Information Design in Concave Games *

Alex Smolin, Takuro Yamashita

August 15, 2022

Abstract

We study information design in games with a continuum of actions such that the players' payoffs are concave in their own actions. A designer chooses an information structure—a joint distribution of a state and a private signal of each player—and evaluates it according to the designer's expected payoff under the equilibrium play in the induced Bayesian game. We show an information structure is designer optimal whenever it induces the equilibrium play that can be implemented by an incentive contract in an auxiliary principal-agent problem with a single agent who observes the state and controls all actions.

We use this result to characterize optimal information structures in a variety of settings, including price competition, first-order Bayesian persuasion, and venture capital fundraising. If the state is normally distributed and the payoffs are quadratic, then in many cases Gaussian information structures are optimal. Fully informing a subset of players can also be optimal and robustly so, for all state distributions.

Keywords: Bayesian persuasion, concave games, first-order approach, Gaussian information structures, information design, selective informing, weak duality.

*Smolin: Toulouse School of Economics, University of Toulouse Capitole and CEPR, alexey.v.smolin@gmail.com. Yamashita: Toulouse School of Economics, University of Toulouse Capitole, tytakuroy@gmail.com. The authors acknowledge funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d'Avenir) program (grant ANR-17-EURE-0010). Takuro Yamashita acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 714693).

1 Introduction

Information flows are vital for the economy and are increasingly controlled by technological giants who decide how to display products to consumers, what prices to recommend to sellers, how to compose news feeds, what consumer characteristics to reveal to bidding advertisers, and so on. These choices guide and structure strategic interactions at an unprecedented scale. To understand these choices, one needs to understand optimal information control, which was recently formalized in the field of Bayesian persuasion or information design (Bergemann and Morris (2019), Kamenica (2019)). Existing methodology enables the designer's problem to be posed and to be solved in important special cases, such as those with a binary state, binary actions, a single player, or for special classes of information structures. However, to date, there exist no solutions for large-scale multiplayer games, with a continuum of states and actions and nonlinear payoffs. In this paper, we develop a universal solution method to find unconstrained-optimal information structures in such large-scale games.

Specifically, we introduce and study information design in *concave* games of incomplete information, i.e., games in which each player's action can take values in a convex set and the payoff of each player is strictly concave in his action, for any state and actions of other players. The information designer takes the players' payoffs and a prior state distribution as given but can design an arbitrary information structure that specifies the joint distribution of the state and private signals of the players. The designer anticipates equilibrium play. In concave games, the best response of each player to his signal can be found by means of a first-order optimality condition, and the joint equilibrium behavior is determined by a system of such conditions. The induced distribution of state and actions is assessed by the designer according to her expected payoffs. The goal is to find an information structure that is optimal for the designer.

In a problem of this scale, a direct search for an optimal information structure is intractable because of the sheer number of optimization parameters and equilibrium constraints. Instead, we develop a solution method to check and certify the optimality of candidate information structures. To do so, we observe a close connection between an information-design problem and an optimal transport problem, and construct a dual problem applying duality results in the optimal transport theory. In the dual problem, a principal faces a single agent who fully controls all actions and perfectly observes the state. The principal chooses an incentive contract that directly affects the agent's payoff: the payoff is a weighted sum of the designer's payoff and the marginal payoffs of the players in the information-design problem; the weights can depend only on individual actions and are specified by the contract. The principal anticipates the agent's best response and chooses the contract to minimize the expected agent's payoff; hence, we can interpret the dual problem as adversarial contracting.

The dual problem is important because its optimal value places an upper bound on the optimal value of the information-design problem and, as such, on the value of information control (Theorem 1). It in turn enables the optimality of any given information structure to be certified: if a given state-action distribution can be implemented by some information structure in the information-design problem and by some contract in the adversarial-contracting problem, then these information structure and contract solve the respective problems (Proposition 1). Moreover, the adversarial-contracting problem can suggest the shape of optimal information provision. We discuss the general properties of certifiably optimal information structures in Section 3.4 and the scope of the certification method in Section 3.5.

The certification solution method can be applied to any concave game. However, its application is particularly simple in games with quadratic payoffs because in such games the players' marginal payoffs are linear. Consequently, we show that in such games an optimal information structure can often be certified by linear contracts.

First, in Section 4, we use this method to solve general information-design problems in which the state is distributed according to a multivariate normal distribution and the designer's and players' payoffs are quadratic in actions and the state. We provide conditions under which an optimal information structure informs each player about a linear combination of state components and explicitly derive its optimal coefficients (Theorem 2). Under these conditions, the optimal information structure is Gaussian, i.e., the private signals, as well as the induced actions, are jointly normally distributed.

In Section 4.1, we apply Theorem 2 to characterize an optimal information structure in a differentiated Bertrand duopoly with linear demand curves and uncertain demand shocks. We show that information structures that maximize a weighted average of the consumer and producer surplus induce normally distributed prices that are linear in demand shocks and correlated between firms. If the weight given to consumer surplus is low, then the optimal information structure induces coordinated pricing; if it is high, the pricing is anticoordinated. The shift between these two modes is discontinuous.

Second, in Section 4.2 we study a first-order persuasion setting in which each player aims to make the best prediction of a common one-dimensional state and the designer aims to polarize players' predictions. We show that there co-exist two qualitatively different classes of optimal information structures. In the first class, the designer fully informs a fraction of players and leaves all other players uninformed. In the second class, the designer chooses a symmetric Gaussian information structure with noises finely tuned across players. Both information structures induce the same aggregate prediction, deterministic in the state, but differ in their egalitarian properties.

Third, in Section 4.3 we apply our solution method to an investment game in which each player decides how much to invest in a project of uncertain quality. The average investment profitability increases in the project quality but decreases in total investment. We characterize information structures that maximize investment profits. Once again, two different classes of solutions co-exist. One solution is to fully inform a single player while leaving all other players uninformed. This simple information structure is optimal irrespectively of the number of players, prior state distribution, and other payoff parameters. Another solution, which exists if the state is normally distributed, is a symmetric and Gaussian information structure. Relative to no information or full information, the optimal information control avoids dissipation of investment rents as the number of players increases.

In Section 4.4, we discuss several features of optimal information structures that we uncovered throughout our analysis: the multiplicity of solutions, the presence of extraneous noise, the role of state dimensionality, and the limits at large economies. Section 5 concludes.

Related Literature The literature on Bayesian persuasion or information design covers the analysis of information control in decision problems (Rayo and Segal (2010), Kamenica and Gentzkow (2011)) and multiplayer games (Bergemann and Morris (2016), Taneva (2019)) and constitutes a vibrant field of research.

Our work is based on the duality methodology. This methodology was applied in the past to solve information-design problems but primarily those with a single receiver, be it a single player or a team.[1] Dworczak and Martini (2019) solve a class of problems in which the receiver cares only about the first moment of a state. Dworczak and Kolotilin (2019) extend this analysis to higher moments and beyond.[2] Malamud and Schrimpf (2021) and Cieslak, Malamud, and Schrimpf (2021) establish some general properties of optimal information structures building on optimal-transportation duality. Kolotilin

[1]The duality methodology is frequently used to solve optimization problems in many different fields. In economics, it has been applied to consumer theory (Krishna and Sonnenschein (1990)), matching problems (Galichon (2018)), multidimensional mechanism design (Cai et al. (2019)), and robust mechanism design (Carroll (2017), Du (2018), Brooks and Du (2020)), among others.

[2]These authors pose the information-design problem in the space of belief distributions rather than information structures; as a result, their dual problems are qualitatively different from ours.

(2017) studies the persuasion of a receiver with uncertain preferences; his formalism is closest to ours, and his problem may be viewed as an instance of our setting with a single player and a one-dimensional state. Many of these works spend much effort on establishing strong duality, whereas we stress that weak duality is sufficient to apply our solution method. None of these works study games, formulate the dual problem as adversarial contracting, or establish the optimality of Gaussian information structures.

We are not familiar with any previous or concurrent work that characterizes an unconstrained-optimal information structure in a fixed game with many players and a continuum of states and actions. For example, Galperti and Perego (2018) and Galperti, Levkun, and Perego (2021) apply duality methodology to study information design in games with general payoffs but with finitely many actions. As such, their incentive constraints feature utility comparisons of all possible deviations, as inequalities. Without any structure on payoffs, the problem is unwieldy and the authors focus on the analysis and interpretation of optimal dual variables rather than on the search for optimal information structures. In contrast, we study continuous games and formulate the incentive constraints as local first-order conditions, as equalities. This is the key step toward tractability and translates into qualitatively different problems, primal or dual.

The strand of economic literature that perhaps comes closest to the understanding of optimal information structures in large-scale games is one that studies optimal parameters of symmetric Gaussian information structures in symmetric games with a normally distributed state (Angeletos and Pavan (2007), Angeletos and Pavan (2009), Bergemann and Morris (2013), Bergemann, Heumann, and Morris (2015), Ui (2020), Bergemann, Heumann, and Morris (2021)). In general, an optimal information structure does not have to be symmetric or Gaussian, even if the game is symmetric and the state is normally distributed. However, our results in Section 4 suggest that symmetric Gaussian information structures may indeed be optimal in some of those settings and may be possibly certifiable by our solution method.[3]

[3] As such, our results conform with the findings of Tamura (2018) who showed the optimality of Gaussian information structures in a setting with a normally-distributed state, a single receiver, and quadratic payoffs, building directly on statistical properties of covariance matrices of posterior expectations. However, his analysis is not directly extendable to multiplayer games: the relevant statistical properties are not clear and, more importantly, the player's state expectation is generally not a sufficient statistic for his best response.

2 Model

Payoffs There are N players indexed by i, $1 \leq N < \infty$, and an information designer. Each player is to take an action $a_i \in A_i = \mathbb{R}$.[4] We denote by A the set of action profiles $A = \times_i A_i$ and write (a_i, a_{-i}) for an action profile when focusing on its i-th component.

A state ω is distributed over a possibly infinite set $\Omega \subseteq \mathbb{R}^K$, $K \geq 1$, according to a prior distribution $\mu_0 \in \Delta(\Omega)$. An action profile $a = (a_1, \ldots, a_N) \in A$ together with the state determine the payoffs of player i according to the payoff function

$$u_i : A \times \Omega \to \mathbb{R}. \tag{1}$$

The primitives $((A_i, u_i)_{i=1}^N, \mu_0)$ constitute a *basic game*. The designer's payoff given action profile a at state ω is described by the payoff function

$$v : A \times \Omega \to \mathbb{R}. \tag{2}$$

Information The players and the designer start with a common prior belief about the state ω that coincides with the prior distribution μ_0. The designer can provide additional information to players by choosing an information structure $\mathcal{I} = (S, \pi)$ that consists of a signal set $S = \times_i S_i$, which is a Polish space, and a likelihood function $\pi \in \Delta(\Omega \times S)$ that has μ_0 as its state marginal distribution. This information structure prescribes the sets of private signals the players can observe and, through the likelihood function, their informational content. The information structure governs information about the state and coordinates the players' actions.

The timing is as follows. First, the designer chooses an information structure $\mathcal{I}$. Second, the state ω and the signal profile $s = (s_1, \ldots, s_N)$ are realized according to the chosen information structure. Finally, each player privately observes his corresponding signal s_i and chooses an action a_i.

The basic game together with the information structure chosen by the designer determine a Bayesian game of incomplete information. In that game, each player's behavior is described by a strategy that maps any received signal to a possibly random action, $\sigma_i : S_i \to \Delta(A_i)$, and we consider as a solution concept a Bayes Nash equilibrium that prescribes the players to form their beliefs via Bayes' rule and to act to maximize their expected payoffs.

[4] Our methodology can easily be extended to the case of multidimensional actions at the expense of additional notation. The methodology is extended to the case of bounded actions in Appendix B.4.

Definition 1. (Bayes Nash Equilibrium) *For a given information structure $\mathcal{I}$, a strategy profile $\sigma = (\sigma_1, \ldots, \sigma_N)$ constitutes a Bayes Nash equilibrium if*

$$\mathbb{E}_{\mathcal{I},\sigma_i,\sigma_{-i}}[u_i(a_i, a_{-i}, \omega)] \geq \mathbb{E}_{\mathcal{I},\sigma_i',\sigma_{-i}}[u_i(a_i', a_{-i}, \omega)] \tag{3}$$

for all i and $\sigma_i' : S_i \to \Delta(A_i)$, where $\mathbb{E}_{\mathcal{I},\sigma_i,\sigma_{-i}}[\cdot]$ denotes a mathematical expectation given information structure $\mathcal{I}$ and strategy profile (σ_i, σ_{-i}).

An information-design problem consists of choosing an information structure that maximizes the expected payoff of the designer without placing any additional restrictions on the sets of signals or the likelihood function. Formally, each strategy profile determines a conditional distribution over the action profiles in each state $\alpha : \Omega \to \Delta(A)$, which we call an *allocation rule*. Each allocation rule together with the prior state distribution μ_0 and the payoff function (2) determines the designer's expected payoff. Therefore, the value of any information structure can be determined as the maximal designer's expected payoff that can arise in equilibrium of the induced Bayesian game.[5] The solution to the information-design problem is an information structure such that there does not exist an information structure with a strictly higher value.

In what follows, we analyze a specific class of basic games in which each player's payoff is everywhere concave in his own action:

Assumption 1. (Concave Payoffs) *For all $i = 1, \ldots, N$, $\omega \in \Omega$, and $a_{-i} \in A_{-i}$, $u_i(a_i, a_{-i}, \omega)$ is continuously differentiable in a_i, strictly concave in a_i, and obtains its maximum at some finite value.*

Assumption 1 resembles the assumption imposed in the seminal work on concave games of complete information by Rosen (1965) but is weaker in that we do not require differentiability or continuity of the player's payoff with respect to the opponents' actions. Assumption 1 is standard in applied economic models with fixed information structures because it simplifies the characterization of equilibrium behavior: the best response of each player at any belief over the state and actions of other players can be found via a first-order condition, and an equilibrium can be characterized by a system of such conditions, one for each player's signal.[6] We show that the same assumption facilitates the analysis of the information-design problem in which the information structure is the object of design, for arbitrary designer's payoffs. We call a basic game in which

[5] If a Bayesian game allows for multiple equilibria, the designer can choose the one she prefers. If no equilibrium exists, the value is undefined.

[6] Our analysis can be applied in any game in which this property holds.

Assumption 1 is satisfied a *concave game*. We call an information-design problem in a concave game a *concave information-design problem*.

3 General Analysis

3.1 Equilibrium Conditions. Primal Problem.

We begin by simplifying the equilibrium conditions (3) utilizing the special payoff structure of concave games. Consider the choice of player i. In equilibrium, for any admissible belief over the state and opponents' actions $\nu \in \Delta(A_{-i} \times \Omega)$, the player must take a best-response action $a_i^*(\nu)$ that maximizes his expected payoff $\mathbb{E}_\nu\left[u_i(a_i, a_{-i}, \omega)\right]$. By Assumption 1, this payoff is continuously differentiable and strictly concave in a_i since it is a convex combination of continuously differentiable and strictly concave functions. Thus, $a_i^*(\nu)$ is a unique solution to a first-order condition. Denote the partial derivative of the player's payoff function by

$$\dot{u}_i(a, \omega) \triangleq \frac{\partial u_i(a, \omega)}{\partial a_i}. \tag{4}$$

Assumption 1 implies that $\dot{u}_i(a, \omega)$ exists, is continuous, and strictly decreases in a_i everywhere. By Leibniz integral rule, the first-order condition that identifies the best response $a_i^*(\nu)$ can be written as

$$\frac{\partial \mathbb{E}_\nu\left[u_i(a_i^*, a_{-i}, \omega)\right]}{\partial a_i} = \mathbb{E}_\nu\left[\frac{\partial u_i(a_i^*, a_{-i}, \omega)}{\partial a_i}\right] = \mathbb{E}_\nu\left[\dot{u}_i(a_i^*, a_{-i}, \omega)\right] = 0. \tag{5}$$

To further simplify the information-design problem, we appeal to the revelation principle (Myerson (1983), Bergemann and Morris (2016)) and focus, without loss of generality, on *direct* information structures that inform each player about a recommended action $S = A$ and induce posterior beliefs such that all players are obedient, i.e., are willing to follow the recommendations. Each direct information structure corresponds to a measure $\pi \in \Delta(A \times \Omega)$ that has μ_0 as its state marginal.

These two simplifications enable us to formulate an information-design problem as:

$$V^P \triangleq \sup_{\pi \in \Delta(A\times\Omega)} \int_{A\times\Omega} v(a,\omega)\mathrm{d}\pi \tag{6}$$

$$\text{s.t.} \int_{A_i'\times A_{-i}\times\Omega} \dot{u}_i(a,\omega)\mathrm{d}\pi = 0 \quad \forall\, i=1,\dots,N, \text{measurable}\, A_i' \subseteq A_i, \tag{7}$$

$$\int_{A\times\Omega'} \mathrm{d}\pi = \int_{\Omega'} \mathrm{d}\mu_0 \quad \forall\, \text{measurable}\, \Omega' \subseteq \Omega. \tag{8}$$

Constraints (7) capture players' obedience and must hold at all measurable subsets $A_i' \subseteq A_i$. They effectively require that for each player i, the linear projection of π on A_i weighted by the marginal utilities is equal to zero measure. This is a proper formulation of first-order conditions (5) in light of a possible continuum of recommended actions. Constraints (8) capture Bayes' plausibility and, likewise, require that the linear projection of π on Ω equals the prior distribution μ_0.

Problem (6) is linear in π. In the spirit of linear programming, we view it as a *primal problem* and call any $\pi \in \Delta(A \times \Omega)$ a *primal measure*. If a primal measure satisfies the constraints of the primal problem, then we call that measure *implementable by information* and call the corresponding value of the objective, V^P, a feasible primal value.

3.2 Dual Problem. Adversarial Contracting.

In this section, we develop a *dual problem*, applying duality from optimal transport theory (see, for example, Villani (2003)). The significance of this problem, and dual problems in general, comes from its ability to provide an upper bound on the information designer's payoffs and, ultimately, certify a solution. The dual problem to (6) is as follows:

$$V^D \triangleq \inf_{\lambda \in \times_i L(A_i), \gamma \in L(\Omega)} \int_{\Omega} \gamma(\omega)\mathrm{d}\mu_0 \tag{9}$$

$$\text{s.t.} \sum_{i=1}^{N} \lambda_i(a_i)\dot{u}_i(a,\omega) + \gamma(\omega) \geq v(a,\omega)\ \forall\, a \in A, \omega \in \Omega,$$

where $L(X)$ denotes the space of measurable real-valued functions on X.[7] The minimization arguments, the dual variables (λ, γ), represent the Lagrange multipliers associated with the primal incentive constraints (7) and the feasibility constraints (8), respectively.

[7] Throughout the paper, we adopt the convention that the value of an integral is set to $+\infty$ whenever the underlying function is not integrable.

Thus, they have a clear economic interpretation: $\lambda_i(a_i)$ measures the marginal benefit for the information designer from pushing action a_i upward, whereas $\gamma(\omega)$ measures the marginal benefit from increasing the prior probability of state ω.

The dual problem (9) can be simplified by reducing the number of optimization parameters and admits an intuitive economic interpretation which, to the best of our knowledge, is novel in the literature on information design. To this end, observe that the objective in (9) is additive separable in $\gamma(\omega)$ and that the constraints at different states ω are linked only through variables λ. Hence, for any λ and ω, an optimal $\gamma(\omega)$ must be minimized across the values above the lower bounds imposed by the dual constraints and hence must be equal to:

$$\gamma^*(\lambda, \omega) = \sup_{a \in A} u^\lambda(a, \omega),$$

where u^λ is a *dual payoff* defined as

$$u^\lambda(a, \omega) \triangleq v(a, \omega) - \sum_{i=1}^{N} \lambda_i(a_i)\dot{u}_i(a, \omega). \tag{10}$$

As a result, the dual problem (9) can be restated as

$$V^D = \inf_{\lambda \in \times_i L(A_i)} \int_\Omega \sup_{a \in A} u^\lambda(a, \omega) \mathrm{d}\mu_0 = \inf_{\lambda \in \times_i L(A_i)} \mathbb{E}_{\mu_0}[\sup_{a \in A} u^\lambda(a, \omega)], \tag{11}$$

Problem (11) can be interpreted as *adversarial contracting* between a principal and a single *agent*.[8] The agent perfectly observes the state and alone controls the whole action profile. To influence the agent's behavior, the principal chooses an incentive contract λ that consists of N functions $\lambda_i(a_i)$ and modifies the agent's payoff according to (10), i.e., the ith component of the contract links the agent's utility to $\dot{u}_i(a, \omega)$. The timing of the adversarial contracting is as follows. First, the designer chooses a contract λ. Second, the state ω is realized and is observed by the agent. Finally, the agent chooses an action profile $a \in A$. If the best responses exist at all states and induce the joint action-state measure $\pi(a, \omega)$, then we say that λ *implements* π *by incentives* so that π *is implementable by incentives*. Whenever the best response does not exist, the interim payoff is assessed as a supremum. The contracting is adversarial in that the designer aims to minimize the agent's expected payoff; equivalently, the game between the designer and the agent is zero sum. [9]

[8] The "agent" is not to be confused with the "player" of the information-design problem.

[9] Note that in adversarial contracting, a benchmark feasible contract is a null contract $\lambda_1(a_1) \equiv$

3.3 Weak Duality. Optimality Certification.

In this section, we show that the optimal values of the primal and dual problems are tightly connected. We demonstrate how this relationship can be used to solve concave information-design problems.

Theorem 1. (Weak Duality) $V^P \leq V^D$.

Proof. Take any dual variables (λ, γ) that satisfy the constraints of dual problem (9). Take any measure π that satisfies the constraints of primal problem (6). Integrating both sides of the dual constraints over $a \in A$ and $\omega \in \Omega$ against measure π yields:

$$\int_{A\times\Omega} v(a,\omega)\mathrm{d}\pi \leq \int_{A\times\Omega} \sum_{i=1}^{N} \lambda_i(a_i)\dot{u}_i(a,\omega)\mathrm{d}\pi + \int_{A\times\Omega} \gamma(\omega)\mathrm{d}\pi = \int_{\Omega} \gamma(\omega)\mathrm{d}\mu_0, \tag{12}$$

where the equality follows because π satisfies the primal constraints. The left-hand side of (12) is the value of the primal problem given measure π. At the same time, the right-hand side of (12) is the value of the dual problem given dual variables (λ, γ). As the inequality (12) holds for any allowed values of primal measure and dual variables, it holds also at the respective maximization and minimization limits. □

Theorem 1 establishes that the adversarial-contracting problem provides an upper bound on the value of information control. Importantly, this result underlies the *certification approach* to solve concave information-design problems.

Proposition 1. (Optimality Certification) *If measure $\pi \in \Delta(A \times \Omega)$ is implementable by information and by incentives, then π is optimal in the information-design problem.*

Proof. Take any primal measure π implementable by information, i.e., that satisfies the constraints of primal problem (6). If it is implementable by incentives, then there exist dual variables λ that implement this measure in dual adversarial-contracting problem

$\cdots \equiv \lambda_N(a_N) \equiv 0$. Faced with this contract, the agent would act to maximize $v(a,\omega)$ in each state, thus implementing the first-best allocation rule of the information designer. The goal of the adversarial principal can then be viewed as adjusting the null contract to minimize the expected payoff starting from this benchmark level.

(11), and

$$V^D = \inf_{\hat{\lambda} \in \times_i B(A_i)} \mathbb{E}_{\mu_0} \left[\sup_{a \in A} u^{\hat{\lambda}}(a, \omega) \right] \tag{13}$$

$$\leq \mathbb{E}_{\mu_0, \pi} \left[u^{\lambda}(a, \omega) \right] \tag{14}$$

$$= \int_{A \times \Omega} v(a, \omega) \mathrm{d}\pi - \int_{A \times \Omega} \sum_{i=1}^{N} \lambda_i(a_i) \dot{u}_i(a, \omega) \mathrm{d}\pi \tag{15}$$

$$= \int_{A \times \Omega} v(a, \omega) \mathrm{d}\pi \leq V^P, \tag{16}$$

where the first inequality follows from the implementability of π in the dual problem and the last three steps follow from the feasibility of π in the primal problem.

Furthermore, by Theorem 1, $V^D \geq V^P$. Combining the two inequalities, we obtain

$$V^D = \int_{A \times \Omega} v(a, \omega) \mathrm{d}\pi = V^P, \tag{17}$$

which proves the optimality of measure π. □

Proposition 1 offers a solution method for concave information-design problems. In the first step, one conjectures an optimal measure π^*. In the second step, one verifies that it can be implemented with information, which is equivalent to its feasibility in the primal problem, and that it can be implemented with incentives, e.g., by explicitly constructing the dual contract λ that implements it. In the last step, one sets an optimal information structure to privately recommend actions to players according to π^*: $\mathcal{I}^* = (A, \pi^*)$. The implementability of π^* with information implies that the players would follow the recommendations. The implementability of π^* with incentives confirms, by Proposition 1, that $\mathcal{I}^*$ is optimal among all information structures. In this case, we say that λ is a (dual) *certificate* of π^*, that λ *certifies* the optimality of π^*, and that π^* is a *certifiably optimal* or, simply, *certifiable* information structure.

3.4 On Certifiable Information Structures

Before proceeding with the application of the certification solution method in specific economic settings, we highlight one general property that holds for all certifiable information structures. This property is based on the observation that an allocation rule induced by a certifiable information structure must be undertaken at will by an agent in possession of full information in the dual problem. It has two consequences. First, the prior state distribution is irrelevant for the implementability of an allocation rule with incentives since the prescribed action profiles must be optimal state-by-state. Second, if a certifiable allocation rule randomizes over several action profiles at some state, the dual agent must be indifferent between these profiles and hence could just as well randomize over these profiles with different probabilities. That is, only the support of an allocation rule is relevant for the implementability with incentives.

Proposition 2. (Robustness to Marginal Distributions) *Consider two concave information-design problems that differ only in their prior state distributions, if at all. Let information structure $\mathcal{I}_1^*$ be certifiably optimal in the first problem and implement an allocation rule α_1^*. If information structure $\mathcal{I}_2$ implements in the second problem an allocation rule α_2 such that $\operatorname{supp}\alpha_2(\omega) \subseteq \operatorname{supp}\alpha_1^*(\omega)$ for all $\omega \in \Omega$, then $\mathcal{I}_2$ is certifiably optimal in the second problem.*

Proposition 2 highlights the robustness of certifiable information structures to their marginal distributions: over states and over actions. Either of these distributions may change without sacrificing optimality as long as the supports of the implemented allocation rules remain the same.[10]

We highlight that Proposition 2 is specific to concave games and does not generally hold in games with finitely many actions. For such problems, a given allocation rule is typically implementable for many prior distributions, yet an optimal information structure continuously changes with the prior. For concreteness, consider a leading example of Kamenica and Gentzkow (2011) in which a designer persuades a single receiver. The state space and the action space are binary: $A = \{a_0, a_1\}$, $\Omega = \{\omega_0, \omega_1\}$. The payoffs are $v(a, \omega) = 1$ if $a = a_1$ and zero otherwise; $u(a, \omega) = 1$ if $a = a_0, \omega = \omega_0$ or $a = a_1, \omega = \omega_1$ and zero otherwise. As long as $\mu_0(\omega_1) \in (0, 1/2)$, an optimal information structure sends two signals s_0, s_1 that induce posterior beliefs that assign probabilities 0 and 1/2 to state ω_1, respectively. The allocation rule induced by the optimal information structure changes with the prior: the higher the prior probability of state ω_1 is, the less likely signal s_1 is sent and action a_1 is taken in that state. However, the same allocation rule is implementable by information for a variety of priors. As a result, there is no robustness to the prior distribution in this example.

Proposition 2 enables us to assess the optimality of full transparency. Namely, say that an information structure is fully informative about the state if each player deduces the state with certainty, i.e., each private signal induces an extreme posterior belief

[10]This property anticipates the multiplicity of optimal information structures that we observe in applications in Section 4 and discuss in detail in Section 4.4.

about the state. Such an information structure can still allow for uncertainty about the actions of other players. We have the following.

Corollary 1. (Full State Information) *An information structure that is fully informative about the state is certifiably optimal if and only if it is certifiably optimal under all prior state distributions.*

This corollary follows from Proposition 2. If an allocation rule is implemented by an information structure that is fully informative about the state under some prior state distribution, then the same rule is implemented by the same information structure under any other prior state distribution because all prior uncertainty is resolved in either case. Hence, if such information is certifiably optimal under one prior distribution, then it implements the same allocation rule and is certifiably optimal under any other prior distribution.

Alternatively, we can use Proposition 2 to assess the support of induced action profiles under certifiable information structures. The larger the support is, the easier it is to construct another information structure that implements an allocation rule within that support. In the extreme case, if the action support covers the whole action space, then the support condition of Proposition 2 has no bite, and any information structure can be certified to be optimal.

Corollary 2. (Full-Support Noise) *If an information structure $\mathcal{I}^*$ is certifiably optimal and induces an allocation rule α^* with $\operatorname{supp} \alpha^*(\omega) = A$ for all $\omega \in \Omega$, then any information structure is certifiably optimal.*

Corollary 2 presents a case against using independent noises that induce full-support individual actions in optimal information structures. These information structures can never be optimal in concave problems with finitely many players, except in trivial cases in which the designer's expected payoff is invariant to the information provided.[11] However, such independent noises may optimally appear in the limit information structure as the number of players grows to infinity, as we show in Section 4.4.

3.5 Scope of the Certification Method

Can any optimal information structure be certified? Observe that by Theorem 1, the difference $G \triangleq V^D - V^P \geq 0$ between the optimal values of primal and dual problems

[11]This finding resonates with the analysis of Taneva (2019), who studied a parameterized setting with two players, a binary state, binary actions, and a symmetric designer's payoff function and showed that sending conditionally independent signals is never strictly optimal in that setting.

is nonnegative and constitutes a *duality gap*. The solution to the information-design problem can be certified if and only if (i) the duality gap is equal to zero, $V^D = V^P$, and (ii) solutions to both primal and dual problems exist. Thus, either all optimal information structures can be certified or none of them can.

While we expect properties (i) and (ii) to hold quite generally, neither is trivial. The former property is referred to as the case of "strong duality" in the literature on optimization. Strong duality always holds in linear programs with a finite number of arguments and constraints. However, for the first-order conditions to determine the best response, the information-design problem necessarily has to feature a continuum of actions and incentive constraints, and establishing strong duality even in well-behaved infinite problems is challenging (e.g., Daskalakis et al. (2017), Dworczak and Martini (2019), Dizdar and Kováč (2020)). The latter property requires the solutions to both problems to exist, which might fail due to a lack of compactness of the underlying optimization spaces.[12]

It is certainly of theoretical interest to understand under which conditions the certification method is guaranteed to work, and much study in the literature is devoted to finding such conditions in various problems. For example, these conditions could help to establish the properties of all optimal information structures without having to solve the information-design problem itself. Thus, in the Appendix, we provide a set of sufficient conditions by building on the Fenchel-Rockafellar duality of optimal transportation theory.[13] However, note that from a perspective of actually solving an information-design problem, knowing that any optimal information structure can be certified does not help in finding an optimal structure or a certifying contract. Conversely, any certifying contract by its very existence proves that the certification method applies, i.e., the duality gap is zero and both primal and dual solutions exist. Therefore, for any concave problem, it may be worth constructing the dual problem and searching for certifiable information structures. If successful, the method leads to the solution. This is exactly what we do in the next section.

[12] To establish properties (i) and (ii) in an infinite problem, one typically has to impose carefully chosen topologies on the relevant spaces, whereas the weak duality result is based solely on the algebraic structure.

[13] Concurrently, Cieslak et al. (2021) and Malamud and Schrimpf (2021) use transportation theory to establish strong duality in a class of settings with a single receiver.

4 Application: Normal-Quadratic Settings

In this section, we apply our theoretical machinery to study a broad subclass of concave information-design problems in which all payoffs are quadratic functions of actions and states and the state is normally distributed.

In particular, addition to the basic structure of the previous section, we impose two assumptions on the environment. First, we assume that the state components are jointly normally distributed so $\Omega = \mathbb{R}^K$ and $\omega \sim N(0, \Sigma)$, where all means are set to zero without loss of generality and Σ is the arbitrary covariance matrix. Second, we assume that there exist vectors $\hat{b}, b \in \mathbb{R}^N$, matrices $\hat{B}, B \in \mathbb{R}^{N \times K}$ and $\hat{C} \in \mathbb{R}^{N \times N}$, and a positive definite matrix $C \in \mathbb{R}^{N \times N}$ such that the designer's and player-i's payoffs are, respectively,

$$v(a, \omega) = a^T(\hat{b} + \hat{B}\omega) - \frac{1}{2} a^T \hat{C} a, , \tag{18}$$

$$u_i(a, \omega) = a^T(b + B\omega) - \frac{1}{2} a^T C a. \tag{19}$$

Elements $\hat{b}_i, B_{i\cdot}$ capture the base benefit of player i's action, elements $\hat{B}_{ij}, B_{ij}$ capture the interaction between player i's action and the jth state component, and elements $\hat{C}_{ij}, C_{ii}$ capture the interaction between the actions of players i and j, for the designer and the players, respectively.[14] This setting allows for asymmetries across players by allowing the rows within matrices B and C to differ from each other.

To maximize her payoffs, the designer chooses an information structure $\mathcal{I} = (S, \pi)$. One class of information structures available to the designer is the class of Gaussian information structures, under which the players' individual signals and the state are jointly normally distributed. This class has been extensively used in the economic literature because of its richness and tractability. However, the focus on Gaussian information structures excludes many other natural choices, e.g., monotone partitions of the state space, and the question of how limiting this focus is has not been addressed.

In the following, we do not restrict the designer to use Gaussian information structures to achieve her objective. Instead, we provide conditions under which a Gaussian information structure is optimal among *all* information structures. In particular, we show the optimality of a direct Gaussian information structure that recommends ac-

[14]For any matrix X, the term $X_{i\cdot}$ denotes its ith row.

tions proportionally to state components:

$$a^*(\omega) = a_0 + R\omega, \tag{20}$$

where $a_0 \in \mathbb{R}^N$ is a constant vector and R is an $N \times N$ *responsiveness matrix* that determines the responsiveness of recommended actions to different state components. Under this information structure, player i observes a recommendation $a_i(\omega) = a_{0i} + R_{i\bullet}\omega$. Thus, the player can only infer the value of a linear combination of the state components and generically, as long as $K > 1$, receives only imperfect information about the state and the actions of other players.

Theorem 2. (Optimal Information) *An information structure that recommends a linear allocation rule $a(\omega) = a_0 + R\omega$ is optimal among all information structures if:*

(i) $a_0 = C^{-1}b$ and $(C_{i\bullet}R - B_{i\bullet})\Sigma R_{i\bullet}^T = 0$ for all $i = 1, \dots, N$, and

(ii) $R = (\hat{C} + 2D(x)C)^{-1}(\hat{B} + D(x)B)$ for some $x \in \mathbb{R}^N$ such that $C + 2D(x)\hat{C}$ is positive definite, where $D(x)$ is a diagonal matrix with $D(x)_{ii} \triangleq x_i$.

To prove the theorem, we use the certification method developed in the previous section. In particular, we show that under the conditions of Theorem 2, allocation rule (20) is implementable both by information in the primal problem and by incentives in the dual problem.

The implementability of the allocation rule by information, i.e., the players' obedience in following recommendations, is captured by the system of first-order conditions, as in any concave game. In a quadratic game, these conditions are linear in state components and actions. These conditions must hold for all recommended actions so they must also hold on average, which uniquely pins down the constant term a_0, forming the first part of condition (i). Next, since the state components are jointly normally distributed, the marginal payoffs present in the first-order conditions are jointly normally distributed with the recommended actions. Thus, the sufficient condition for obedience is that the recommended actions and the marginal payoffs are uncorrelated, forming the second part of condition (i).

To summarize, a linear allocation rule (20) is implementable by information whenever its coefficients a_0 and R satisfy the condition (i) of Theorem 2. However, the fact that a given allocation rule is implementable does not mean that it is optimal. After all, there are many implementable linear allocation rules and, potentially more importantly, there are many implementable nonlinear allocation rules, which may be preferred by the designer.

This is where the second condition of Theorem 2 plays a role. Under this condition, the linear allocation rule can be implemented by incentives in the dual problem. This can be achieved using a contract such that $\lambda_i(a_i)$ is linear in a_i with the proportionality coefficient being equal to x_i. Under the linear contract, the dual payoff is quadratic in the action profile. The positive-definiteness requirement of condition (ii) ensures that this payoff is concave and thus the best-response allocation rule exists and unique. The first part of condition (ii) guarantees that this allocation rule responds to the state according to the responsiveness matrix R. The constant terms of the linear contract can then always be set such that the constant terms of the best-response allocation rule match a_0. The result follows.

Theorem 2 provides a two-step procedure for finding optimal information structures in normal-quadratic environments. In the first step, one uses condition (i) to identify the parameters a_0, R of the candidate information structure. In the second step, one searches for $x \in \mathbb{R}^N$ that satisfies condition (ii); if such x exists, then it dual-certifies the optimality of the candidate information structure. In the next section, we apply this procedure to characterize optimal information regulation in a market with differentiated product competition.[15]

4.1 Differentiated Bertrand Competition

We apply our solution method in a setting of differentiated product competition, in which a designer controls the demand information available to firms. One can think of this designer as a platform, such as Amazon or AliExpress, that organizes the marketplace in which the firms compete. The platform has more detailed knowledge about demand conditions than firms do, for instance, because it has access to a larger and more recent sales data set or higher processing capabilities. The platform can communicate this information privately to each firm, for instance, by giving it access to personalized data analysis or by direct price recommendations. By programming its algorithms, the platform can basically design and commit to any information structure. We characterize the information structure such a platform would optimally design and the resulting allocation distortions.

Formally, the market consists of two firms and a continuum of consumers. Each firm sells a single product and competes in price with its opponent, so action a_i is the

[15] In Sections 4.2 and 4.3, we present alternative settings in which the positive-definiteness part of condition (ii) is not satisfied and in which, as a result, the optimal allocation rule can feature an extraneous noise.

price set by firm i. Demand is ex ante symmetric across firms and is generated by a continuum of consumers that differ in their tastes. Each consumer has a type $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$ and decides how much of the firms' products to consume, $q = (q_1, q_2)$. The type components are independently and identically distributed according to a normal distribution with mean $\bar{\theta}$ and variance σ^2. The ex post payoff of a type-θ consumer who consumes quantities q at prices a is:

$$w(q, a, \theta) \triangleq w_0 + \frac{1}{2}(\theta - q)^T W^{-1}(\theta - q) - a^T q, \tag{21}$$

where w_0 is a constant shift parameter henceforth normalized to zero and W is an $N \times N$ negative semidefinite matrix with $W_{11} = W_{22}$. Thus, the consumer's type θ determines her consumption bliss points, optimal at zero prices, whereas W captures the substitution effects across products.

For any price vector $a \in A$, the quantity of good i demanded by a consumer of type θ is equal to:[16]

$$q_i(a, \theta) = \theta_i + \eta a_i + \xi a_{-i}, \tag{22}$$

where $\eta \triangleq W_{ii} < 0$ and $\xi \triangleq W_{-ii}$; thus, equivalently, $q(a, \theta) = \theta + Wa$. Equation (22) reveals that the chosen type structure microfounds linear demand; the consumer's type determines the intercept of the demand curve for each product. We refer to η as *own-price sensitivity* and to ξ as *cross-price sensitivity*.

The firms have quadratic costs of production such that their profits are:

$$u_i(a, \theta) = a_i q_i(a, \theta) - c q_i(a, \theta)^2. \tag{23}$$

The resulting ex post valuations for consumer surplus and total profits are:

$$CS(a, \theta) = -a^T\theta - \frac{1}{2}a^T W a, \tag{24}$$

$$\mathit{\Pi}(a, \theta) \triangleq u_1 + u_2 = -c\theta^T\theta + a^T(1 - 2cW)\theta - \frac{1}{2}a^T(2W + 2cW^2)a. \tag{25}$$

The designer's payoff is a convex combination of consumer surplus and total profits with $\delta \in [0, 1]$ measuring the weight placed on consumer surplus:

$$v(a, \theta) = \delta \times CS(a, \theta) + (1 - \delta) \times \mathit{\Pi}(a, \theta). \tag{26}$$

[16] As standard, this specification allows the prices and quantities to be negative.

Given (26), the optimal designer's choices in the extreme cases $\delta = 1$ and $\delta = 0$ correspond to consumer-optimal and producer-optimal information structures, respectively, whereas the choice in the case $\delta = 1/2$ corresponds to the socially efficient information structure. As the welfare weight δ spans the interval $[0, 1]$, the corresponding solutions span the Pareto frontier in the space of consumer surplus and total profits.

Clearly, this setting is normal-quadratic because (i) firms' payoffs are quadratic and concave, (ii) designer's payoffs are quadratic, and (iii) consumer types are normally distributed.[17] Consequently, if the conditions of Theorem 2 are satisfied, then an optimal information structure recommends a linear allocation rule and can be characterized in closed form. At the end of this section, we show that this optimal characterization indeed works, but first, we discuss several natural benchmarks to give a sense of the trade-offs faced by the designer.

Direct Price Control We begin the analysis by studying a hypothetical scenario in which the designer can directly control the prices set by the firms. This scenario constitutes a first-best benchmark; it provides an upper bound on the designer's payoff and illustrates the designer's preferred pricing.

This problem admits a solution only if δ is not excessively high, i.e., only if the seller does not overly value the consumer's welfare. Namely, there is a threshold value δ^{FB}:

$$\delta^{FB} = \frac{2 + 2c(-\eta - |\xi|)}{3 + 2c(-\eta - |\xi|)}, \tag{27}$$

such that if $\delta > \delta^{FB}$, then the designer can arbitrarily increase her payoff by setting arbitrarily large negative prices, because the monetary transfer to consumers outweighs any allocation inefficiency. In contrast, if $\delta < \delta^{FB}$, then the designer's problem is well-behaved: it is concave and admits a unique solution that can be found by first-order conditions to (26). The solution is proportional to the type and can be written in matrix form as

$$a^{FB} = R^{FB}\theta, \tag{28}$$

where the entries of matrix R^{FB} are nonlinear functions of the problem's parameters and their exact formulation is relegated to the Appendix. We use this solution to illustrate the first-best benchmark later in the section.

[17] The exact mapping between the settings requires state normalization $\omega_i \triangleq \theta_i - \bar{\theta}$ and is presented in detail in the Appendix.

In our setting, the designer does not control prices directly but rather indirectly through demand information she supplies to firms. Before deriving the generally optimal policy, it is instructive to analyze two extreme information benchmarks that are particularly easy to implement in practice: not informative and completely informative information structures.

No Information If the firms obtain no demand information, $S_1 = S_2 = \{s_0\}$, then their beliefs stay at the prior, and the equilibrium prices satisfy the first-order conditions derived from (24). In equilibrium, each firm sets a price:

$$a_i^{NI} = \frac{1 - 2c\eta}{-2\eta(1 - c\eta) - \xi(1 - 2c\eta)}\bar{\theta}. \tag{29}$$

Lacking demand information, the firms fix their prices at a level proportional to the expected consumer type. The equilibrium prices do not depend on finer details about the type distribution because the demand is linear.

Full Information If the firms obtain full demand information, $S_1 = S_2 = \Theta$ and π is concentrated on event $s_1 = s_2 = \theta$, then the consumer type is always commonly known. In equilibrium, each firm responds linearly to the type components perfectly anticipating the price of its opponent:

$$a_i^{FI}(\theta) = \frac{-2\eta(1 - c\eta)(1 - 2c\eta)}{4\eta^2(1 - c\eta)^2 - (1 - 2c\eta)^2\xi^2}\theta_i + \frac{(1 - 2c\eta)^2\xi}{4\eta^2(1 - c\eta)^2 - (1 - 2c\eta)^2\xi^2}\theta_{-i}. \tag{30}$$

In a sense, this behavior generalizes price-setting under no information. If $\theta_1 = \theta_2 = \mu$, then prices are the same as those under no information. If $\theta_1 \neq \theta_2$, then demand is asymmetric across firms, and prices are adjusted to reflect the competitive advantages.

Optimal Information Structure The choice of any of the extreme information structures has drawbacks. Providing no information misses the opportunity to strengthen the link between consumer type and allocation and thus potentially limits efficiency. Providing full information may exacerbate competition and dissipate firm profits. Providing partial information may alleviate the individual shortcomings of extreme information structures and, as we will show, is the best option in most cases.

We find an optimal structure using the certification method developed for the normal-quadratic setting. The sufficiency conditions of Theorem 2 stipulate the existence of certification parameters $(x_1, x_2) \in \mathbb{R}^2$. Given the symmetry of the environment, it is natural to conjecture $x_1 = x_2 = x$. By condition (ii) of the Theorem, any such x uniquely pins down the responsiveness matrix $R(x)$, whose elements are quadratic functions of x. Furthermore, condition (i) requires equation $(C_{i\cdot}R(x) - B_{i\cdot})\Sigma R(x)_i^T = 0$ to hold for both firms. Due to the symmetry of the environment, this condition becomes a single equation:

$$f(x) = 0, \tag{31}$$

where $f(x)$ is a degree-four polynomial whose coefficients depend on the parameters of the problem and are explicitly defined in the Appendix. Equation (31) admits up to four real solutions. If any of these solutions makes the matrix $C + 2D((x,x))\hat{C}$ positive semidefinite, then by Theorem 2 the information structure that recommends the linear allocation rule $a(\theta) = a_0(x) + R(x)\theta$ is optimal.

Proposition 3. (Optimal Demand Information) *Polynomial $f(x)$ certifies optimal demand information: if there exists $x \in \mathbb{R}$ such that $f(x) = 0$ and $C + 2D((x,x))\hat{C}$ is positive semidefinite, then an information structure that recommends allocation rule $a(\theta) = a_0(x) + R(x)\theta$ is optimal.*

We use Proposition 3 to derive optimal information structures, to understand how they differ from the benchmarks presented above and to see how they depend on the weight the designer attaches to consumer surplus. For concreteness, in what follows, we fix the basic parameters of the problem to $c = 1$, $\bar{\theta} = 3$, $\sigma^2 = 1$, $\eta = -1$, and $\xi = 1/2$.

The equilibrium strategies under no information and under full information can be immediately calculated as:

$$a^{NI}(\theta) = \frac{6}{5},\ a^{FI}(\theta) = \frac{48}{55}\theta_i + \frac{18}{55}\theta_{-i}. \tag{32}$$

The first-best benchmark can also be immediately calculated and is used in the upcoming illustrations.

To solve for an optimal information structure for any given $\delta \in [0,1]$, we construct the polynomial $f(x)$ and solve for its roots, which is possible to do in closed form in radicals. We compute this solution and show that for all $\delta \in [0,1]$, with the single exception of $\delta^{cr} = 11/18$, which we call a critical value, there exists a unique root of $f(x)$ that makes the matrix $C + 2D((x,x))\hat{C}$ positive definite. This root, henceforth denoted by $x(\delta)$, forms a certifying parameter and is plotted in the Appendix as a function of δ.

Once we find the certifying parameter, we can immediately construct an optimal

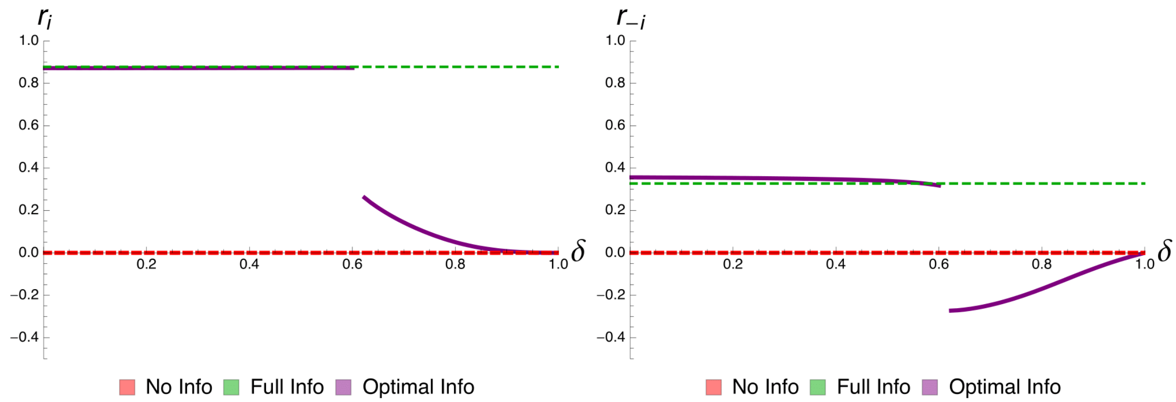

Figure 1: Price responsiveness to own demand (left) and opponent's demand (right) under no information, full information, and optimal information structure. Calculated at $c = 1$, $\bar{\theta} = 3$, $\sigma^2 = 1$, $\eta = -1$, and $\xi = 1/2$.

information structure. This structure recommends a linear allocation rule

$$a_i^*(\theta) = a_{i0} + r_i\theta_i + r_{-i}\theta_{-i}. \tag{33}$$

We refer to responsiveness coefficients r_i, r_{-i} as *own-responsiveness* and *cross-responsiveness*, respectively. We plot the optimal responsiveness coefficients in Figure 1, together with their counterparts under full information and no information.

For $\delta < \delta^{cr}$, the optimally induced behavior resembles that under full information. The own responsiveness is, in fact, exactly the same. However, the cross-responsiveness differs, showing that providing full information is not optimal. At $\delta = 0$, to dampen competition, the designer induces larger responses to the opponent's demand. As δ increases, the cross-responsiveness decreases. Around $\delta = \delta^{cr}$, the optimal information structure undertakes a discontinuous structural change. The own-responsiveness r_i plummets in absolute value, whereas the cross-responsiveness r_{-i} changes its sign, so firms respond oppositely to the same demand shock. As δ further increases in the region $\delta > \delta^{cr}$, both responsiveness parameters gradually decrease in their absolute values. At $\delta = 1$, both parameters equal zero: a designer who wishes to maximize consumer surplus should not reveal any demand information.

The induced equilibrium behavior translates into distinct patterns of price volatility and price cross-correlation (Figure 2). For lower consumer weights $\delta < \delta^{cr}$, the price volatility measured by price's standard deviation $\sigma_i = \sigma\sqrt{r_i^2 + r_{-i}^2}$ is high, and the prices are highly positively correlated, as witnessed by the high value of the Pearson correlation coefficient $\rho_{i,-i} = \frac{\text{cov}(a_i, a_{-i})}{\sigma_i\sigma_{-i}} = \frac{2r_ir_{-i}}{r_i^2 + r_{-i}^2}$. This region is marked by coordination

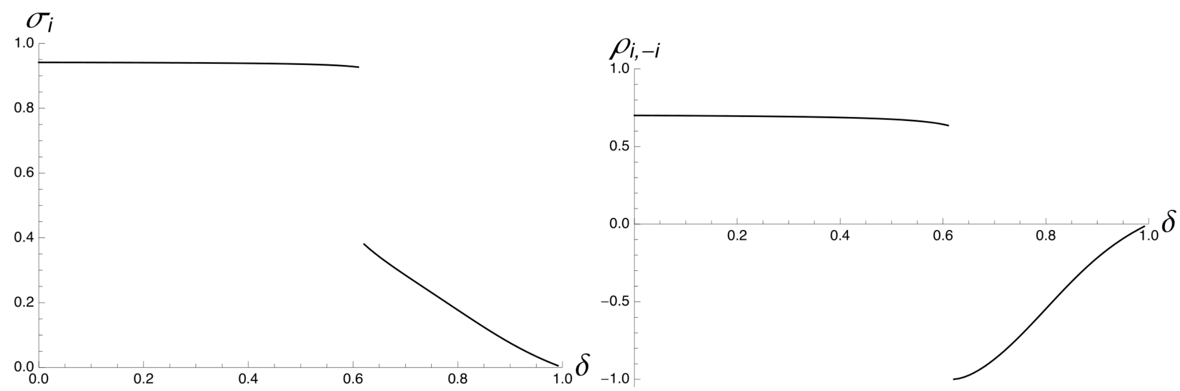

Figure 2: Price volatility (left) and price correlation (right) under the optimal information structure. Calculated at $c = 1$, $\bar{\theta} = 3$, $\sigma^2 = 1$, $\eta = -1$, and $\xi = 1/2$.

and high volatility of prices. In contrast, for higher consumer weights $\delta > \delta^{cr}$, the price volatility is substantially lower and the product prices are negatively correlated. This region is associated with anticoordination and low volatility of prices. These distinct price patterns may be easier to observe in practice than are firms' strategy parameters and can serve as an indicator of the underlying information structure and interests of the designer.

What occurs near the critical value of consumer weight δ? Why does the optimal information structure change discontinuously? The formal explanation is as follows. The certifying parameter x changes continuously for all $\delta \in (0, 1)$. However, the matrix $C + 2D((x, x))\hat{C}$ evaluated at $x = x(\delta^{cr})$ loses a rank and becomes noninvertible. As a result, the optimal best response of an agent in a dual problem, proportional to the inverse of that matrix whenever the matrix is invertible, changes discontinuously. In other words, even though a certifying contract changes continuously, the allocation rule that it implements exhibits a jump.

To intuitively understand the economic reasoning behind this discontinuity, it is instructive to compare the induced pricing under the optimal information structure to its first-best counterpart under direct price control (Figure 3). Under direct price control, both responsiveness coefficients start high at $\delta = 0$ and progressively decrease, diverging to negative values as δ approaches δ^{FB}. The pricing behavior induced by optimal information control overall follows the same responsiveness pattern. However, information control has limits, as it needs to account for firms' willingness to follow recommendations. As a result, there are important caveats. For $\delta < \delta^{cr}$, only the cross-responsiveness decreases, while the own-responsiveness remains at the full-information level. At the critical value $\delta = \delta^{cr}$, the first-best responsiveness levels become too low

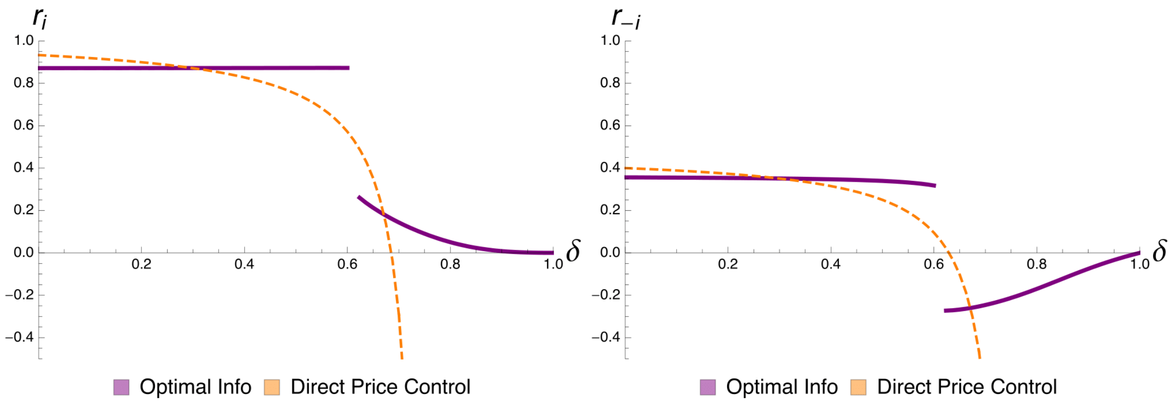

Figure 3: Price responsiveness to own demand shock (left) and opponent's demand shock (right) in the first-best benchmark and under the optimal information structure. Calculated at $c = 1$, $\bar{\theta} = 3$, $\sigma^2 = 1$, $\eta = -1$, and $\xi = 1/2$.

to be approached in a coordinated fashion, with both responsiveness coefficients being positive. As a result, both responsiveness coefficients discontinuously drop and the cross-responsiveness becomes negative. For $\delta > \delta^{cr}$, the own-responsiveness begins to gradually decrease while the cross-responsiveness increases, with both values converging to zero as δ approaches 1. Not being able to directly funnel monetary surplus to consumers, the principal provides less and less consumer information to minimize price targeting.

4.2 First-Order Persuasion

In this section, we apply our solution method to a game in which multiple players try to correctly predict a common underlying state. The state is one-dimensional $\omega \in \mathbb{R}$ and distributed according to the prior distribution μ_0. There are $N \geq 2$ players that make predictions, $A_i = \mathbb{R}$, and the ex post payoff of each player is

$$u_i(a, \omega) = -(a_i - \omega)^2. \tag{34}$$

Each player's payoff depends only on his action and the state; there is no strategic interaction across the players. Information design in such games is referred to as *first-order Bayesian persuasion* and studied by Arieli, Babichenko, Sandomirskiy, and Tamuz (2021) in the case of a binary state.[18]

[18]The name is derived from the fact that in such a game, the player's first-order belief about the state is the sufficient statistic for a best response. It is not to be confused with the first-order approach to information design put forward in the current manuscript.

An alternative way of presenting this environment is to note that for any given belief $\nu \in \Delta(A_{-i} \times \Omega)$, player i's best response is simply the posterior expectation of the state:

$$a_i^*(\nu) = \mathbb{E}_\nu[\omega]. \tag{35}$$

As such, the information-design problem is strongly connected with the question of what distributions of posterior expectations can be induced by information structures.

If the designer's objective $v(a, \omega)$ is additively separable across players' actions, then the first-order Bayesian persuasion reduces to a collection of single-receiver Bayesian persuasion problems. However, if the designer's objective features interaction across players' actions, then the whole action profile needs to be tracked at the same time and the problem cannot be separated. In the following, we consider two classes of such designer objectives: expectation polarization and co-movement.

4.2.1 Expectation Polarization

Following Arieli et al. (2021), we consider the question of inducing maximal polarization of the players' posterior beliefs. To this end, we set the designer's payoff as

$$v(a, \omega) = \sum_{i,j} (a_i - a_j)^2. \tag{36}$$

Given the payoff, the designer aims to make the agents' rational predictions as far as possible from each other. In this situation, the designer benefits from sending private signals: any public information structure, including full information or no information, leads to the players to have the same predictions and consequently minimizes the designer's objective. At the same time, the designer cannot induce arbitrary action distributions because (i) by Bayes' plausibility, each individual prediction must be a martingale and (ii) each individual prediction drifts toward the common true state (Doval and Smolin (2021)). On the one hand, an optimal information structure must provide some state information to move players' predictions and, on the other hand, should heterogeneously obfuscate the information to counterbalance truth drifting.

Proposition 4. (Optimal Polarizing Information) *For any prior distribution μ_0, if an information structure $\mathcal{I}$ induces a best-response action profile such that for all $\omega \in \Omega$*

$$\frac{1}{N} \sum_{i=1}^{N} a_i(\omega) = \frac{1}{2}(\omega + \mathbb{E}[\omega]), \tag{37}$$

then this information structure is optimal.

The proof is based on the duality certification method. We show that allocation (37) can be implemented by incentives in the adversarial problem by linear contracts with $\lambda_i(a_i) = -N(a_i - \mathbb{E}[\omega])$. Hence, if this allocation can be implemented by some information structure, then that information structure is optimal.

Proposition 4 provides a sufficient condition for the optimality of a given information structure. At the same time, the condition is placed only on the aggregate prediction across all players. As such, it allows freedom with respect to how individual predictions are distributed and anticipates the possible multiplicity of optimal information structures. We illustrate this multiplicity by presenting two classes of optimal information structures.

Corollary 3. (Polarizing by Selective Informing) *Let the number of players N be even. For any prior distribution μ_0, an information structure that fully reveals the state to half of the players and provides no information to the other half is optimal.*

Corollary 3 shows that in a large variety of settings, expectation polarization is achieved by a remarkably simple information structure that informs only half of a population. This result extends the findings of Arieli et al. (2021) from two players and a binary state to a general number of players and states. At the same time, this simple policy is not symmetric and, moreover, not feasible for an odd number of players. It is also not uniquely optimal. At least for the case of a normally distributed state, and for any number of players, there exists an optimal information structure that is Gaussian and symmetric.

Corollary 4. (Polarizing by Coordinated Informing) *Let the state ω be normally distributed with mean $\bar{\omega}$ and variance σ^2. For any number of players $N \geq 2$, the information structure that recommends the following actions as functions of the state is optimal:*

$$a_i(\omega) = \frac{\omega + \bar{\omega}}{2} + \varepsilon_i - \frac{1}{N-1} \sum_{j \neq i} \varepsilon_j, \tag{38}$$

where ε_i are i.i.d. Gaussian noises with mean 0 and variance $\frac{N-1}{4N}\sigma^2$.

Corollary 4 showcases an alternative way to polarize the posterior expectations of a group of players—by carefully designing the correlation structure across their signals. The optimal information structure is symmetric and provides imperfect information to each player; the individual noises are negatively correlated and designed in a way to ensure that the aggregate prediction follows optimality condition (37).

4.2.2 Co-Movement and Miscoordination

In the previous section, the designer had a pure polarization objective, irrespective of the state. In this section, we enrich her incentives by incorporating the willingness to co-move the players' actions with the state. Namely, assume that the prior state expectation is zero, $\mathbb{E}[\omega] = 0$, so that by the martingale property, each prediction is zero on average, $\mathbb{E}[a_i] \equiv 0$, and let the designer's payoff be:

$$v(a, \omega) = \frac{\sum_{i=1}^{N} a_i}{N}\omega - \rho\frac{\sum_{i=1}^{N}\sum_{j\neq i} a_i a_j}{N^2}. \tag{39}$$

for $\rho \geq 0$. The first element of (39) captures the designer's willingness to co-move the players' average action with the state. The second element captures the designer's willingness to miscoordinate the individual actions of different players and is similar to the polarizing payoff. The parameter ρ captures the intensity of the latter. These two designer's objectives are in conflict. To maximize the former, the designer should provide full state information. However, doing so would make the actions perfectly correlated and would dampen the miscoordination payoff component.

Proposition 5. (Optimal Co-Movement Information) *For any prior distribution μ_0: if $\rho \leq \frac{N}{2N-1}$, then providing full information to all players is optimal; if $\rho > \frac{N}{2N-1}$ and an information structure $\mathcal{I}$ induces a best-response action profile such that for all $\omega \in \Omega$*

$$\frac{1}{N}\sum_{i=1}^{N} a_i(\omega) = \left(\frac{1}{2\rho} + \frac{1}{2N}\right)\omega, \tag{40}$$

then this information structure is optimal.

As before, the proof is based on the duality certification method. We show that allocation (40) can be implemented by incentives in the adversarial problem by linear contracts with $\lambda_i(a_i) = -\frac{\rho}{2N^2}a_i$. Hence, if that allocation can be implemented by some information structure, then that information structure is optimal.

Note how Proposition 5 mirrors Proposition 4. Both results certify the optimality of an information structure by linear dual contracts. Both results pin down only aggregate action. However, condition (40) differs in its responsiveness of the aggregate action in state—it depends on ρ and on the number of players. Naturally, the responsiveness to the state increases as the miscoordination motive captured by ρ decreases. At the critical value $\rho = \frac{N}{2N-1}$, the responsiveness is maximal and equal to 1, corresponding to full state information. For all $\rho \leq \frac{N}{2N-1}$, full state information is optimal.

In the following, we focus on the case $\rho \geq \frac{N}{2N-1}$. Similar to the case of polarizing information, we can implement the optimal allocation (40) in several ways.

Corollary 5. (Co-Moving by Selective Informing) *If $N^* \triangleq N(\frac{1}{2\rho} + \frac{1}{2N})$ is a natural number, then for any prior distribution μ_0, an information structure that fully reveals the state to N^* players and provides no information to others is optimal.*

The simple selective informing policy may again be optimal. The condition of Corollary 5 is quite restrictive because it is not satisfied for "generic" ρ. However, note that as the number of players grows, then the set of such ρ becomes denser and denser. As such, selective informing would deliver an approximately optimal payoff for any prior state distribution. At the same time, if the prior uncertainty is Gaussian, then there exists an exactly optimal Gaussian information structure:

Corollary 6. (Co-Moving by Coordinated Informing) *Let the state ω be normally distributed with mean 0 and variance σ^2. The information structure that recommends the following actions as the functions of the state is optimal:*

$$a_i(\omega) = \left(\frac{1}{2\rho} + \frac{1}{2N}\right)\omega + \varepsilon_i - \frac{1}{N-1}\sum_{j \neq i} \varepsilon_j, \tag{41}$$

where $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ are symmetric noises independent of each other and the state, and $\sigma_\varepsilon^2 = \frac{1}{4\rho^2}\frac{N-1}{N}\left(1 + \frac{\rho}{N}\right)\left(\frac{(2N-1)\rho}{N} - 1\right)\sigma^2$.

According to Corollary 6, an optimal information structure provides each player with a Gaussian estimate of the state. The estimate errors are correlated across players in a way that ensures that an average action is a deterministic and linear function of the state. The estimate precision is chosen to achieve an optimal trade-off between the coordination of the average action with the state and the anticoordination across players. Naturally, the precision decreases as the designer's anticoordination motives increase. For $\rho \leq \frac{N}{2N-1}$, providing perfect information is optimal, $\sigma_\varepsilon^2 = 0$. As ρ increases, the noise variance σ_ε^2 increases and converges in the limit to $\frac{(N-1)(2N-1)}{4N^2}\sigma^2$. This simple Gaussian information structure is optimal across all possible information structures.

4.3 Investment Game

In this section, we demonstrate the applicability of our solution method in an investment game in the spirit of the literature on large games (e.g., Angeletos and Pavan (2007) and Bergemann and Morris (2013)).

There are N players, who simultaneously decide how much to invest in a project, $A_i = \mathbb{R}$. The profitability of a project is uncertain; it depends on the unknown project quality $\omega \in \mathbb{R}$ and on the total amount of investment. The ex post payoff of player i is

$$u_i(a, \omega) = (\omega - rA)a_i - ca_i, \tag{42}$$

where $r > 0$ is a congestion rate, $c > 0$ is an opportunity cost of investment, and $A \triangleq \sum_{i=1}^{N} a_i$ is a total investment in the project.[19] As $r > 0$, the project has decreasing returns to scale: its average profitability decreases in the total investment.

It is convenient to rewrite payoff (42) as

$$u_i(a, \theta) = r(\theta - A)a_i, \tag{43}$$

where $\theta \triangleq \omega/r - c$ is a normalized project quality. Given (43), for any belief $\nu \in \Delta(A_{-i} \times \Theta)$, the player i's best response can be found via first-order condition to equal

$$a_i^*(\nu) = \mathbb{E}_\nu \left[\frac{\theta - A_{-i}}{2} \right], \tag{44}$$

where $A_{-i} \triangleq \sum_{j \neq i} a_j$, so that the best response linearly increases in the normalized state expectation and linearly decreases in the expected amount of total investment made by other players. The player's actions are thus strategic substitutes.

Applying the ex ante expectation to both sides of (44) and using the law of iterated mathematical expectations, we observe that the profile of expected individual investments must satisfy a system of linear equations which doesn't depend on the information structure. This system has a unique and symmetric solution according to which

$$\mathbb{E}[a_i] \equiv \frac{1}{N+1} \mathbb{E}[\theta]. \tag{45}$$

Hence, for any information structure, the expected individual investment of each player is given by (45). The resulting total investment is fixed at $\mathbb{E}[A] \equiv \frac{N}{N+1}\mathbb{E}[\theta] = \frac{N}{N+1}(\mathbb{E}[\omega]/r - c)$. Naturally, the total investment increases in the expected project quality, and decreases in the congestion rate r and in the opportunity costs c. It also increases in the number of players, converging to the expected normalized state as this number

[19] Note that this setting can also be interpreted as a Cournot competition with linear demand and linear production costs.

goes to infinity. Intuitively, as the number of players grows, each individual investment becomes smaller and each player internalizes less the congestion effect induced by his investment.

The information designer can fully control the information about the project quality and can privately inform each player about it. While the designer cannot affect the total amount of investment, she can direct the investment towards more productive projects. The designer aims to maximize the total profits generated by the project, her ex post payoff is[20]

$$v(a, \theta) = \sum_{i=1}^{N} (\theta - A) a_i = (\theta - A) A = \theta A - A^2. \tag{46}$$

As such, the designer must balance two conflicting objectives. On the one hand, the designer wants to make the aggregate investment correlated with the project quality to maximize $\mathbb{E}[\theta A]$. On the other hand, the designer wants to minimize the investment volatility, to maximize the term $\mathbb{E}[-A^2]$. If she could direct the players' actions directly, then she would set the total investment to respond to the state as $A(\theta) = \theta/2$. However, as we will see, this first-best outcome cannot be achieved with information control.

Before proceeding with derivation of the optimal information structure, we consider two extreme information structures that present natural benchmarks.

No Information Under no additional information, all players base their investment decisions only on the prior estimate of project quality. The unique equilibrium is one in which each player invests

$$a_i^{NI} = \frac{1}{N+1} \mathbb{E}[\theta]. \tag{47}$$

The investment is uniform across projects with different qualities, resulting in total investment $A = \frac{N}{N+1} \mathbb{E}[\theta]$ and the designer's payoff

$$v^{NI} = \mathbb{E}\left[(\theta - \frac{N}{N+1} \mathbb{E}[\theta]) \frac{N}{N+1} \mathbb{E}[\theta] \right] = \frac{N}{(N+1)^2} \mathbb{E}^2[\theta]. \tag{48}$$

Full Information In contrast, under a fully informative information structure, each player has a complete information about the project quality and takes it into account

[20] Because the expected investment and the expected investment costs are invariant to information, this payoff also captures the objective of maximizing the total investor welfare.

when investing. For any commonly known state, the ensuing game admits a strict potential and, thus, has a unique equilibrium (Neyman (1997)). This equilibrium is symmetric and each player invests proportionally to the normalized state as

$$a_i^{FI}(\theta) = \frac{1}{N+1}\theta. \tag{49}$$

The investment is correlated with the project quality and is the same for all players, resulting in total investment $A(\theta) = \frac{N}{N+1}\theta$. The designer's payoff is

$$v^{FI} = \mathbb{E}\left[(\theta - \frac{N}{N+1}\theta)\frac{N}{N+1}\theta\right] = \frac{N}{(N+1)^2}\mathbb{E}[\theta^2] = \frac{N}{(N+1)^2}(\mathbb{E}^2[\theta] + \mathbb{V}[\theta]). \tag{50}$$

Comparing the designer's payoffs in these two benchmarks, we see that providing full information outperforms providing no information. The relative benefit increases in the variance of normalized state θ. As such, it increases in the variance of the project quality $\mathbb{V}[\omega]$ and decreases in congestion rate r.

In both benchmarks, as the number of players goes to infinity, the designer's payoff which captures the total project profit converges to zero. Intuitively, since the aggregate investment remains the same, as the number of players grows their individual contributions become increasingly small and the project must become increasingly less attractive. In the limit, the individual rent as well as the total profit are dissipated.[21] However, in what follows, we show that this problem can be alleviated by a careful design of the game's information structure.

Optimal Information The design of an optimal information structure must reflect the trade-offs present in designer's objective (46). On the one hand, the designer should provide information in order to make the investment better correlated with the project quality and to boost the investment efficiency. On the other hand, providing information correlates individual decisions and exacerbates investment congestion.

Proposition 6. (Optimal Investment Information) *For any prior distribution μ_0, if an information structure $\mathcal{I}$ induces a best-response action profile such that for all $\theta \in \Theta$*

$$\sum_{i=1}^{N} a_i(\theta) = \frac{1}{2}\left(\theta + \frac{N-1}{N+1}\mathbb{E}[\theta]\right), \tag{51}$$

[21]If the setting is interpreted as Cournot competition, then this rent dissipation reflects the zero-profit property of large competitive markets.

then this information structure is optimal.

The proof is based on the duality certification method and demonstrates that allocation (37) can be implemented by incentives in the adversarial-contracting problem. However, in contrast to our other settings, the implementation is achieved by a constant contract with $\lambda_i(a_i) = \frac{N-1}{r(N+1)^2}\mathbb{E}[\theta]$.

As in Section 4.2, the optimality condition of Proposition 6 features only the aggregate action and leads to multiplicity of optimal information structures.

Corollary 7. (Guiding Investment by Selective Informing) *For any prior distribution μ_0 and any number of players, an information structure that fully reveals the project quality to a single player and provides no information to all others is optimal.*

Corollary 7 shows that, remarkably, in a large variety of settings, the same information structure is optimal among all possible information structures. The designer can simply designate a single player and provide full state information to him. This player will be the only one making informed decisions. All other players will be investing the same amount irrespective of the project quality. This information structure ensures that the total investment follows (51).

The selective informing policy is robustly optimal across many environments, yet, it is inegalitarian in that it places all information in the hands of a single player. However, similarly to the results in Section 4.2, if the prior state is normally distributed, then there is another optimal information structure which provides information symmetrically across players.

Corollary 8. (Guiding Investment by Coordinated Informing) *Let the normalized state θ be normally distributed with mean $\bar{\theta}$ and variance σ^2. The information structure that recommends the following actions as the functions of the state is optimal:*

$$a_i(\theta) = \frac{1}{2N}\theta + \frac{N-1}{2N(N+1)}\bar{\theta} + \varepsilon_i - \frac{1}{N-1}\sum_{j\neq i}\varepsilon_j, \tag{52}$$

where $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ are independent noises with $\sigma_\varepsilon^2 = \frac{N-1}{8N^2}\sigma^2$.

The symmetric Gaussian information structure (52) obtains the same allocation rule as the optimal selective informing by carefully coordinating the players' individual noises.

Comparing the allocation under the optimal information structures with the allocations under no information and full information, we observe that the total investment responds to the project quality more than under no information, $1/2 > 0$ but less than under full information $1/2 < N/(N+1)$. Thus, the information design is used to guide the investment but to a limited extent to avoid investment congestion.

Importantly, the information control avoids rent dissipation observed under the no- and full- information benchmarks as the number of players grows to infinity. Indeed, the optimal designer's payoff equals

$$v^* = \mathbb{E}\left[\left(\theta - \frac{\theta}{2} - \frac{N-1}{N+1}\mathbb{E}[\theta]\right)\left(\frac{\theta}{2} + \frac{N-1}{N+1}\mathbb{E}[\theta]\right)\right] = \frac{N}{(N+1)^2}\mathbb{E}^2[\theta] + \frac{1}{4}\mathbb{V}[\theta]. \quad (53)$$

As such, the project's total profit stays above $\mathbb{V}[\theta]/4$ and converges to this level as $N \to \infty$. The limit payoff naturally decreases in the congestion rate r. This payoff does not depend on the individual investment cost c and increases in the variance of the project quality. Under optimal information structure, the riskier projects lead to higher realized profits even if they have the same expected quality.

4.4 Discussion

Our analysis showcased several notable features of optimal information structures in multiplayer games. In this section, we discuss some of these features in detail.

Information Multiplicity One feature discovered in Sections 4.2 and 4.3 is the multiplicity of optimal information structures. This multiplicity is substantive. Different optimal information structures induce different allocation rules and lead to different players' payoffs.

To the best of our knowledge, such multiplicity has rarely been observed in the existing literature on information design and Bayesian persuasion. One way to explain this is to note that despite their multiplicity, the optimal information structures induce the same aggregate action behavior. In "small-scale" settings, which have been most studied in the literature to date, e.g., with a single receiver, or binary receivers and binary actions, the scope in which the same aggregate behavior can be implemented by information in different ways is limited and often translated into a unique optimal information structure. In contrast, in the "large-scale" settings that we study, with multiple receivers and a continuum of actions and states, there are more ways to implement the same aggregate behavior.

This observation suggests that in "large-scale" settings, one can possibly encounter the multiplicity of optimal information structures driven by the multitude of ways to implement an optimal behavior of some coarse statistic of players' actions.

Extraneous Noise An open question in Bayesian persuasion literature is when does an optimal information structure feature an extraneous noise and when does it not, i.e., when it is a deterministic function of a state.[22] Our analysis presented both of these cases. The certifiably optimal Gaussian structures in Section 4.1 did not feature an extraneous noise whereas the certifiably optimal Gaussian structures of Sections 4.2 and 4.3 did.

One broad conjecture is that extraneous uncertainty is introduced by the designer to compensate for the lack of prior uncertainty captured by the state. We show some support for this conjecture but specify that one should be careful regarding exactly what the "lack" of prior uncertainty means. A naive approach would be to measure the richness of prior uncertainty by the state variance: if the variance is low, then extraneous noise is introduced to increase the behavior volatility. However, this approach overlooks the Bayesian nature of belief updating. When the information structure becomes more noisy, rational players take noise into account and, in fact, their posterior beliefs generally become *less* volatile, because each signal is less informative about the state. Accordingly, in the optimal information structures that we found, the qualitative presence of extraneous noise does not depend on the absolute value of the state variance.

Instead, we suggest measuring the richness of prior uncertainty by its effective dimensionality, i.e., the number of state components that are not perfectly correlated. If this dimensionality is low, then extraneous noises may need to be introduced. In Section 4.1, the number of state components equals the number of players and the optimal Gaussian information structure does not require extraneous noise; in Section 4.2, the number of state components is smaller than the number of players, and the optimal Gaussian structure features extraneous noise.

This view is further corroborated by the following observation which we formally articulate in Appendix B.3. In the co-movement setting in Section 4.2.2, if the state had several components that were not perfectly correlated and each player cared about an individual state component, then there would exist an optimal Gaussian information structure that would be a deterministic function of the state. This deterministic structure would, however, converge to the Gaussian stochastic structure that we found. Intuitively, as long as the Gaussian state components are not perfectly correlated, one can extract the relevant individual noises from the states themselves and then successfully use them to implement the optimal behavior. In other words, our analysis suggests that what matters for the presence of exogenous noise is not the variation of each state

[22]For a recent example of the latter, see Candogan and Strack (forthcoming).

component but rather the variation across state components.

Infinite Economies In Section 4.2, under certifiably optimal information structures, the players' aggregate action is a deterministic function of the state. This feature is achieved with finitely many players by either selective informing or precise coordination of their individual noises. In contrast, in games with infinitely many players, the same feature can appear due to the law of large numbers even if the individual noises are independent.

For concreteness, consider the polarization setting of Section 4.2.1 with a normally distributed state. Select an optimal symmetric Gaussian information structure (38) and consider its limit as the number of players goes to infinity. As $N \to \infty$, the correlated components of the individual noises vanish by the law of large numbers, and in the limit, each player is provided with a conditionally independent estimate of the state. The optimal aggregate behavior is ensured by the large population size rather than by noise correlation. This analysis suggests that the Gaussian information structures with independent noises commonly considered in the literature on infinite economies, e.g., by Angeletos and Pavan (2007) and Bergemann and Morris (2013), may indeed be optimal among all information structures. However, as our analysis clarifies, there may exist alternative optimal information structures that induce the same aggregate behavior, and moreover, independent noises are likely to not be optimal in finite economies (Section 3.4).

5 Conclusion

In this paper, we developed a universal solution method for concave information-design problems. This method builds upon the duality between information design and adversarial contracting. We illustrated the power and tractability of the solution method in quadratic environments with a normally distributed state. Our applications offer insights into the determinants of price coordination in markets, the limits of belief polarization, and the disclosure practices that encourage investment. Along the way, we provide justification for the use of Gaussian information structures from the optimality perspective. Overall, our analysis paves the way and provides tools to study information design in "large-scale" games.

References

Aliprantis, C. D. and K. C. Border (2006): *Infinite Dimensional Analysis*, Springer.

Angeletos, G.-M. and A. Pavan (2007): "Efficient Use of Information and Social Value of Information," *Econometrica*, 75, 1103–1142.

——— (2009): "Policy with Dispersed Information," *Journal of the European Economic Association*, 7, 11–60.

Arieli, I., Y. Babichenko, F. Sandomirskiy, and O. Tamuz (2021): "Feasible Joint Posterior Beliefs," *Journal of Political Economy*, 129, 2546–2594.

Bergemann, D., T. Heumann, and S. Morris (2015): "Information and Volatility," *Journal of Economic Theory*, 158, 427–465.

——— (2021): "Information, Market Power, and Price Volatility," *The RAND Journal of Economics*, 52, 125–150.

Bergemann, D. and S. Morris (2013): "Robust Predictions in Games with Incomplete Information," *Econometrica*, 81, 1251–1308.

——— (2016): "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, 11, 487–522.

——— (2019): "Information Design: A Unified Perspective," *Journal of Economic Literature*, 57, 44–95.

Brooks, B. A. and S. Du (2020): "Optimal Auction Design with Common Values: An Informationally-Robust Approach," *Working paper.*

Cai, Y., N. R. Devanur, and S. M. Weinberg (2019): "A Duality-based Unified Approach to Bayesian Mechanism Design," *SIAM Journal on Computing.*

Candogan, O. and P. Strack (forthcoming): "Optimal Disclosure of Information to a Privately Informed Receiver," *Theoretical Economics.*

Carroll, G. (2017): "Robustness and Separation in Multidimensional Screening," *Econometrica*, 85, 453–488.

Cieslak, A., S. Malamud, and A. Schrimpf (2021): “Optimal Transport of Information,” *Working paper.*

Daskalakis, C., A. Deckelbaum, and C. Tzamos (2017): “Strong Duality for a Mutliple Good Monopolist,” *Econometrica*, 85, 735–767.

Dizdar, D. and E. Kováč (2020): “A Simple Proof of Strong Duality in the Linear Persuasion Problem,” *Games and Economic Behavior*, 122, 407–412.

Doval, L. and A. Smolin (2021): “Information Payoffs: An Interim Perspective,” *arXiv preprint arXiv:2109.03061.*

Du, S. (2018): “Robust Mechanisms under Common Valuation,” *Econometrica*, 86, 1569–1588.

Dworczak, P. and A. Kolotilin (2019): “The Persuasion Duality,” *Working paper.*

Dworczak, P. and G. Martini (2019): “The Simple Economics of Optimal Persuasion,” *Journal of Political Economy*, 127, 1993–2048.

Galichon, A. (2018): *Optimal Transport Methods in Economics*, Princeton University Press.

Galperti, S., A. Levkun, and J. Perego (2021): “The Value of Data Records,” *Working paper.*

Galperti, S. and J. Perego (2018): “A Dual Perspective on Information Design,” *Working Paper.*

Kamenica, E. (2019): “Bayesian Persuasion and Information Design,” *Annual Review of Economics.*

Kamenica, E. and M. Gentzkow (2011): “Bayesian Persuasion,” *American Economic Review*, 101, 2590–2615.

Kolotilin, A. (2017): “Optimal Information Disclosure: A Linear Programming Approach,” *Theoretical Economics*, forthcoming.

Krishna, V. and H. Sonnenschein (1990): “Duality in Consumer Theory,” in *Preferences, Uncertainty and Optimality*, ed. by J. Chipman, D. McFadden, and M. Richter, Westview Press Boulder, chap. 2, 44–55.

Malamud, S. and A. Schrimpf (2021): "Persuasion by Dimension Reduction," *Working paper.*

Myerson, R. (1983): "Mechanism Design by an Informed Principal," *Econometrica*, 51, 1767–1797.

Neyman, A. (1997): "Correlated Equilibrium and Potential Games," *International Journal of Game Theory*, 26, 223–227.

Rayo, L. and I. Segal (2010): "Optimal Information Disclosure," *Journal of Political Economy*, 118, 949–987.

Rosen, J. (1965): "Existence and Uniqueness of Equilibrium Points for Concave N-Person Games," *Econometrica*, 33, 520–534.

Tamura, W. (2018): "Bayesian Persuasion with Quadratic Preferences," *Working paper.*

Taneva, I. (2019): "Information Design," *American Economic Journal: Microeconomics*, 11, 151–85.

Ui, T. (2020): "LQG Information Design," *Working paper.*

Villani, C. (2003): *Topics in Optimal Transportation*, 58, American Mathematical Society.

A Omitted Proofs and Derivations

Proof of Theorem 2. We show that under the conditions of Theorem 2, allocation rule (20) is implementable both by information in the primal problem and by incentives in the dual problem.

The implementability of the allocation rule in the information-design problem is captured by the system of first-order conditions. An allocation rule $\alpha : \Omega \to \Delta(A)$ is incentive compatible if and only if

$$\mathbb{E}[b_i + B_{i\cdot}\omega - C_{i\cdot}a \mid a_i] = 0 \quad \forall i = 1, \ldots, N, a_i \in A_i, \tag{54}$$

for all actions a_i recommended under α.

Because conditions (54) must hold for all a_i, they must also hold, on average; thus,

$$\mathbb{E}_{\mu_0}[b_i + B_{i\bullet}\omega - C_{i\bullet}a] = 0 \quad \forall i = 1, \ldots, N. \tag{55}$$

The recommended actions are linear in the state components. Since $\mathbb{E}[\omega_k] \equiv 0$, the constant term of the linear information structure is uniquely pinned down as

$$a_0 = C^{-1}b. \tag{56}$$

The recommended action profile a^* and the variable $C_{i\bullet}a - B_{i\bullet}\omega - b_i$ are jointly normally distributed. Hence, the sufficient condition for (54) is that a_i and $C_{i\bullet}a - B_{i\bullet}\omega - b_i$ are uncorrelated for all i:

$$\begin{aligned}\operatorname{cov}(C_{i\bullet}a^* - B_{i\bullet}\omega - b_i, a_i^*) &= \mathbb{E}[(C_{i\bullet}R - B_{i\bullet})\omega R_{i\bullet}\omega] = \mathbb{E}[(C_{i\bullet}R - B_{i\bullet})\omega\omega^T R_{i\bullet}^T] \\ &= (C_{i\bullet}R - B_{i\bullet})\mathbb{E}[\omega\omega^T]R_{i\bullet}^T = (C_{i\bullet}R - B_{i\bullet})\mathit{\Sigma} R_{i\bullet}^T = 0. \end{aligned} \tag{57}$$

Conditions (56) and (57) together form condition (i) of the theorem.

It remains to show that under the second set of conditions, the linear allocation rule is implementable in the dual adversarial problem. The dual payoff is

$$\begin{aligned} u^\lambda(a,\omega) &= v(a,\omega) - \sum_i \lambda_i(a_i)(C_{i\bullet}a - B_{i\bullet}\omega - b_i) && (58) \\ &= a^T(\hat{b} + \hat{B}\omega) - \frac{1}{2}a^T\hat{C}a - \lambda^T(a)(Ca - B\omega - b). && (59) \end{aligned}$$

Under the conditions of the theorem, the linear allocation rule can be implemented in the dual problem by a contract $\lambda = x_0 + x * a$, where "$*$" is the Hadamard product. Given this contract, the optimal best response of the dual agent solves, at any state $\omega \in \mathit{\Omega}$:

$$\begin{aligned} &\max_{a\in A} \; a^T(\hat{b} + \hat{B}\omega) - \frac{1}{2}a^T\hat{C}a - (x_0 + x * a)^T(Ca - B\omega - b), \\ &\max_{a\in A} \; a^T\left(\hat{b} + D(x)b - C^T x_0 + (\hat{B} + D(x)B)\omega\right) - \frac{1}{2}a^T\left(\hat{C} + 2D(x)C\right)a + x_0^T(B\omega + b), \end{aligned}$$

where $D(x)$ is a diagonal matrix with $D(x)_{ii} = x_i$. This is a quadratic optimization problem. Since matrix $\hat{C} + 2D(x)C$ is positive definite, the agent's best response equals

$$a^*(\omega) = (\hat{C} + 2D(x)C)^{-1}(\hat{b} + D(x)b - C^T x_0) + (\hat{C} + 2D(x)C)^{-1}(\hat{B} + D(x)B)\omega. \tag{60}$$

The conditions of the theorem ensure that x implements the best response with the responsiveness matrix R. It remains to construct x_0 to capture the constant vector $a_0 = C^{-1}b$, i.e.,

$$(\hat{C} + 2D(x)C)^{-1}(\hat{b} + D(x)b - C^T x_0) = C^{-1}b$$
$$C(\hat{C} + 2D(x)C)^{-1}C^T x_0 = c_0,$$

for some vector c_0. However, $(\hat{C} + 2D(x)C)^{-1}$ is positive definite and, by the maintained assumption, C is positive semidefinite. Hence, $C(\hat{C} + 2D(x)C)^{-1}C^T$ is invertible, and one can always find x_0 that implements the targeted linear allocation rule.

Proof of Proposition 4 Given Proposition 1, it suffices to show that allocation (37) is implementable by incentives. The dual payoff is

$$u^{\lambda}(a,\omega) = v(a,\omega) - \sum_{i=1}^{N} \lambda_i(a_i)\dot{u}_i(a,\omega) = \sum_{i,j}(a_i - a_j)^2 + 2\sum_i \lambda_i(a_i)(a_i - \omega). \tag{61}$$

Set the contract to be $\lambda(a_i) = -N(a_i - \mathbb{E}[\omega])$. The agent's payoff can be rewritten as:

$$\sum_{i,j}(a_i - a_j)^2 - 2N\sum_i(a_i - \mathbb{E}[\omega])(a_i - \omega) = -2(\sum_i a_i - N\mathbb{E}[\omega])(\sum_i a_i - N\omega). \tag{62}$$

For any state ω, the agent's payoff is concave and quadratic in the aggregate action $\sum_{i=1}^{N} a_i$ with a unique optimal aggregate action satisfying $\sum_{i=1}^{N} a_i(\omega) \equiv \frac{N}{2}(\omega + \mathbb{E}[\omega])$. Hence, allocation (37) is implementable by incentives, and the result follows.

Proof of Corollary 4 Note that by construction of the optimal information structure, the noise variance is equal to the variance of the state component:

$$\mathbb{V}[\frac{\omega + \bar{\omega}}{2}] = \mathbb{V}[\varepsilon_i - \frac{\sum_{j\neq i}\varepsilon_j}{N-1}] = \frac{\sigma^2}{4}. \tag{63}$$

Hence, by Gaussian belief updating, the information structure is obedient:

$$\mathbb{E}[\frac{\omega + \bar{\omega}}{2} \mid a_i] = \frac{1}{2}(\mathbb{E}[\frac{\omega + \bar{\omega}}{2}] + a_i) = \frac{1}{2}(\bar{\omega} + a_i), \tag{64}$$

and, consequently, $\mathbb{E}[\omega \mid a_i] = a_i$.

At the same time, the aggregate action satisfies the optimality condition (37):

$$\sum_{i=1}^{N} a_i(\omega) = \sum_{i=1}^{N} \frac{\omega+\bar{\omega}}{2} + \sum_{i=1}^{N}(\varepsilon_i - \frac{\sum_{j\neq i}\varepsilon_j}{N-1}) = N\frac{\omega+\bar{\omega}}{2} + \sum_{i=1}^{N}\varepsilon_i - \sum_{i=1}^{N}\varepsilon_i = \frac{N}{2}(\omega + \mathbb{E}[\omega]). \tag{65}$$

The result thus follows from Proposition 4.

Proof of Proposition 5 Consider a contract with $\lambda_i(a_i) = -\frac{\rho}{2N^2}a_i$. The corresponding dual payoff is

$$v(a,\omega) - \sum_i \lambda_i(a_i)\dot{u}_i(a,\omega) = \frac{\sum_{i=1}^{N} a_i}{N}\omega - \rho\frac{\sum_{i=1}^{N}\sum_{j\neq i} a_i a_j}{N^2} + 2\sum_{i=i}^{N}\lambda_i(a_i)(a_i-\omega) \tag{66}$$

$$= \frac{\sum_{i=1}^{N} a_i}{N}\frac{N+\rho}{N}\omega - \rho\left(\frac{\sum_{i=1}^{N} a_i}{N}\right)^2. \tag{67}$$

Thus, the agent's best response in state ω is any action profile a that satisfies

$$\frac{\sum_{i=1}^{N} a_i}{N} = \frac{N+\rho}{2\rho N}\omega. \tag{68}$$

The proposed contract implements the candidate allocation rule and, by Proposition 1, certifies the optimality of the information structure that induces it.

Proof of Corollary 6 Allocation rule (41) is obedient, $\mathbb{E}[\omega|a_i] \equiv a_i$. Indeed, $a_i(\omega)$ and ω are jointly normally distributed. Therefore, the obedience reduces to $\mathbb{E}[\omega] = \mathbb{E}[a_i]$, satisfied since $\mathbb{E}[\omega] = 0$, and to

$$\mathbb{E}[a_i(\omega - a_i)] = 0. \tag{69}$$

By the definition of the recommended actions:

$$\mathbb{E}[a_i(\omega - a_i)] = \left(\frac{1}{2\rho} + \frac{1}{2N}\right)\left(\frac{1}{2\rho} + \frac{1}{2N} - 1\right)\mathbb{V}[\omega] + \mathbb{V}[\varepsilon_i - \frac{1}{N-1}\sum_{j\neq i}\varepsilon_j], \tag{70}$$

and

$$\mathbb{V}[\varepsilon_i - \frac{1}{N-1}\sum_{j\neq i}\varepsilon_j] = \frac{N}{N-1}\sigma_\varepsilon. \tag{71}$$

Hence, the obedience is satisfied for the chosen variance σ_ε^2.

At the same time, the aggregate action satisfies the optimality condition (40):

$$\sum_{i=1}^N a_i(\omega) = \sum_{i=1}^N \left(\frac{1}{2\rho}+\frac{1}{2N}\right)\omega + \sum_{i=1}^N(\varepsilon_i - \frac{\sum_{j\neq i}\varepsilon_j}{N-1}) \tag{72}$$

$$= N\left(\frac{1}{2\rho}+\frac{1}{2N}\right)\omega + \sum_{i=1}^N \varepsilon_i - \sum_{i=1}^N \varepsilon_i = N\left(\frac{1}{2\rho}+\frac{1}{2N}\right)\omega. \tag{73}$$

The result thus follows from Proposition 5.

Derivations for Section 4.3

No Information: Under no additional information, each player's action cannot depend on the state and is thus uniquely determined by best-response condition (44).

Full Information: For any θ, the ensuing game admits a strictly concave potential $\Psi(a,\theta) = (\theta - \frac{A}{2})A - \sum_{i=1}^N a_i^2$ and thus has a unique equilibrium. Parameterize a symmetric linear strategy profile as

$$a_i(\theta) = k_0 + k_1\theta. \tag{74}$$

The best-response condition (44) can be rewritten as

$$a_i(\theta) = -\frac{N-1}{2}k_0 + \frac{1-k_1(N-1)}{2}\theta, \tag{75}$$

that pins down the equilibrium parameters at $k_0 = 0$ and $k_1 = \frac{1}{N+1}$.

Proof of Proposition 6 Given Proposition 1, it suffices to show that allocation (51) can be implemented by incentives. The dual payoff is

$$u^\lambda(a,\theta) = v(a,\omega) - \sum_{i=1}^N \lambda_i(a_i)\dot{u}_i(a,\theta) = (\theta - A)A - \sum_{i=1}^N \lambda_i(a_i) r(\theta - A - a_i). \tag{76}$$

Set the contract to be $\lambda(a_i) = \frac{N-1}{r(N+1)^2}\mathbb{E}[\theta]$. The agent's payoff can be rewritten as:

$$u^\lambda(a,\theta) = (\theta - A)A - \frac{N-1}{(N+1)^2}\mathbb{E}[\theta](N(\theta - A) - A). \tag{77}$$

For any normalized state θ, the payoff is concave and quadratic in A with a unique maximizer satisfying

$$A(\theta) = \frac{1}{2}\left(\theta + \frac{N-1}{N+1}\mathbb{E}[\theta]\right), \tag{78}$$

that corresponds to (51).

Proof of Corollary 7 Under selective informing, each uninformed player i plays the same action irrespectively of a state, $a_i = a_i^{NI} = \frac{1}{N+1}\mathbb{E}[\theta]$. Thus, a fully informed player optimally responds to the normalized state according to best-response condition (44) as

$$a_i(\theta) = \frac{\theta}{2} - \frac{N-1}{2}\frac{\mathbb{E}[\theta]}{N+1}, \tag{79}$$

leading to total investment (51). The result thus follows by Proposition 6.

Proof of Corollary 8 If obedient, information structure (52) results in total investment (51). It if left to demonstrate obedience.

Because a_i and $\theta - A - a_i$ are jointly normal, the obedience condition $\mathbb{E}[\theta - A - a_i|a_i] = 0$ is equivalent to $\mathbb{E}[\theta - A - a_i] = 0$ and $\mathbb{E}[a_i(\theta - A - a_i)] = 0$. These conditions are easily verified:

$$\mathbb{E}[\theta - A - a_i] = \mathbb{E}[\theta - (N+1)(\frac{\theta}{2N} + \frac{(N-1)\mu}{2N(N+1)})] = \mathbb{E}[\frac{(N-1)(\theta-\mu)}{2N}] = 0, \tag{80}$$

and,

$$\begin{aligned}\mathbb{E}[a_i(\theta - A - a_i)] &= \mathbb{E}[\frac{\theta}{2N}\frac{(N-1)(\theta-\mu)}{2N} - (\varepsilon_i - \frac{\sum_{j\neq i}\varepsilon_j}{N-1})^2] &(81)\\ &= \frac{(N-1)\sigma^2}{4N^2} - 2\sigma_\varepsilon^2 = 0. &(82)\end{aligned}$$

The result thus follows by Proposition 6.

B Supplementary Appendix

B.1 Derivations for Section 4.1

The firms' profits are equal to

$$u_i(a,\theta) = a_i(\theta_i + \eta a_i + \xi a_{-i}) - c(\theta_i + \eta a_i + \xi a_{-i})^2,$$

so the derivative with respect to their own actions is:

$$\begin{aligned}\frac{\partial u_i(a,\theta)}{\partial a_i} &= \theta_i(1-2c\eta) + 2a_i\eta(1-c\eta) + a_{-i}\xi(1-2c\eta),\\ &= \mu_i(1-2c\eta) + (\theta_i - \mu_i)(1-2c\eta) + 2a_i\eta(1-c\eta) + a_{-i}\xi(1-2c\eta).\end{aligned}$$

By comparison with the F.O.C. of (55), we recover the parameters of Section 4:

$$b = (1-2c\eta)\bar{\theta}, \quad B = (1-2c\eta)I, \quad C = \begin{pmatrix} -2\eta(1-c\eta) & -\xi(1-2c\eta) \\ -\xi(1-2c\eta) & -2\eta(1-c\eta) \end{pmatrix}.$$

Regarding the designer's payoff, consumer surplus can be written as:

$$CS(a,\theta) = -a^T\theta - \frac{1}{2}a^TWa = -\bar{\theta} - a^T(\theta - \bar{\theta}) - \frac{1}{2}a^TWa.$$

Comparing it with the payoff function (18), we recover

$$\hat{b}_{CS} = -\bar{\theta}, \quad \hat{B}_{CS} = -I, \quad \hat{C}_{CS} = W = \begin{pmatrix} \eta & \xi \\ \xi & \eta \end{pmatrix}.$$

Similarly, the total profits can be written as

$$\begin{aligned}\Pi(a,\theta) &= a^T(\theta + Wa) - c(\theta + Wa)^T(\theta + Wa)\\ &= -c\theta^T\theta + a^T(1-2cW)\theta - \frac{1}{2}a^T(-2W + 2cW^2)a\\ &\approx a^T(1-2cW)\bar{\theta} + a^T(1-2cW)(\theta - \bar{\theta}) - \frac{1}{2}a^T(-2W + 2cW^2)a,\end{aligned}$$

where the last line ignores the action-independent term $-c\theta^T\theta$. Comparing it with the payoff function (18), we recover:

$$\hat{b}_{\Pi} = (I - 2cW)\bar{\theta}, \quad \hat{B}_{\Pi} = I - 2cW, \quad \hat{C}_{\Pi} = -2W + 2cW^2.$$

For any given $\delta \in [0,1]$, the parameters of the designer's problem are the weighted averages:

$$\hat{b} = \delta \hat{b}_{CS} + (1-\delta)\hat{b}_{\Pi},$$
$$\hat{B} = \delta \hat{B}_{CS} + (1-\delta)\hat{B}_{\Pi},$$
$$\hat{C} = \delta \hat{C}_{CS} + (1-\delta)\hat{C}_{\Pi}.$$

Direct Price Control: The designer's first-order condition is

$$\hat{B}\theta - \hat{C}a = 0,$$

which results in the first-best responsiveness matrix

$$R^{FB} = \hat{C}^{-1}\hat{B}.$$

The threshold value δ^{FB} is the one that equalizes the determinant of $\hat{C}$ to zero.

Full Information and No Information: Equilibrium pricing behavior satisfies and can be derived from the system of first-order conditions

$$\mathbb{E}_{\mu}[q_i(a_i, a_{-i}, \theta) + \frac{\partial q_i(a_i, a_{-i}, \theta)}{\partial a_i}(a_i - 2ca_i)] = 0, \quad i = 1, 2, \tag{83}$$

after substituting the linear form (22) of demand function q_i and setting the belief μ equal to μ_0 for no-information equilibrium and equal to the belief concentrated on θ for full-information equilibrium.

Optimal Information Structure: As discussed in the main text, the optimal information structure can be certified by a contract with certifying parameters $(x_1, x_2) = x$. By Theorem 2, the parameter x and the corresponding responsiveness matrix $R(x)$ must satisfy the conditions $(C_{i\bullet}R - B_{i\bullet})\Sigma R_{i\bullet}^T = 0$ and $R = (\hat{C} + 2D(x)C)^{-1}(\hat{B} + D(x)B)$. Plugging in the parameters of the differentiated Bertrand competition, we obtain the certifying condition $f(x) = 0$, where $f(x)$ is the following polynomial:

$$f(x) = b_0 + b_1 x + b_2 x^2 + b_3 x^3 + b_2 x^4,$$

$$
\begin{aligned}
b_0 &= -8c^3(\delta-1)^3(\xi-\eta)(\eta+\xi)(\delta\eta^4+(\delta-1)\xi^4+3(2\delta-1)\eta^2\xi^2) \\
&\quad + 4c^2(\delta-1)^2\eta(\delta(5-8\delta)\eta^4+(\delta(8\delta-11)+4)\xi^4-2(\delta(16\delta-19)+6)\eta^2\xi^2) \\
&\quad + 2c(3\delta-2)(\delta-1)(\delta(7\delta-4)\eta^4+(\delta-1)\delta\xi^4+(3\delta(8\delta-9)+8)\eta^2\xi^2) \\
&\quad + (2\delta-1)(-(2-3\delta)^2)\eta(\delta\eta^2+(3\delta-2)\xi^2),
\end{aligned}
$$

$$
\begin{aligned}
b_1 &= 32c^5(\delta-1)^3(\eta^2+\xi^2)(\eta^3-\eta\xi^2)^2 \\
&\quad - 16c^4(\delta-1)^2\eta(\xi-\eta)(\eta+\xi)((8-11\delta)\eta^4+2(\delta-1)\xi^4+(2-3\delta)\eta^2\xi^2) \\
&\quad + 8c^3(\delta-1)(\xi-\eta)(\eta+\xi)((-42\delta^2+64\delta-25)\eta^4+(\delta-1)^2\xi^4+\delta(2-3\delta)\eta^2\xi^2) \\
&\quad + 4c^2((\delta(2(78-31\delta)\delta-133)+38)\eta^5+(\delta-1)(\delta(\delta+7)-6)\eta\xi^4 \\
&\quad + (\delta(5\delta(25\delta-58)+226)-60)\eta^3\xi^2)-2c((\delta((79-17\delta)\delta-88)+28)\eta^4+(\delta-1)\delta(5\delta-4)\xi^4 \\
&\quad + 2(\delta(\delta(102\delta-211)+146)-34)\eta^2\xi^2)+(3\delta-2)\eta((\delta(7\delta+4)-4)\eta^2+(\delta(41\delta-52)+16)\xi^2),
\end{aligned}
$$

$$
\begin{aligned}
b_2 &= -2((\eta-1)\eta^3(\delta^2(8(\eta-2)\eta(10(\eta-2)\eta+13)+29)-4\delta(\eta-1)(2\eta-1)(20(\eta-2)\eta+13) \\
&\quad + 20(2\eta^2-3\eta+1)^2)+2(\delta-1)(1-2\eta)^2r^4(2(\delta-1)\eta^2-\delta+2\eta) \\
&\quad + \eta r^2(\delta^2(\eta(4\eta(\eta(8(11-3\eta)\eta-145)+135)-271)+56) \\
&\quad + 2\delta(\eta-1)(4\eta(\eta(8\eta(3\eta-7)+63)-37)+37)-8(\eta-1)(3(\eta-1)\eta+1)(4(\eta-1)\eta+3))),
\end{aligned}
$$

$$
\begin{aligned}
b_3 &= -16\eta(1-2c\eta)^2(c\eta-1)(\xi^2(4c^2(\delta-1)\eta^2+2c(2-3\delta)\eta+3\delta-2) \\
&\quad - 2\eta^2(c\eta-1)(2c(\delta-1)\eta-3\delta+2)),
\end{aligned}
$$

$$
b_4 = -8\eta(1-2c\eta)^2(c\eta-1)(4\eta^2(c\eta-1)^2-\xi^2(1-2c\eta)^2).
$$

For the parameters of the numerical example, the polynomial becomes

$$
\begin{aligned}
f(x) &= (2996\delta^4-7880\delta^3+7490\delta^2-3000\delta+414)+(6728\delta^3-20948\delta^2+20234\delta-6210)x \\
&\qquad + (-52780\delta^2+88084\delta-36368)x^2+(77184\delta-62208)x^3-31680x^4,
\end{aligned}
$$

and the condition $f(x)=0$ can be solved in radicals for any $\delta\in[0,1]$. The solution x that makes $\hat{C}+2D(x)C$ positive definite is a certifying parameter. Calculations show that such x is unique for all $\delta\neq\delta^{cr}$, and its value is plotted in Figure 4. The value δ^{cr}

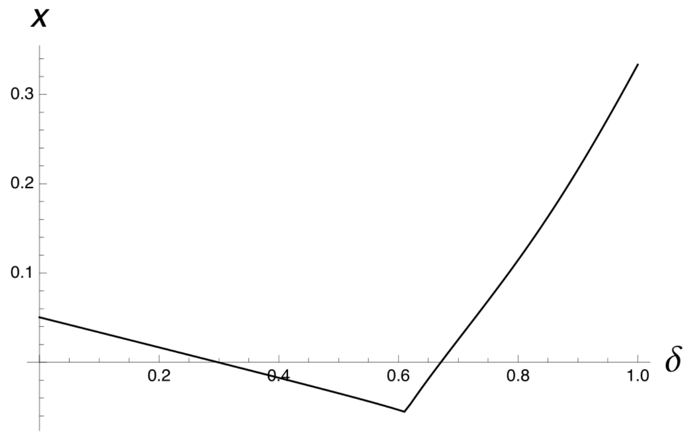

Figure 4: Parameter x that certifies an optimal information structure plotted as a function of the consumer surplus weight δ.

is the one that makes the determinant of matrix $\hat{C} + 2D(x(\delta))C$ equal to zero:

$$\delta^{cr} = \frac{2(c(|\xi| + \eta)(-2c\eta\,|\xi| + |\xi| + \eta(-2c\eta + 3)) - \eta)}{-\,|\xi|\,(-2c\eta(-4c\eta + 5) + 1) - \eta(4c - \eta(-c\eta + 2) + 5) + 2c\xi^2(-2c\eta + 1)} = \frac{11}{18}. \tag{84}$$

B.2 Strong Duality

Notation Denote by $\mathbb{R}$ the set of real numbers and by $\mathbb{N}$ the set of strictly positive integers. For a Polish space X, denote by $\mathscr{P}(X)$ the space of its measurable subsets, by $M(X)$ the space of Radon measures on X, by $\Delta(X) \subseteq M(X)$ the space of probability measures on X, by $B(X)$ the space of real-valued bounded functions on X, by $C(X)$ the space of measurable real-valued continuous functions on X equipped with the uniform norm $\|\cdot\|_\infty$, $\|f\|_\infty \triangleq \sup_{x\in X} |f(x)|$. For an arbitrary collection of sets $\{X_i\}$, denote their product set by $\times_i X_i$.

In this section, we prove the applicability of the certification method for a subclass of concave games, i.e., that (i) the duality gap is equal to zero, $V^D = V^P$, and (ii) solutions to both primal and dual problems exist.

Assumption 2. (Compactness) *The state space Ω is finite. For each $i = 1, \dots, N$, there exists a convex compact subset $\hat{A}_i \subseteq A_i$ such that all actions $a_i \notin \hat{A}_i$ are strictly dominated for player i.*

Assumption 3. (Responsiveness) *There exists $\varepsilon > 0$ such that for each $i = 1, \dots, N$, and $a_i \in A_i$: (i) there exist $\omega^- \in \Omega$, $a_{-i}^- \in A_{-i}$ such that $\dot{u}_i(a_i, a_{-i}^-, \omega^-) < -\varepsilon$ and*

$\dot{u}_j(a_i, a_{-i}^-, \omega^-) = 0$ *for all* $j \neq i$, *and (ii) there exist* $\omega^+ \in \Omega$, $a_{-i}^+ \in A_{-i}$ *such that* $\dot{u}_i(a_i, a_{-i}^+, \omega^+) > \varepsilon$ *and* $\dot{u}_j(a_i, a_{-i}^+, \omega^+) = 0$ *for all* $j \neq i$.

Assumption 3 implies responsiveness of players' actions to the state. It requires that each action may be "too high" in some states and "too low" in others, under complete information if all others respond optimally. This assumption guarantees that in the dual problem, the designer would never use over-powered incentives; hence, the domain of contracts can be bounded.

Theorem 3. (Strong Duality) *If Assumption 2 holds, then the optimal value of the information-design problem (6) is equal to the optimal value of the dual adversarial-contracting problem (11):*

$$V^P = V^D. \tag{85}$$

Moreover, in this case, an optimal value in (6) is achieved by some information structure. If, in addition, Assumption 3 holds, then an optimal value in (11) is achieved by some contract.

Proof: Define auxiliary functions $\phi, \psi : B(A \times \Omega) \to \mathbb{R} \cup \{+\infty\}$ as follows:

$$\phi(f) \triangleq \begin{cases} 0, & \text{if } f(a,\omega) \geq v(a,\omega)\ \forall\, a \in A, \omega \in \Omega, \\ +\infty, & \text{o/w.} \end{cases}$$

$$\psi(f) \triangleq \inf_{\lambda \in \times_i B(A_i), \gamma \in B(\Omega)} \begin{cases} \int_\Omega \gamma(\omega) \mathrm{d}\mu_0, & \text{if } f(a,\omega) = \sum_{i=1}^N \lambda_i(a_i)\dot{u}_i(a,\omega) + \gamma(\omega) \\ & \qquad \forall\, a \in A, \omega \in \Omega, \\ +\infty, & \text{o/w.} \end{cases}$$

In Lemma 1, we show that (i) $\phi(f)$ and $\psi(f)$ are convex. Moreover, by Assumption 2, $v(a,\omega)$ is bounded from above by $\overline{V} \triangleq \sup_{a\in A, \omega \in \Omega} v(a,\omega) < +\infty$. Hence, there exists $f_0(a,\omega) = \overline{V} + 1$ such that (ii) $\phi(f_0) < +\infty$ (as $\phi(f_0) = 0$), (iii) $\psi(f_0) < +\infty$ (as $\psi(f_0) \leq f_0$ since one can set $\lambda \equiv 0$ and $\gamma(\omega) \equiv \overline{V} + 1$), and (iv) ϕ is continuous at f_0 (as $\phi(f) \equiv 0$ for all f with $\|f - f_0\|_\infty < 1$). Consequently, by the Fenchel–Rockafellar duality (Villani (2003), Theorem 1.9), there exists a solution to $\sup_{\pi \in M(A\times\Omega)}(-\phi^*(-\pi) - \psi^*(\pi))$ and:

$$\max_{\pi \in M(A\times\Omega)} (-\phi^*(-\pi) - \psi^*(\pi)) = \inf_{f \in C(A\times\Omega)} (\phi(f) + \psi(f)), \tag{86}$$

where ϕ^* and ψ^* are Legendre-Fenchel transforms of ϕ and ψ, respectively:

$$\phi^*(\pi) \triangleq \sup_{f \in C(A\times\Omega)} \left(\int f \mathrm{d}\pi - \phi(f) \right),$$
$$\psi^*(\pi) \triangleq \sup_{f \in C(A\times\Omega)} \left(\int f \mathrm{d}\pi - \psi(f) \right),$$

and where we used the fact that $M(A \times \Omega)$ is a topological dual space of $C(A \times \Omega)$ (Aliprantis and Border (2006), Corollary 14.15) since $A \times \Omega$ is a compact metrizable space (Assumption 2).

In Lemma 2, we show that the left-hand side of (86) is in fact equal to V^P:

$$\max_{\pi \in M(A\times\Omega)} (-\phi^*(-\pi) - \psi^*(\pi)) = V^P.$$

Moreover, the right-hand side of (86) is at least as large as V^D:

$$\inf_{f \in C(A\times\Omega)} (\phi(f) + \psi(f)) \geq \inf_{f \in B(A\times\Omega)} (\phi(f) + \psi(f)) = V^D,$$

where the inequality follows from $C(A \times \Omega) \subseteq B(A \times \Omega)$ and the equality follows from the definition of ϕ and ψ. As a result, $V^P \geq V^D$; thus, by Theorem 1, $V^P = V^D$.

Lemma 1. ϕ and ψ are convex.

Proof. ϕ: Towards a contradiction, assume that ϕ is not convex. Then, there exist $f_1, f_2 \in C(A \times \Omega)$ and $\alpha \in (0,1)$ such that $\phi(\alpha f_1 + (1-\alpha) f_2) > \alpha\phi(f_1) + (1-\alpha)\phi(f_2)$. It is possible only if the right-hand side is finite, that is, only if $f_1(a,\omega) \geq v(a,\omega)$ and $f_2(a,\omega) \geq v(a,\omega)$ for all $a \in A, \omega \in \Omega$. However, in that case, $\alpha f_1(a,\omega) + (1-\alpha) f_2(a,\omega) \geq v(a,\omega)$; thus, $\phi(\alpha f_1 + (1-\alpha) f_2) = 0 = \alpha\phi(f_1) + (1-\alpha)\phi(f_2)$, which is a contradiction.

ψ: Take any $f_1, f_2 \in C(A \times \Omega)$. If either $\psi(f_1) = +\infty$ or $\psi(f_2) = +\infty$, then $\psi(\alpha f_1 + (1-\alpha) f_2) \leq \alpha\psi(f_1) + (1-\alpha)\psi(f_2) = +\infty$ for all $\alpha \in (0,1)$. If both $\psi(f_1), \psi(f_2) < +\infty$, then for any $n \in \mathbb{N}$, there exist $\lambda_1^n, \gamma_1^n, \lambda_2^n, \gamma_2^n$ such that for all $a \in A, \omega \in \Omega$:

$$\int_{\omega\in\Omega} \gamma_1^n(\omega) \mathrm{d}\mu_0(\omega) \leq \psi(f_1) + 1/n, \quad f_1(a,\omega) = \sum_{i=1}^N \lambda_{1i}^n(a_i)\dot{u}_i(a,\omega) + \gamma_1^n(\omega),$$
$$\int_{\omega\in\Omega} \gamma_2^n(\omega) \mathrm{d}\mu_0(\omega) \leq \psi(f_2) + 1/n, \quad f_2(a,\omega) = \sum_{i=1}^N \lambda_{2i}^n(a_i)\dot{u}_i(a,\omega) + \gamma_2^n(\omega).$$

Hence, for any $\alpha \in (0,1)$, for all $a \in A, \omega \in \Omega$:

$$\alpha f_1(a,\omega) + (1-\alpha) f_2(a,\omega) = \sum_{i=1}^{N} \lambda_{\alpha i}^n(a_i)\dot{u}_i(a,\omega) + \gamma_\alpha^n(\omega),$$

where $\lambda_{\alpha i}^n(a_i) \triangleq \alpha\lambda_{1i}^n(a_i) + (1-\alpha)\lambda_{2i}^n(a_i)$ and $\gamma_\alpha^n(\omega) \triangleq \alpha\gamma_1^n(\omega) + (1-\alpha)\gamma_2^n(\omega)$. Consequently,

$$\begin{aligned}\psi(\alpha f_1 + (1-\alpha) f_2) &\leq \sum_{\omega\in\Omega} \gamma_\alpha^n(\omega)\mu_0(\omega) \\ &\leq \sum_{\omega\in\Omega} (\alpha\gamma_1^n(\omega) + (1-\alpha)\gamma_2^n(\omega))\mu_0(\omega) \\ &\leq \alpha\psi(f_1) + (1-\alpha)\psi(f_2) + 1/n.\end{aligned}$$

Since this inequality holds for arbitrarily large $n \in \mathbb{N}$, the result follows. □

Lemma 2. $V^P = \max_{\pi\in M(A\times\Omega)}(-\phi^*(-\pi) - \psi^*(\pi))$.

Proof. We can write:

$$\begin{aligned}-\phi^*(-\pi) &= -\sup_{f\in C(A\times\Omega)} \left(\int f \mathrm{d}(-\pi) - \phi(f)\right) \\ &= \inf_{f\in C(A\times\Omega), f\geq v} \left(\int f\mathrm{d}\pi\right) \\ &= \begin{cases} \int v\mathrm{d}\pi, & \text{if } \pi \in \Delta(A\times\Omega), \\ -\infty, & \text{o/w.}\end{cases}\end{aligned}$$

$$\begin{aligned}-\psi^*(\pi) &= -\sup_{f\in C(A\times\Omega)} \left(\int f\mathrm{d}\pi - \psi(f)\right) \\ &= \inf_{f\in C(A\times\Omega)} \left(-\int f\mathrm{d}\pi + \psi(f)\right) \\ &= \begin{cases} 0, & \text{if (7) and (8) hold,} \\ -\infty, & \text{o/w.}\end{cases}\end{aligned}$$

The last line in the derivation of $-\phi^*(-\pi)$ holds because (i) if $\pi \notin \Delta(A\times\Omega)$, then π assigns a negative measure to some set and one can diverge the value of $-\phi^*(-\pi)$ to $-\infty$ by choosing $f \in C(A\times\Omega)$ that assigns arbitrarily large values to that set and (ii) $\pi \in \Delta(A\times\Omega)$, then the infimum is obtained by setting $f \equiv v \in C(A\times\Omega)$. Similarly, to establish the last line in the derivation of $-\psi^*(\pi)$, note that for any given f, by the definition of ψ:

$$\begin{aligned}-\int f\mathrm{d}\pi + \psi(f) &\geq -\int \left(\sum_{i=1}^{N} \lambda_i(a_i)\dot{u}_i(a,\omega) + \gamma(\omega)\right)\mathrm{d}\pi + \int_\Omega \gamma(\omega)\mathrm{d}\mu_0 \\ &= \int_\Omega \gamma(\omega)\mathrm{d}\mu_0 - \int \gamma(\omega)\mathrm{d}\pi - \sum_{i=1}^{N}\int \lambda_i(a_i)\dot{u}_i(a,\omega)\mathrm{d}\pi,\end{aligned}$$

for any λ, γ such that $\sum_{i=1}^N \lambda_i(a_i)\dot{u}_i(a,\omega) + \gamma(\omega) = f(a,\omega)$ for all $a \in A$, $\omega \in \Omega$. If (7) and (8) hold, then $-\int f \mathrm{d}\pi + \psi(f) \geq 0$. Furthermore, the zero value can be achieved by setting $f \equiv 0$. If (7) or (8) do not hold, then it is possible to diverge the value to $-\infty$ by assigning arbitrarily large absolute values, positive or negative, of γ or λ to a set with a non-zero measure such that $f(a,\omega)$ set equal to $\sum_{i=1}^{N} \lambda_i(a_i)\dot{u}_i(a,\omega) + \gamma(\omega)$ is continuous.

As a result, the maximization problem

$$\max_{\pi \in M(A \times \Omega)} (-\phi^*(-\pi) - \psi^*(\pi))$$

is identical to the primal problem (6). The result follows. $\square$

Thus far, we have been able to use the Fenchel–Rockafellar duality to establish the strong duality between the primal and dual problems. To complete the proof of Theorem 3, we must confirm that the solutions to these problems exist under the stated conditions.

Lemma 3. (Existence of Solutions) Solutions to the primal problem (6) and to the dual problem (9) exist.

Proof. For the primal problem (6), equip $\Delta(\Omega \times A)$ with a weak* topology. In this topology, $\Delta(\Omega \times A)$ is compact and the objective and the constraints are continuous due to Assumption 1. The solution then exists by the extreme value theorem. (Alternatively, observe that the solution existence follows directly from Fenchel-Rockafellar duality (86) and the proof of Lemma 2.)

For the dual problem (9), define

$$\underline{\gamma}(\lambda,\omega) \triangleq \sup_{a \in A} (v(a,\omega) - \sum_{i=1}^{N} \lambda_i(a_i)\dot{u}_i(a,\omega)),$$

and observe that for any given $\lambda \in \times_i B(A_i)$, setting $\gamma(\omega)$ equal to $\underline{\gamma}(\lambda,\omega)$ obtains the infimum of the objective in (9): any $\gamma(\omega) < \underline{\gamma}(\lambda,\omega)$ is infeasible and any $\gamma(\omega) > \underline{\gamma}(\lambda,\omega)$ can be improved upon by decreasing $\gamma(\omega)$. Hence, the dual problem can be equivalently stated as

$$V^D = \inf_{\lambda \in \times_i B(A_i)} \int_{\Omega} \underline{\gamma}(\lambda,\omega) \mathrm{d}\mu_0.$$

To establish the existence of the solution, we first show that the domain can be bounded. To bound the domain from above, define

$$\overline{\lambda} \triangleq \frac{V^P + 1 - \underline{V}}{\varepsilon \inf_{\omega \in \Omega} \mu_0(\omega)},$$

where $\underline{V} = \inf_{a,\omega} v(a,\omega) > -\infty$ by Assumption 2 and ε is that of Assumption 3. Consider any λ such that for some $i = 1, \dots, N$ and $a_i \in A_i$, $\lambda_i(a_i) > \overline{\lambda}$. By Assumption 3, there exist ω^- and a_{-i} such that:

$$\begin{aligned}
\underline{\gamma}(\lambda, \omega^-) &= \sup_{a' \in A} (v(a', \omega^-) - \sum_{i=1}^N \lambda_i(a_i') \dot{u}_i(a', \omega^-)) \\
&\geq v(a_i, a_{-i}, \omega^-) - \lambda_i(a_i) \dot{u}_i(a_i, a_{-i}, \omega^-) \\
&\geq \underline{V} + \varepsilon \overline{\lambda},
\end{aligned}$$

Moreover, for any $\omega \in \Omega$, $\underline{\gamma}(\lambda, \omega) \geq \underline{V}$. Indeed, according to the Glicksberg-Fan theorem, there exists an allocation $\underline{a}$ such that $\dot{u}_i(\underline{a}, \omega) \equiv 0$ that achieves $v(\underline{a}, \omega) \geq \underline{V}$ irrespective of λ. Hence, under such λ, the value of the dual problem is at least $\underline{v} + \mu_0(\omega^-)\varepsilon\overline{\lambda} > V^P + 1$. However, by strong duality, $V^D = V^P$; consequently, such λ may be excluded from the optimization domain without any loss. An analogous argument bounds the optimization domain from below.

Second, by the definition of the infimum, there exists a sequence $\{\lambda^n\}_{n=1}^\infty$ such that $\lim_{n\to\infty} \int_\Omega \underline{\gamma}(\lambda^n, \omega) \mathrm{d}\mu_0 = V^D$. As the domain is bounded, there exists a bounded point-wise limit of this sequence, $\lambda^* \in \times_i B(A_i)$, $\lambda_i^*(a) \triangleq \lim_{n\to\infty} \lambda_i^n(a)$ for all $i = 1, \dots, N$, $a \in A$. We have:

$$\begin{aligned}
V^D &= \lim_{n\to\infty} \int_\Omega \underline{\gamma}(\lambda^n, \omega) \mathrm{d}\mu_0 \\
&= \int_\Omega \lim_{n\to\infty} \sup_{a \in A} \left(v(a, \omega) - \sum_{i=1}^N \lambda_i^n(a_i) \dot{u}_i(a, \omega) \right) \mathrm{d}\mu_0 \\
&\geq \int_\Omega \sup_{a \in A} \lim_{n\to\infty} \left(v(a, \omega) - \sum_{i=1}^N \lambda_i^n(a_i) \dot{u}_i(a, \omega) \right) \mathrm{d}\mu_0 \\
&= \int_\Omega \underline{\gamma}(\lambda^*, \omega) \mathrm{d}\mu_0 \\
&\geq V^D,
\end{aligned}$$

where the third line follows from the order of the supremum and the last line follows from the definition of V^D as the optimal value of the dual problem. Hence, $\{\lambda, \gamma\} = \{\lambda^*, \underline{\gamma}(\lambda^*, \omega)\}$ solve the dual problem. This concludes the proof. □

As argued in the main text, the dual problem (9) is equivalent to the dual adversarial-contracting problem (11). This concludes the proof of Theorem 3.

B.3 Multidimensional State Perturbation

In Sections 4.2 and 4.3, we observed a multiplicity of optimal information structures driven by the many ways in which one can implement an optimal aggregate action. Moreover, a symmetric Gaussian information structure featured an exogenous noise. In this section, we argue that these features may be driven by a limited dimensionality of a state space and they disappear in a "nearby" multi-dimensional setting.

In particular, we consider the following multidimensional version of a co-movement setting of Section 4.2.2. The state is N-dimensional, $\omega = (\omega_1, \ldots, \omega_N) \in \mathbb{R}^N$ and is normally distributed according to a normal distribution with mean 0 and covariance matrix Σ such that $\Sigma_{ii} = 1$ and $\Sigma_{ij} = 1 - \Delta^2$ for $j \neq i$, where $\Delta \geq 0$.

Player i aims to predict the ith state component: $A_i = \mathbb{R}$ and

$$u_i(a, \omega) = -(a_i - \omega_i)^2. \tag{87}$$

The designer's payoff features the co-movement and the miscoordination components:

$$v(a, \omega) = \frac{\sum_{i=1}^N a_i \omega_i}{N} - \rho \frac{\sum_{i=1}^N \sum_{j \neq i} a_i a_j}{N^2}, \tag{88}$$

so that it is analogous to payoff (39) with the difference that the co-movement component accounts for the state's multi-dimensionality and keeps track of the average of individual co-movement components. We focus on the case $\rho \geq \frac{N}{2N-1}$.

For all $\Delta \geq 0$, the state's cumulative probability distribution is point-wise continuous in Δ. At $\Delta = 0$ the setting is equivalent to the setting of Section 4.2.2. As such, the settings for small $\Delta > 0$ can be viewed as small perturbations of the original setting. We show, however, that a certifiably optimal information structure in the perturbed settings with $\Delta > 0$ does not feature an exogenous noise. Moreover, the certifiably optimal allocation rule is uniquely optimal under the linear certifying contracts.

For any $\Delta > 0$, consider the dual problem and set the contract as $\lambda_i(a_i) = -\frac{q}{2N^2} a_i$, parameterized by $q \in \mathbb{R}$. The dual payoff is

$$v(a, \omega) - \sum_i \lambda_i(a_i) \dot{u}_i(a, \omega) = \frac{\sum_{i=1}^N a_i \omega_i}{N} - \rho \frac{\sum_{i=1}^N \sum_{j \neq i} a_i a_j}{N^2} - \frac{q}{N^2} \sum_{i=i}^N \lambda_i(a_i)(a_i - \omega) \tag{89}$$

$$= \frac{1}{N^2}((q + N) a^T \omega - a^T M_N(q, \rho) a), \tag{90}$$

where $M_N(x, y)$ is an N-dimensional rectangular matrix in which each on-diagonal element equals x and each off-diagonal element equals y. For $q > \rho$, matrix $M_N(q, \rho)$ is positive-semidefinite and the agent's best response can be found from first-order conditions as:

$$a^*(\omega; q) = \frac{q+N}{2} M_N^{-1}(q, \rho)\omega = \frac{q+N}{2(q-\rho)(q+(N-1)\rho)} M_N(q+(N-2)\rho, -\rho)\omega, \quad (91)$$

or, equivalently,

$$a_i^*(\omega; q) = \frac{q+N}{2(q-\rho)(q+(N-1)\rho)}((q+(N-2)\rho)\omega_i - \rho\sum_{j\neq i}\omega_j), \quad (92)$$

Any such allocation rule $a^*(\omega; q)$ is implementable by a linear contract with parameter $q > \rho$. Since the state is normally distributed and $\mathbb{E}[a_i - \omega_i]$=0, for such allocation rule to be implementable with incentives, it suffices that for all i

$$\mathbb{E}[a_i^*(\omega; q)(a_i^*(\omega; q) - \omega_i)] = 0. \quad (93)$$

Condition (93) is equivalent to

$$\begin{aligned} &\mathbb{E}[((q+(N-2)\rho)\omega_i - \rho\sum_{j\neq i}\omega_j)* \quad (94)\\ &\quad (2(q-\rho)(q+(N-1)\rho)\omega_i - (N+q)(q+(N-2)\rho)\omega_i + (N+q)\rho\sum_{j\neq i}\omega_j)] = 0, \end{aligned}$$

which in turn is equivalent to

$$\begin{aligned} &(q+(N-2)\rho)[2(q-\rho)(q+(N-1)\rho) - (N+q)(q+(N-2)\rho) + (N+q)\rho(N-1)(1-\varepsilon^2)] \quad (95)\\ &\quad -\rho(N-1)[2(q-\rho)(q+(N-1)\rho)(1-\varepsilon^2) - (N+q)(q+(N-2)\rho)(1-\varepsilon^2)]\\ &\quad\quad -\rho^2(N+q)[(N-1)+(N-1)(N-2)(1-\varepsilon^2)] = 0. \end{aligned}$$

This condition is satisfied if and only if q satisfies the following:

$$\frac{2}{q+N} - \frac{1}{q-\rho} - \frac{1}{q+(N-1)\rho} + \frac{1}{q-\rho+\Delta^2\rho(N-1)} = 0. \quad (96)$$

There exists a unique real solution to this equation, $q^*(\Delta)$, which indeed satisfies $q^*(\Delta) > \rho$. As such, for all $\Delta > 0$ the information structure that recommends $a_i^*(\omega; q^*(\Delta))$ is optimal. It is symmetric, Gaussian, and does not feature any exogenous noise.

Moreover, as $\varDelta \to 0$, this information structure converges to the noisy symmetric Gaussian structure (41) that we identified as optimal at $\varDelta = 0$ in the main text. To see this, note that by Taylor expansion, $q^*(\varDelta) = \rho + \gamma\varDelta + o(\varDelta)$ where $\gamma = \sqrt{\frac{\rho^2 N(N-1)(N+\rho)}{\rho(2N-1)-N}}$. Furthermore, the Gaussian state admits a decomposition into a common component and independent noises as:

$$\omega_i = \sqrt{1-\varDelta^2}\omega_0 + \varDelta\varepsilon_i, \tag{97}$$

where ω_0 and ε_i are independently distributed according to $N(0,1)$. Then, as $\varDelta \to 0$, the recommended actions (92) converge almost surely to:

$$\begin{aligned} a_i(\omega) &\sim \frac{\rho+N}{2\varDelta\gamma N\rho}\left((\varDelta\gamma + (N-1)\rho)(\omega_0 + \varDelta\varepsilon_i) - \rho((N-1)\omega_0 + \varDelta\sum_{j\neq i}\varepsilon_i)\right) & (98)\\ &= \frac{\rho+N}{2N\rho}\omega_0 + \frac{\rho+N}{2N\gamma}\varepsilon_i - \frac{\rho+N}{2N\gamma}\sum_{j\neq i}\varepsilon_j. & (99)\end{aligned}$$

Substituting γ, we obtain that this is exactly the symmetric Gaussian structure (41).

The analysis in this section suggests a modification of our certification method for the environments in which an optimal information structure features extraneous noise. In such environments, if the optimal noise structure is difficult to guess, it might help considering a perturbed model which approaches the original model in the limit and in which the allocation rule is uniquely pinned down in the adversarial contracting problem. The right noise structure could then be obtained as the limit of this allocation rule.

B.4 Bounded Action Spaces

In the main text, we presented and studied the information-design problem in games with unbounded action spaces, $A_i = \mathbb{R}$ for all i. This enabled us to ignore the possibility of corner solutions and to present the analysis in its simplest form. In this section, we extend our methodology to bounded action spaces, which may be more natural in some applications, and show that our solution method easily generalizes to this case.

To this end, consider the concave information-design problem as in the main text, but let the action space of each player to be bounded from below and from above, $A_i = [\underline{a}_i, \overline{a}_i]$, $-\infty < \underline{a}_i < \overline{a}_i < +\infty$.[23] In this setting, for any given belief over the

[23] The extension to half-bounded spaces is straightforward.

state and opponents' actions $\nu \in \Delta(A_{-i} \times \Omega)$, the player's best-response action $a_i^*(\nu)$, if interior, must be unimprovable by local deviations to lower and higher actions and hence satisfies the same first-order condition as in the unbounded setting,

$$\mathbb{E}_\nu \left[\dot{u}_i(a_i^*, a_{-i}, \omega)\right] = 0. \tag{100}$$

At the same time, the optimal boundary actions must only be unimprovable by one-sided local deviations. As such, the player's best response if located on the boundary must satisfy the first-order conditions in the inequality form:

$$\mathbb{E}_\nu \left[\dot{u}_i(\underline{a}_i^*, a_{-i}, \omega)\right] \leq 0, \tag{101}$$

$$\mathbb{E}_\nu \left[\dot{u}_i(\overline{a}_i^*, a_{-i}, \omega)\right] \geq 0. \tag{102}$$

We can write the primal information-design problem with the modified obedience constraints as:

$$V^P \triangleq \sup_{\pi \in \Delta(A \times \Omega)} \int_{A \times \Omega} v(a, \omega) \mathrm{d}\pi \tag{103}$$

$$\text{s.t.} \int_{A_i' \times A_{-i} \times \Omega} \dot{u}_i(a, \omega) \mathrm{d}\pi = 0 \quad \forall\, i = 1, \ldots, N, \text{measurable}\, A_i' \subseteq (\underline{a}_i, \overline{a}_i), \tag{104}$$

$$\int_{A_i' \times A_{-i} \times \Omega} \dot{u}_i(a, \omega) \mathrm{d}\pi \leq 0 \quad \forall\, i = 1, \ldots, N, \text{measurable}\, A_i' \subseteq [\underline{a}_i, \overline{a}_i), \tag{105}$$

$$\int_{A_i' \times A_{-i} \times \Omega} \dot{u}_i(a, \omega) \mathrm{d}\pi \geq 0 \quad \forall\, i = 1, \ldots, N, \text{measurable}\, A_i' \subseteq (\underline{a}_i, \overline{a}_i], \tag{106}$$

$$\int_{A \times \Omega'} \mathrm{d}\pi = \int_{\Omega'} \mathrm{d}\mu_0 \quad \forall\, \text{measurable}\, \Omega' \subseteq \Omega. \tag{107}$$

This primal problem entails the dual problem

$$V^D \triangleq \inf_{\lambda \in \times_i L(A_i), \gamma \in L(\Omega)} \int_\Omega \gamma(\omega) \mathrm{d}\mu_0 \tag{108}$$

$$\text{s.t.} \sum_{i=1}^N \lambda_i(a_i) \dot{u}_i(a, \omega) + \gamma(\omega) \geq v(a, \omega)\ \forall\, a \in A, \omega \in \Omega,$$

$$\lambda_i(\underline{a}_i) \leq 0,\ \lambda_i(\overline{a}_i) \geq 0 \quad \forall\, i = 1, \ldots, N. \tag{109}$$

The presence of additional obedience constraints (105), (106) in the primal problem translates into the sign constraints on the Lagrange multipliers (109) in the dual problem. Intuitively, the marginal benefit from pushing the action a_i upward is negative at the lower bound and positive at the upper bound.

As was in the case of unbounded actions, the dual problem (108) can be simplified and rewritten as an adversarial-contracting problem:

$$V^D = \inf_{\lambda \in \times_i L(A_i)} \mathbb{E}_{\mu_0}[\sup_{a \in A} u^\lambda(a, \omega)] \tag{110}$$

$$\text{s.t. } \lambda_i(\underline{a}_i) \leq 0, \; \lambda_i(\overline{a}_i) \geq 0 \quad \forall\, i = 1, \ldots, N. \tag{111}$$

That is, the adversarial-contracting interpretation remains intact, but the space of allowed contracts is limited at the boundary actions by the presence of sign constraints on contract parameters.

In the case of bounded actions, say that the measure $\pi \in \varDelta(A \times \varOmega)$ is implementable by information if it satisfies the constraints of the primal problem (103) and that it is implementable by incentives if there exists a feasible contract in the dual problem (110) that induces this measure as a best response. The results on weak duality and the certification solution method continue to hold verbatim.

Theorem 4. (Weak Duality with Bounded Action Spaces) $V^P \leq V^D$.

Proof. Take any dual variables (λ, γ) that satisfy the constraints of dual problem (108). Take any measure π that satisfies the constraints of primal problem (103). Integrating both sides of the dual constraints over $a \in A$ and $\omega \in \varOmega$ against measure π yields:

$$\int_{A \times \varOmega} v(a, \omega) \mathrm{d}\pi \leq \int_{A \times \varOmega} \sum_{i=1}^{N} \lambda_i(a_i) \dot{u}_i(a, \omega) \mathrm{d}\pi + \int_{A \times \varOmega} \gamma(\omega) \mathrm{d}\pi \leq \int_{\varOmega} \gamma(\omega) \mathrm{d}\mu_0, \tag{112}$$

where the second inequality follows because π satisfies the primal constraints and the Lagrange multipliers satisfy the dual constraints. (This inequality holds as equality in the case of unbounded action spaces.) The left-hand side of (112) is the value of the primal problem given measure π. At the same time, the right-hand side of (112) is the value of the dual problem given dual variables (λ, γ). As inequality (112) holds for any allowed values of primal measure and dual variables, it also holds at the respective maximization and minimization limits. □

Proposition 7. (Optimality Certification with Bounded Action Spaces) *If measure $\pi \in \varDelta(A \times \varOmega)$ is implementable by information and by incentives, then π is optimal in the information-design problem.*

Proof. Take any primal measure π implementable by information, i.e., that satisfies the constraints of primal problem (103). If it is implementable by incentives, then there exist dual variables λ that implement this measure in the dual problem (110), and

$$V^D = \inf_{\hat{\lambda}\in\times_i L(A_i)} \mathbb{E}_{\mu_0}\left[\sup_{a\in A} u^{\hat{\lambda}}(a,\omega)\right] \tag{113}$$
$$\leq \mathbb{E}_{\mu_0,\pi}\left[u^{\lambda}(a,\omega)\right] \tag{114}$$
$$= \int_{A\times\Omega} v(a,\omega)\mathrm{d}\pi - \int_{A\times\Omega}\sum_{i=1}^{N}\lambda_i(a_i)\dot{u}_i(a,\omega)\mathrm{d}\pi \tag{115}$$
$$\leq \int_{A\times\Omega} v(a,\omega)\mathrm{d}\pi \leq V^P, \tag{116}$$

where the first inequality follows from the implementability of π in the dual problem and the last three steps follow from the feasibility of π in the primal problem and the sign constraints on dual variables. (The one but last step holds as equality in the case of unbounded actions.)

Furthermore, by Theorem 4, $V^D \geq V^P$. Hence,

$$V^D = \int_{A\times\Omega} v(a,\omega)\mathrm{d}\pi = V^P, \tag{117}$$

which proves the optimality of measure π. □

To summarize, information design in concave games with bounded action spaces can be approached with the same methodology as the information design in concave games with unbounded action spaces presented in the main text.