness to incorrect priors in partially observed stochastic control. *SIAM Journal on Control and Optimization*, 57(3), 1929–1964.

Kara, A.D. and Yüksel, S. (2020). Robustness to incorrect system models in stochastic control. *SIAM Journal on Control and Optimization*, 58(2), 1144–1182.

Khalil, I., Doyle, J.C., and Glover, K. (1996). *Robust and Optimal Control*. Prentice Hall.

Wiesemann, W., Kuhn, D., and Rustem, B. (2013). Robust Markov decision processes. *Mathematics of Operations Research*, 38(1), 153–183.

Yüksel, S. and Linder, T. (2012). Optimization and convergence of observation channels in stochastic control. *SIAM Journal on Control and Optimization*, 50(2), 864–887.

Zheng, Y. and Başar, T. (1982). Existence and derivation of optimal affine incentive schemes for Stackelberg games with partial information: A geometric approach. *International Journal of Control*, 35(6), 997–1011.

Zheng, Y., Başar, T., and Cruz, J.B. (1984). Stackelberg strategies and incentives in multiperson deterministic decision problems. *IEEE Transactions on Systems, Man, and Cybernetics*, (1), 10–24.