WORKING PAPERS

N° 1313

March 2022

"Information Design in Concave Games"

Alex Smolin and Takuro Yamashita

Toulouse School of Economics

Information Design in Concave Games *

Alex Smolin, Takuro Yamashita

February 23, 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. The information structure induces a Bayesian game and is evaluated according to the expected designer's payoff under the equilibrium play.

We develop a method that facilitates the search for an optimal information structure, i.e., one that cannot be outperformed by any other information structure, however complex. We show an information structure is optimal whenever it induces the strategies that can be implemented by an incentive contract in a dual, principal-agent problem which aggregates marginal payoffs of the players in the original game. We use this result to establish the optimality of Gaussian information structures in settings with quadratic payoffs and a multivariate normally distributed state. We analyze the details of optimal structures in a differentiated Bertrand competition and in a prediction game.

1 Introduction

Information flows are vital for the digital economy and are increasingly controlled by technological giants. Amazon decides how to display products on its shopping website. Google decides how to aggregate user reactions on its video platform. Meta decides how to compose

*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).

the news feed on its social network. All these companies decide what customer characteristics to reveal to bidding advertisers. These choices raise the question of optimal information control, which was formalized in the field of information design or Bayesian persuasion (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 binary state, binary actions, a single player, or for special classes of information structures. In this paper, we expand this methodology by developing a solution method to find unconstrained-optimal information structures in large-scale games with a continuum of states and actions and nonlinear payoffs.

Specifically, we study information design in *concave* games of incomplete information, i.e., games in which each player's action can take values on the real line 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 equilibrium behavior is determined by a system of such conditions. The induced joint 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 amount of optimization parameters and constraints. Instead, we develop a solution method to check and certify the optimality of candidate information structures. To do so, we construct a dual problem, as defined in the field of convex optimization, and show that it admits an interpretation of adversarial contracting. In that problem, a contract designer faces a single agent who fully controls all actions and perfectly observes the state. The designer chooses a contract that affects the agent's incentives: the agent's payoff is the sum of the information-designer payoff and the marginal payoffs of the players scaled by the weights specified by the contract. The contract designer anticipates the agent's best response and chooses the contract to minimize the expected agent's payoff, hence the name "adversarial."

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 the information structure implements a state-action distribution that can be implemented by some contract in the dual problem, the information structure is optimal (Proposition 1). 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. 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 tractable 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.

We apply Theorem 2 to characterize optimal information structures in two concrete settings. First, we study a differentiated Bertrand duopoly with linear demand curves and uncertain demand shocks that determine the curves' intercepts. We show that information structures that maximize a weighted average of the consumer and produce surpluses 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, we study a stylized prediction game in which each player aims to predict a one-dimensional state and the designer wants to increase the discordance of players' predictions. We show that an optimal information structure introduces exogenous Gaussian noise that is correlated across players. However, as the number of players increases, the noise correlation decreases. In the limit, the optimal individual noises are independent.

Related Literature The literature on information design or Bayesian persuasion 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

[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), Brooks and Du (2020)), among others.

this analysis to higher moments and beyond.[2] Malamud and Schrimpf (2021) and Cieslak et al. (2021) establish some general properties of optimal information structures building on optimal-transportation duality. Kolotilin (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. None of these works study games 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.[3] The literature that perhaps comes closest to this goal is one that studies optimal parameters of symmetric Gaussian information structures in symmetric games with a normally distributed state (Bergemann and Morris (2013), Bergemann et al. (2015), Bergemann et al. (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 4.2 suggest that symmetric Gaussian information structures may indeed be optimal in some of those settings and may be possibly certified by our solution method.

2 General 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 $\Omega \subseteq \mathbb{R}^K$ 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

[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.

[3]Galperti and Perego (2018) and Galperti et al. (2021) apply duality methodology to study information design in general games with finitely many actions. However, they focus on the analysis and the interpretation of optimal dual variables rather than on the study of optimal information structures.

[4]Our methodology can easily be extended to the case of multidimensional actions at the expense of additional notation.

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 $S = \times_i S_i$ 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.[5] The information structure can provide information about the state and coordinate 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 $s = (s_1, \dots, s_N)$ 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.

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

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

for all i, $s_i \in S_i$, $a_i \in \operatorname{supp} \sigma_i(\cdot \mid s_i)$, and $a_i' \in A_i$, where $\mathbb{E}_{\mathcal{I}}[\cdot | s_i]$ denotes a conditional mathematical expectation given $\mathcal{I}$ and s_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

[5] We impose a standard requirement that for any signal, each player can form a regular conditional probability over $A_{-i} \times \Omega$.

arise in equilibrium of the induced Bayesian game.[6] 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, strictly concave in a_i, and obtains its maximum at some finite value.*

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 per each player's signal.[7] 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 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 given 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 $\mathbb{E}_{\nu_i}\left[u_i(a_i, a_{-i}, \omega)\right]$. 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

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

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

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

Assumption 1 implies that $\dot{u}_i(a, \omega)$ exists, is continuous, and strictly decreases in a_i everywhere. The first-order condition can be written as

$$\frac{\partial \mathbb{E}_\nu \left[u_i(a_i^*, a_{-i}, \omega)\right]}{\partial a_i} = \mathbb{E}_\nu \left[\frac{\partial u_i(a_i^*, a_{-i}, \omega)}{\partial a_i}\right] = \mathbb{E}_\nu \left[\dot{u}_i(a_i^*, a_{-i}, \omega)\right] = 0, \tag{4}$$

where the first equation follows from the Leibniz integral rule and the second by definition. Equation (4) identifies the best response of each player given any belief.

Next, we simplify the information-design problem by appealing to the revelation principle (Myerson (1983), Bergemann and Morris (2016)) to restrict the space of induced posterior beliefs. According to this principle, the designer in any information-design problem can 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; providing any other kind of information is redundant and may only complicate the players' incentives. Each direct information structure corresponds to a distribution $\pi \in \Delta(A \times \Omega)$ that has μ_0 as its state marginal.

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

$$V^P \triangleq \sup_{\pi \in \Delta(A\times\Omega)} \int_{A\times\Omega} v(a, \omega) \mathrm{d}\pi \tag{5}$$

$$\text{s.t.} \int_{A_i' \times A_{-i} \times \Omega} \dot{u}_i(a, \omega) \mathrm{d}\pi = 0 \quad \forall\, i = 1, \ldots, N, \text{measurable}\, A_i' \subseteq A_i, \tag{6}$$

$$\int_{A\times\Omega'} \mathrm{d}\pi = \int_{\Omega'} \mathrm{d}\mu_0 \quad \forall\, \text{measurable}\, \Omega' \subseteq \Omega. \tag{7}$$

Constraints (6) capture players' obedience and must hold at all measurable subsets $A_i' \subseteq A_i$ and 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 (4) in light of a possible continuum of recommended actions. Constraints (7) capture Bayes' plausibility and, likewise, require that the linear projection of π on Ω equals the prior distribution μ_0.

The problem (5) 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*, which is dual to the primal problem. The significance of this problem, and dual problems in general (see, for example, Villani (2003)), comes from its ability to provide an upper bound on the information designer's payoffs and, ultimately, certify a solution. The dual problem to (5) is as follows:

$$V^D \triangleq \inf_{\lambda \in \times_i B(A_i), \gamma \in B(\Omega)} \int_\Omega \gamma(\omega) \mathrm{d}\mu_0 \qquad (8)$$
$$\text{s.t. } \sum_{i=1}^N \lambda_i(a_i) \dot{u}_i(a, \omega) + \gamma(\omega) \geq v(a, \omega) \; \forall\, a \in A, \omega \in \Omega,$$

where $B(X)$ denotes the space of bounded real-valued functions on X. The minimization arguments, the dual variables (λ, γ), represent the Lagrange multipliers associated with the primal incentive constraints (6) and the feasibility constraints (7), respectively. Thus, they have a clear economic interpretation: $\lambda_i(a_i)$ measures the marginal benefit for the information designer from pushing the action a_i downwards, whereas $\gamma(\omega)$ measures the marginal benefit from increasing the prior probability of state ω.

The dual problem (8) itself can be simplified and rewritten to admit an economic interpretation. To this end, observe that the objective in (8) 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

$$u^\lambda(a, \omega) \triangleq v(a, \omega) - \sum_{i=1}^N \lambda_i(a_i) \dot{u}_i(a, \omega). \qquad (9)$$

As a result, the dual problem (8) can be restated as

$$V^D = \inf_{\lambda \in \times_i B(A_i)} \int_\Omega \sup_{a \in A} u^\lambda(a, \omega) \mathrm{d}\mu_0 = \inf_{\lambda \in \times_i B(A_i)} \mathbb{E}_{\mu_0}[\sup_{a \in A} u^\lambda(a, \omega)]. \qquad (10)$$

The problem (10) can be interpreted as *adversarial contracting* between a contract designer and a single agent. The agent perfectly observes the state and alone controls the whole action profile. To influence the agent's behavior, the contract designer chooses an incentive contract λ that consists of N functions $\lambda_i(a_i)$ and modifies the agent's payoff according to (9), i.e., the i-th 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*. 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.

To better understand the relationship between the primal and dual problems, observe that the contract designer in a sense aims to minimize the payoffs of the information designer. More precisely, note that in adversarial contracting, a benchmark feasible contract is a null contract $\lambda_1(a_1) \equiv \cdots \equiv \lambda_N(a_N) \equiv 0$. 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 contract designer can then be viewed as adjusting the null contract to minimize the expected payoff starting from this benchmark level.

In the next section, we show that this rough intuition about the relationship between the primal and the dual problems goes in the right direction; the optimal values of these problems are tightly connected.

3.3 Weak Duality. Optimality Certification.

In this section, we establish the general relation between the values of the primal and dual problems and demonstrate how it can be used to solve concave information-design problems.

Theorem 1. (Weak Duality) *The optimal value of a concave information-design problem is weakly below the optimal value of the dual adversarial contracting:*

$$V^P \leq V^D. \tag{11}$$

Proof. Take any dual variables (λ,γ) that satisfy the constraints of dual problem (8):

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

Take any measure π that satisfies the constraints of primal problem (5). Integrating both sides of the inequality over $a \in A$ and $\omega \in \Omega$ against the measure $\pi(a,\omega)$ yields:

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

$$= \int_{\Omega} \gamma(\omega)\mathrm{d}\mu_0, \tag{13}$$

where the equality follows because $\pi(a,\omega)$ 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 without any additional assumptions, the adversarial contracting problem provides an upper bound on the value of information control. Perhaps more importantly, this result underlies the *certification approach* to solve concave information-design problems.

Proposition 1. (Optimality Certification) *Consider any measure $\pi \in \Delta(A\times\Omega)$ implementable by information. If π is implementable 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 (5). If it is implementable by incentives, then there exist dual variables λ that implement this measure in the dual adversarial contracting (10), and

$$V^D = \inf_{\lambda\in\times_i B(A_i)} \mathbb{E}_{\mu_0}\left[\sup_{a\in A} u^{\lambda}(a,\omega)\right] \tag{14}$$

$$\leq \mathbb{E}_{\mu_0,\pi}\left[u^{\lambda}(a,\omega)\right] \tag{15}$$

$$= \int_{A\times\Omega} v(a,\omega)\mathrm{d}\pi - \int_{A\times\Omega} \sum_{i=1}^{N} \lambda_i(a_i)\dot{u}_i(a,\omega)\mathrm{d}\pi \tag{16}$$

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

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

Furthermore, by Theorem 1, $V^D \geq V^P$. Hence,

$$V^D = \int_{A\times\Omega} v(a,\omega)\mathrm{d}\pi = V^P, \tag{18}$$

which proves the optimality of measure π. $\square$

Proposition 1 offers the following solution method for concave information-design problems. In the first step, one conjectures an optimal allocation 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 adversarial contracting. In the last step, one sets an optimal information structure to privately recommend actions to players according to the conjectured measure π^*: $\mathcal{I}^* = (A, \pi^*)$. The implementability of π^* with information implies that the players would find it optimal to follow the recommendations. The implementability of π^* with incentives confirms, by Proposition 1, that $\mathcal{I}^*$ is optimal among all possible 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 $supp\alpha_2(\omega) \subseteq supp\alpha_1^*(\omega)$ 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.

In general information-design problems, the prior distribution plays two roles. First, it affects what allocation rules are implementable by information. Second, it affects what allocation rules among those that can be implemented are optimal. Proposition 2 highlights that in concave information-design problems, the first channel is more important than the second one: once optimal under some prior distribution, an information structure remains optimal under other prior distributions, as long as it implements an allocation rule with the same support in each state.

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 prior distribution in this example.

Proposition 2 enables us to assess the optimality of 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 about the state. Such an information structure can still allow for strategic uncertainty, i.e., uncertainty about the actions of other players. We have the following.

Corollary 1. (Complete 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 immediately from Proposition 2. Indeed, 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. Consequently, 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) *Consider a concave information-design problem. If an information structure $\mathcal{I}^*$ is certifiably optimal and induces an allocation rule α^* with $supp\alpha^*(\omega) = A$ for all $\omega \in \Omega$, then any information structure is certifiably optimal.*

Corollary 2 presents a strong case against using independent noises that induce full-support individual actions in optimal information structures. These information structures induce full support action profiles and can never be optimal in concave problems with finitely many players, except in the trivial cases in which the designer's expected payoff is invariant to the information provided. 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 4.2.

3.5 Scope of the Certification Method

Can any optimal information structure can 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 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) the 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)). The latter property requires the solutions to both problems to exist, which might fail due to a lack of compactness.

It is certainly of theoretical interest to understand under which conditions the certification method is guaranteed to work. Thus, in the Appendix, we provide a set of sufficient conditions by building on the Fenchel-Rockafellar duality of optimal transportation theory.[8] However, from a practical perspective, knowing that an optimal information structure can be certified does not help in finding the optimal structure or the certifying contract. Vice versa, 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, it may be worth constructing the dual problem and searching for certifiable information structures in any concave problem at hand; if successful, the method leads to the solution. This is exactly what we do in the next section.

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 and the state is normally distributed. In particular, as before, we consider a setting with N players, each player taking an action $a_i \in A_i = \mathbb{R}$, so an action profile is $a \in \mathbb{R}^N$. The payoff relevant uncertainty is captured by a multidimensional state $\omega = (\omega_1, \dots, \omega_K) \in \Omega \subseteq \mathbb{R}^K$, $K \geq 1$, that can be viewed as a collection of state components ω_k, $k = 1, \dots, K$.

In addition to that basic structure, we impose two assumptions on the environment. First, we assume that the state components are jointly normally distributed so $\Omega = \mathbb{R}^K$ and $\mu_0 \sim N(0, \Sigma)$, where all means are set to zero without loss of generality and Σ is the arbitrary covariance matrix. Second, we assume that all payoff functions are quadratic. Namely, 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,

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

$$u_i(a, \omega) = a^T(b + B\omega) - \frac{1}{2}a^TCa. \tag{20}$$

Elements $\hat{b}_i, B_{i\boldsymbol{\cdot}}$ capture the base benefit of player i's action, elements $\hat{B}_{ij}, B_{ij}$ capture the interaction between player i's action and 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 for the players, respectively.[9] Note how this setting allows for asymmetries across players.

To maximize her payoffs, the designer chooses an information structure $\mathcal{I} = (S, \pi)$. As we argued in the previous section, without loss of generality, we can focus on direct information

[8]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.

[9]A useful mnemonic to link the variables correctly is to think of a designer as wearing a hat.

structures with $S = A$ under which the designer recommends to players what actions to take and the players are willing to follow the recommendations.

In what follows, we *do not* restrict the designer to use Gaussian, i.e., normally distributed, 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 an information structure that recommends actions proportionally to state components:

$$a^*(\omega) = a_0 + R\omega, \tag{21}$$

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 $a_i(\omega) = a_{0i} + R_{i\bullet}\omega$, where for any matrix X, the term $X_{i\bullet}$ denotes its ith row. 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 $i = 1, \ldots, N$, 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 $D(x)_{ii} \triangleq x_i$.

Proof. To establish the optimality of the information structure that recommends allocation rule (21), we use the certification method developed in the previous section. In particular, we show that under the conditions of Theorem 2, allocation rule (21) 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, i.e., the players' willingness to follow recommendations, is captured by the system of first-order conditions, as in any concave game. In a quadratic game, this is a system of linear equations. An allocation rule $\alpha : \Omega \to \Delta(A)$ is incentive compatible if and only if

$$\mathbb{E}[b_i + B_{i\bullet}\omega - C_{i\bullet}a \mid a_i] = 0 \quad \forall i = 1, \ldots, N, a_i \in A_i, \tag{22}$$

where, as before, conditions (22) are set to be automatically satisfied for actions a_i never recommended under α.

Because conditions (22) must hold for all a_i, they must also hold, on average; thus,

$$\mathbb{E}_{\mu_0}[b_i + B_{i\bullet}\omega - C_{i\bullet}a] = 0 \quad \forall i = 1, \ldots, N. \tag{23}$$

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

$$a_0 = C^{-1}b. \tag{24}$$

That is, the constant term is the same for all implementable linear allocation rules, and they all lead to the same expected actions; however, the action distributions may differ.

Since the state components are jointly normally distributed, the recommended action profile a^*, as well as the variable $C_{i\bullet}a - B_{i\bullet}\omega - b_i$, are also normally distributed. Hence, the sufficient condition for first-order conditions (22) is that a_i and $C_{i\bullet}a - B_{i\bullet}\omega - b_i$ are uncorrelated for all i. The zero-correlation condition can be written as

$$\begin{aligned}
\operatorname{cov}(C_{i\bullet}a^* - B_{i\bullet}\omega - b_i, a_i^*) &= \mathbb{E}[(C_{i\bullet}R - B_{i\bullet})\omega R_{i\bullet}\omega] \\
&= \mathbb{E}[(C_{i\bullet}R - B_{i\bullet})\omega\omega^T R_{i\bullet}^T] \\
&= (C_{i\bullet}R - B_{i\bullet})\mathbb{E}[\omega\omega^T]R_{i\bullet}^T \\
&= (C_{i\bullet}R - B_{i\bullet})\Sigma R_{i\bullet}^T = 0.
\end{aligned} \tag{25}$$

To summarize, a linear allocation rule (21) is implementable by information whenever its coefficients a_0 and R satisfy the conditions (24) and (25) that together form condition (i) of the theorem. However, clearly, 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 (ii) of Theorem 2 plays a role. Under this condition, as we show in the Appendix, the linear allocation rule can be implemented by incentives in the dual problem, and in fact, it can be done by linear contracts with the proportionality coefficients x. According to Proposition 1, any allocation rule implementable in both problems is optimal in both problems, which concludes the proof.

□

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.

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 design and commit to any information structure, however complex. We characterize the information structure such a designer 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 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:

$$w(q, a, \theta) \triangleq w_0 + \frac{1}{2}(\theta - q)^T W^{-1}(\theta - q) - a^T q, \tag{26}$$

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:[10]

$$q_i(a, \theta) = \theta_i + \eta a_i + \xi a_{-i}, \tag{27}$$

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

where $\eta \triangleq W_{ii} < 0$ and $\xi \triangleq W_{-ii}$; thus, equivalently, $q(a, \theta) = \theta + Wa$. Equation (27) reveals that the chosen type structure micro-founds linear demand; the consumer's type simply determines the intercept of the demand curve for each product. Parameters η and ξ capture demand sensitivity to the own price and to the competitor's price, respectively; 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:

$$u_i(a, \theta) = a_i q_i(a, \theta) - c q_i(a, \theta)^2. \tag{28}$$

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

$$CS(a, \theta) = -a^T \theta - \frac{1}{2} a^T W a, \tag{29}$$

$$\Pi(a, \theta) \triangleq u_1 + u_2 = -c\theta^T \theta + a^T (1 - 2cW)\theta - \frac{1}{2} a^T (2W + 2cW^2) a. \tag{30}$$

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

$$v(a, \theta) = \delta \times CS(a, \theta) + (1 - \delta) \times \Pi(a, \theta). \tag{31}$$

The firms share a prior belief about the consumer type, equal to the true distribution. The designer can provide additional information about the type privately to each firm: she can choose an arbitrary information structure and does so to maximize her objective.

Given (31), 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 solution 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 optimal information structures span the Pareto frontier of feasible equilibrium payoffs 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.[11] 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.

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

Direct Price Control We begin the analysis by studying a hypothetical scenario in which the designer can directly control 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.

The first-best benchmark 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}:

$$\delta^{FB} = \frac{2 + 2c(-\eta - |\xi|)}{3 + 2c(-\eta - |\xi|)}, \tag{32}$$

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. Clearly, this outcome is a modeling artifact and can never be implemented through information control. Nevertheless, it highlights the designer's willingness to decrease prices as she cares progressively more about consumer welfare.

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 optimality conditions to (31). The solution is proportional to the type and can be written in matrix form as

$$a^{FB} = R^{FB}\theta, \tag{33}$$

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.

Full Disclosure and No Disclosure 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.

Under the not informative information structure, each firm obtains no additional information, $S_1 = S_2 = \{s_0\}$. Thus, each firm's belief stays at the prior, and the equilibrium prices satisfy the first-order condition derived from (29):

$$\mathbb{E}_{\mu_0}[q_i(a_i, a_{-i}, \theta) + \frac{\partial q_i(a_i, a_{-i}, \theta)}{\partial a_i}(a_i - 2ca_i)] = 0. \tag{34}$$

In a symmetric equilibrium, each firm sets a price:

$$a_i^{NI} = \frac{1 - 2c\eta}{-2\eta(1 - c\eta) - \xi(1 - 2c\eta)}\bar{\theta}. \tag{35}$$

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. The proportionality coefficient, naturally, depends on the own-price and cross-price sensitivities.

In contrast, under a fully informative information structure, the firms obtain perfect demand information, $S_1 = S_2 = \Theta$, and π is concentrated on the event $s_1 = s_2 = \theta$. The consumer type is always commonly known to firms. Under this information structure, there is an equilibrium such that each firm responds linearly to the type components perfectly anticipating the price of its opponent:

$$a_i^{FI}(\theta) = \frac{-2\eta(1 - c\eta)(1 - 2c\eta)}{4\eta^2(1 - c\eta)^2 - (1 - 2c\eta)^2\xi^2}\theta_i + \frac{(1 - 2c\eta)^2\xi}{4\eta^2(1 - c\eta)^2 - (1 - 2c\eta)^2\xi^2}\theta_{-i}. \tag{36}$$

In a sense, this behavior generalizes the 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 the demand is asymmetric across firms, and the prices are adjusted. At the same time, $\mathbb{E}[a_i^{FI}(\theta)] = a_i^{NI}$. This is an instance of a general observation made in the previous section that all implementable linear allocation rules lead to the same average actions.

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. However, the space of all possible information structures is too large for a direct search to be tractable. Instead, we find an optimal structure using the certification method developed for the normal-quadratic setting and which application we now outline.

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 the equation $(C_{i\bullet}R(x) - B_{i\bullet})\Sigma R(x)_i^T = 0$ to hold for both firms. Due to symmetry of the environment, this condition becomes a single equation:

$$f(x) = 0, \tag{37}$$

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 (37) admits up to four real solutions. If any of these solutions makes the matrix $C + 2D((x, x))\hat{C}$ positive semidefinite, then Theorem 2 allows us to conclude that the information structure that recommends the linear allocation rule $a(\theta) = a_0(x) + R(x)\theta$ is optimal. This discussion is summarized in the proposition below.

Proposition 3. (Optimal Demand Information) *Polynomial $f(x)$ certifies a solution to the information-design problem in differentiated product competition: 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 $\xi = 1/2$, so $\theta_i \sim N(3, 1)$ and the goods are imperfect substitutes with the demand function for product i being

$$q_i(a, \theta) = \theta_i - a_i + \frac{1}{2}a_{-i}. \tag{38}$$

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

$$a^{NI}(\theta) = \frac{6}{5}, \tag{39}$$

$$a^{FD}(\theta) = \frac{48}{55}\theta_i + \frac{18}{55}\theta_{-i}. \tag{40}$$

Under full information, each firm increases its price in response to higher demand for either product yet, naturally, is more sensitive to the demand for its own product. As a result, the prices are more dispersed and volatile under full information than they are under no information. 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.

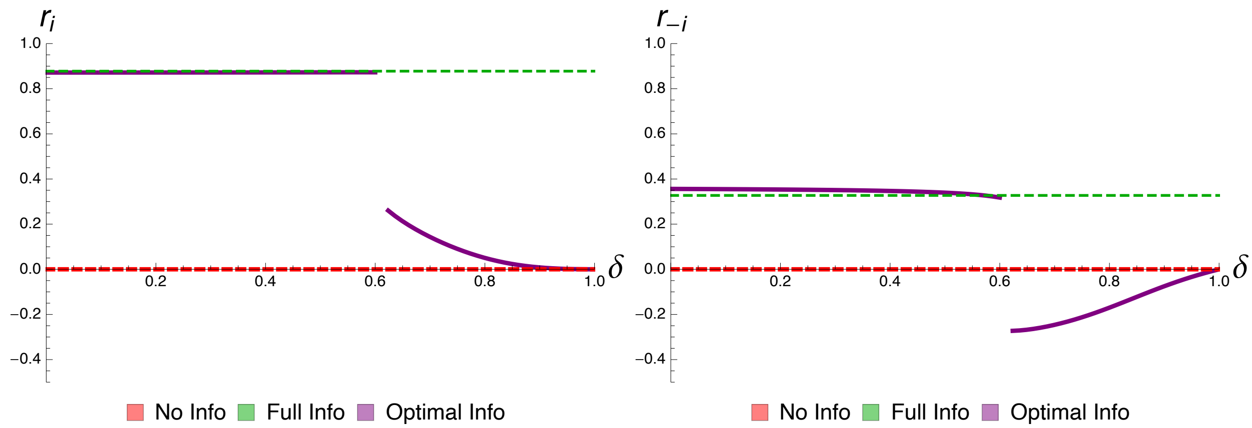

Figure 1: Price responsiveness to own demand (left) and opponent's demand (right) under no information, full information, and optimal information structure. Calculated at $c = 1$, $\bar{\theta} = 3$, $\sigma^2 = 1$, $\eta = -1$, and $\xi = 1/2$.

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 information structure. This structure recommends a linear allocation rule symmetrically across firms:

$$a_i^*(\theta) = a_{i0} + r_i \theta_i + r_{-i} \theta_{-i}, \tag{41}$$

with the constant term a_{i0} being pinned down by the responsiveness coefficient to satisfy the necessary condition for implementability of linear allocation, $\mathbb{E}[a_i^*(\theta)] = a^{NI}$. We refer to responsiveness coefficients r_i, r_{-i} as to *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 complete information is not optimal. Even if $\delta = 0$, i.e., if the designer aims to maximize firms' profits, the designer benefits from information control: to dampen competition, the designer induces larger responses to the opponent's demand. As δ increases, the cross-responsiveness decreases, even if moderately. 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

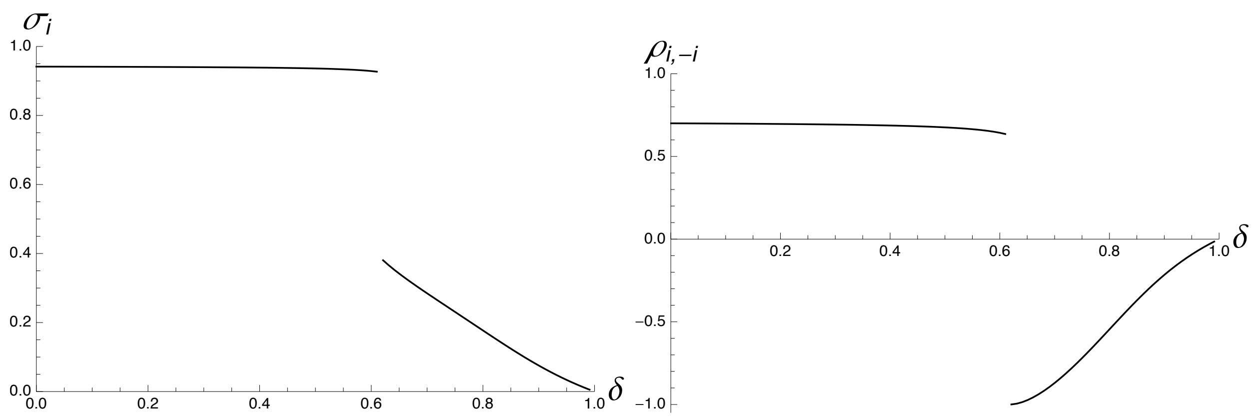

Figure 2: Price volatility (left) and price correlation (right) under the optimal information structure. Calculated at $c = 1$, $\bar{\theta} = 3$, $\sigma^2 = 1$, $\eta = -1$, and $\xi = 1/2$.

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, indicating that, in this case, providing no information is optimal: a designer who wishes to maximize consumer surplus should keep the firms in the dark.

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 $\rho_{i,-i} = \frac{\text{cov}(a_i, a_{-i})}{\sigma_i \sigma_{-i}} = \frac{2r_i r_{-i}}{r_i^2 + r_{-i}^2}$. This region is marked by coordination and high volatility of prices. By 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 around the critical value. In other words, even though a certifying contract changes continuously, the allocation rule that it implements exhibits a jump.

To understand intuitively the economic reasoning behind this discontinuity, it is instruc-

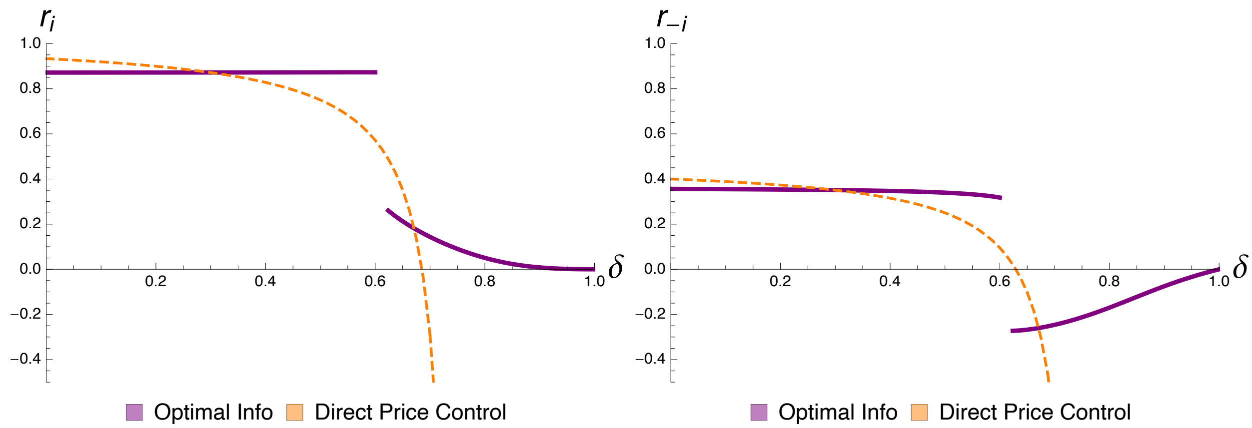

Figure 3: Price responsiveness to own demand shock (left) and opponent's demand shock (right) in the first-best benchmark and under the optimal information structure. Calculated at $c = 1$, $\bar{\theta} = 3$, $\sigma^2 = 1$, $\eta = -1$, and $\xi = 1/2$.

tive 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 more the designer cares about consumers, the lower the responsiveness she wants to induce. This logic underlies the structure of pricing behavior induced by optimal information control, which overall follows the decreasing 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 to this overall picture. 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 to be approached in a coordinated fashion, with both responsiveness coefficients being positive; the better way to approach them is to switch to an anticoordination structure, with the cross-responsiveness being negative. As a result, both responsiveness coefficients discontinuously drop. 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.

To summarize, in this section we fully characterized optimal information structures that inform imperfectly competing firms about market demand conditions. We showed that these structures induce normally distributed prices that are linear in consumer types. The optimal price responsiveness depends on the weight that the designer attaches to consumer surplus relative to firms' profits. For low consumer weight, optimal information induces coordinated pricing and is not far from full information. For high consumer weight, optimal information induces anticoordinated pricing. The shift between these two modes is discontinuous.

4.2 A Prediction Game

Thus far, we have considered only the case in which the optimally induced action profile is a deterministic function of a state, i.e., the designer does not use randomization in the optimal information structure. One may wonder, especially given the private noise results of Section 3.4, whether this is a general feature of certifiable information structures. In this section, we show that this is not the case. We introduce and study a stylized prediction game in which a certifiably optimal information structure features exogenous noise.

The setting is as follows. The state is one-dimensional and normally distributed, $\theta \sim N(0,1)$. There are $N \geq 2$ players, each of whom tries to match his action with the state:

$$u_i(a,\theta) = -\frac{1}{2}(a_i - \theta)^2. \tag{42}$$

The designer would like to, on the one hand, coordinate the average action with the state and, on the other hand, make the individual actions sufficiently dispersed; her payoff function is

$$v(a,\theta) = \frac{\sum_{i=1}^{N} a_i}{N}\theta - 2\rho\frac{\sum_{i=1}^{N}\sum_{j\neq i} a_i a_j}{N^2}. \tag{43}$$

The first element of (43) captures the designer's willingness to correlate the players' average action with the state and corresponds to the coordination motive. The second element captures the designer's willingness to anticorrelate the actions of different players and corresponds to the anticoordination motive. The parameter ρ captures the intensity of the latter and is assumed to be sufficiently high, $\rho \geq \frac{N}{2N-1}$, such that the anticoordination motives are sufficiently important.

Given the designer's incentives, it is natural to conjecture that her objectives may be achieved by introducing exogenous noise to each player's estimate that is independent of the state but negatively correlated across players. The next proposition confirms this intuition and pins down the parameters of an optimal information structure.

Proposition 4. (Optimal Prediction Information) *The information structure that recommends to each player the following action as the function of a state is optimal:*

$$a_i(\theta) = \left(\frac{1}{2\rho} + \frac{1}{2N}\right)\theta + \varepsilon_i - \frac{1}{N-1}\sum_{j\neq i}\varepsilon_j, \tag{44}$$

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

Proof. To prove this result, we cannot apply Theorem 2 because it requires the existence of unique best responses in the dual problem, which would preclude randomization. Instead, we directly use Proposition 1.

The allocation rule (44) is incentive compatible, $\mathbb{E}[\theta|a_i] \equiv a_i$. Indeed, $a_i(\theta)$ and θ are jointly normally distributed.Therefore, the incentive compatibility is

$$\mathbb{E}[a_i(\theta - a_i)] = 0, \tag{45}$$

the condition satisfied for the chosen variance σ_ε^2.

To certify the optimality of an incentive-compatible rule, it suffices to find a contract that implements it in the dual problem (10). Consider a linear contract with $\lambda_i(a_i) = -\frac{\rho}{N^2}a_i$. The dual payoff of the agent in state θ becomes:

$$v(a,\theta) - \sum_i \lambda_i(a_i)u_i(a,\theta) = \frac{\sum_{i=1}^N a_i}{N}\theta - 2\rho\frac{\sum_{i=1}^N \sum_{j\neq i} a_i a_j}{N^2} + \sum_{i=i}^N \lambda_i(a_i)(a_i - \theta) \tag{46}$$

$$= \left(\frac{\sum_{i=1}^N a_i}{N}\right)\left(1 + \frac{\rho}{N}\right)\theta - \rho\left(\frac{\sum_{i=1}^N a_i}{N}\right)^2. \tag{47}$$

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

$$\frac{\sum_{i=1}^N a_i}{N} = \frac{N+\rho}{2\rho N}\theta, \tag{48}$$

including the profile specified by rule (44). Consequently, according to Proposition 1, the proposed contract implements the candidate allocation rule and certifies the optimality of the information structure that recommends it. □

According to Proposition 4, 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 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. The precision decreases as the designer's anticoordination motives increase. At $\rho = \frac{N}{2N-1}$, providing perfect information is optimal, $\sigma_\varepsilon^2 = 0$. As $\rho \to \infty$, the noise variance σ_ε^2 converges to $\frac{(N-1)(2N-1)}{4N^2}$: to achieve maximally anticoordinated actions, the designer must provide moderately useful information. This simple Gaussian information structure is optimal across all possible information structures, however complex.

Under the optimal information structure, the average players' action is perfectly informative about the state. This feature is achieved with finitely many players by precise coordination of their individual noises. In contrast, in games with infinitely many players and independent noise, the same feature could appear due to the law of large numbers. To make a comparison, we consider the limit of our setting as the number of players goes to infinity.

When $\rho > 1/2$ and $N \to \infty$, the correlation component of the individual noise in (44) vanishes almost surely by the law of large numbers, and the recommended allocation rule converges to

$$a_i(\theta) = \frac{1}{2\rho}\theta + \varepsilon_i, \tag{49}$$

where $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ and $\sigma_\varepsilon^2 = \frac{2\rho-1}{4\rho^2}$. In the limit, each player is provided a conditionally independent estimate of the state, and the accurate behavior of the average action is ensured by the large population size rather than by noise correlation. This analysis suggests that the Gaussian information structures with conditionally independent noises commonly considered in the literature, e.g., by Bergemann and Morris (2013), may indeed be optimal among all information structures in games with a continuum of players. However, as our analysis clarifies, such structures fail to be optimal in a finite economy.

5 Conclusion

In this paper, we developed a solution method for information-design problems in concave games. This method builds upon the duality between the information-design problem 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 and the role of private noise in prediction games. We expect our analysis to open many possibilities for future studies of information design in games.

References

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

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.*

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.

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.

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

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

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

A Appendix

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 $\times_i X_i$.

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 certain responsiveness of players' actions to the state. Namely, 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 (5) is equal to the optimal value of the dual adversarial-contracting problem (10):*

$$V^P = V^D. \tag{50}$$

Moreover, in this case, an optimal value in (5) is achieved by some information structure. If, in addition, Assumption 3 holds, then an optimal value in (10) is achieved by some contract.

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

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

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

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:

$$\max_{\pi \in M(A\times\Omega)}(-\phi^*(-\pi) - \psi^*(\pi)) = \inf_{f \in C(A\times\Omega)}(\phi(f) + \psi(f)), \tag{51}$$

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 (51) 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 (51) 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. $\square$

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

Proof. We can write:

$$\begin{aligned}
-\phi^*(-\pi) &= -\sup_{f\in C(A\times\Omega)} \left(\int f\mathrm{d}(-\pi) - \phi(f)\right) \\
&= \inf_{f\in C(A\times\Omega), f\geq v} \left(\int f\mathrm{d}\pi\right) \\
&= \begin{cases} \int v\mathrm{d}\pi, & \text{if } \pi \in \Delta(A\times\Omega), \\ -\infty, & \text{o/w.} \end{cases}
\end{aligned}
\qquad
\begin{aligned}
-\psi^*(\pi) &= -\sup_{f\in C(A\times\Omega)} \left(\int f\mathrm{d}\pi - \psi(f)\right) \\
&= \inf_{f\in C(A\times\Omega)} \left(-\int f\mathrm{d}\pi + \psi(f)\right) \\
&= \begin{cases} 0, & \text{if (6) and (7) hold,} \\ -\infty, & \text{o/w.} \end{cases}
\end{aligned}$$

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 (6) and (7) hold, then $-\int f \mathrm{d}\pi + \psi(f) \geq 0$. Furthermore, the zero value can be achieved by setting $f \equiv 0$. If (6) or (7) 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 (5). The result follows. □

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 (5) and to the dual problem (8) exist.

Proof. For the primal problem (5), 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 (51) and the proof of Lemma 2.)

For the dual problem (8), define

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

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 (8): 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 pointwise 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. $\square$

As argued in the main text, the dual problem (8) is equivalent to the dual adversarial contracting problem (10). This concludes the proof of Theorem 3.

Proof of Proposition 2. We showed in the main text that under the first set of conditions, the linear allocation rule is implementable in the primal problem. It remains to show that under the second set of conditions, the linear allocation rule is implementable in the dual adversarial problem. The dual problem can be written as:

$$\min_{\lambda} \ \mathbb{E}_{\mu_0}[\max_{a \in A} v(a, \omega) - \sum_i \lambda_i(a_i)(C_{i\bullet} a - B_{i\bullet}\omega - b_i)] \tag{52}$$

$$\min_{\lambda} \ \mathbb{E}_{\mu_0}[\max_{a \in A} a^T(\hat{b} + \hat{B}\omega) - \frac{1}{2} a^T \hat{C} a - \lambda^T(a)(Ca - B\omega - b)]. \tag{53}$$

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

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

where $D(x)$ is a diagonal matrix with $D(x)_{ii} = x_i$. This is a quadratic optimization problem. Under the conditions of the theorem, matrix $\hat{C} + 2D(x)C$ is positive definite; hence, the agent's best response can be found via the system of first-order conditions

$$a^*(\omega) = (\hat{C} + 2D(x)C)^{-1}(\hat{b} + D(x)b - C^T x_0) + (\hat{C} + 2D(x)C)^{-1}(\hat{B} + D(x)B)\omega. \tag{54}$$

The conditions of the theorem ensure that x implements the best response with the responsiveness matrix R. It remains to show that we can 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 independent of x_0. However, according to the conditions of the theorem, $(\hat{C} + 2D(x)C)^{-1}$ 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. The result follows.

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 (23), 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:

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

Comparing it with the payoff function (19), 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 (19), 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 Disclosure and No Disclosure: Equilibrium pricing behavior satisfies and can be derived from the system of first-order conditions

$$\mathbb{E}_{\mu}[q_i(a_i, a_{-i}, \theta) + \frac{\partial q_i(a_i, a_{-i}, \theta)}{\partial a_i}(a_i - 2ca_i)] = 0, \quad i = 1, 2, \tag{55}$$

after substituting the linear form (27) of demand function q_i and setting the belief μ equal to μ_0 for no disclosure equilibrium and equal to the belief concentrated on θ for full disclosure 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:

$$
\begin{aligned}
f(x) &= b_0 + b_1 x + b_2 x^2 + b_3 x^3 + b_2 x^4, \\
b_0 &= -8c^3(\delta-1)^3(\xi-\eta)(\eta+\xi)(\delta\eta^4 + (\delta-1)\xi^4 + 3(2\delta-1)\eta^2\xi^2) \\
&\quad + 4c^2(\delta-1)^2\eta(\delta(5-8\delta)\eta^4 + (\delta(8\delta-11)+4)\xi^4 - 2(\delta(16\delta-19)+6)\eta^2\xi^2) \\
&\quad + 2c(3\delta-2)(\delta-1)(\delta(7\delta-4)\eta^4 + (\delta-1)\delta\xi^4 + (3\delta(8\delta-9)+8)\eta^2\xi^2) \\
&\quad + (2\delta-1)(-(2-3\delta)^2)\eta(\delta\eta^2 + (3\delta-2)\xi^2), \\
b_1 &= 32c^5(\delta-1)^3(\eta^2+\xi^2)(\eta^3-\eta\xi^2)^2 \\
&\quad - 16c^4(\delta-1)^2\eta(\xi-\eta)(\eta+\xi)((8-11\delta)\eta^4 + 2(\delta-1)\xi^4 + (2-3\delta)\eta^2\xi^2) \\
&\quad + 8c^3(\delta-1)(\xi-\eta)(\eta+\xi)((-42\delta^2+64\delta-25)\eta^4 + (\delta-1)^2\xi^4 + \delta(2-3\delta)\eta^2\xi^2) \\
&\quad + 4c^2((\delta(2(78-31\delta)\delta-133)+38)\eta^5 + (\delta-1)(\delta(\delta+7)-6)\eta\xi^4 \\
&\quad + (\delta(5\delta(25\delta-58)+226)-60)\eta^3\xi^2) - 2c((\delta((79-17\delta)\delta-88)+28)\eta^4 + (\delta-1)\delta(5\delta-4)\xi^4 \\
&\quad + 2(\delta(\delta(102\delta-211)+146)-34)\eta^2\xi^2) + (3\delta-2)\eta((\delta(7\delta+4)-4)\eta^2 + (\delta(41\delta-52)+16)\xi^2), \\
b_2 &= -2((\eta-1)\eta^3(\delta^2(8(\eta-2)\eta(10(\eta-2)\eta+13)+29) - 4\delta(\eta-1)(2\eta-1)(20(\eta-2)\eta+13) \\
&\quad + 20(2\eta^2-3\eta+1)^2) + 2(\delta-1)(1-2\eta)^2 r^4(2(\delta-1)\eta^2 - \delta + 2\eta) \\
&\quad + \eta r^2(\delta^2(\eta(4\eta(\eta(8(11-3\eta)\eta-145)+135)-271)+56) \\
&\quad + 2\delta(\eta-1)(4\eta(\eta(8\eta(3\eta-7)+63)-37)+37) - 8(\eta-1)(3(\eta-1)\eta+1)(4(\eta-1)\eta+3))), \\
b_3 &= -16\eta(1-2c\eta)^2(c\eta-1)(\xi^2(4c^2(\delta-1)\eta^2 + 2c(2-3\delta)\eta + 3\delta - 2) \\
&\quad - 2\eta^2(c\eta-1)(2c(\delta-1)\eta - 3\delta + 2)), \\
b_4 &= -8\eta(1-2c\eta)^2(c\eta-1)(4\eta^2(c\eta-1)^2 - \xi^2(1-2c\eta)^2).
\end{aligned}
$$

For the parameters of the numerical example, the polynomial becomes

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

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

$$
\delta^{cr} = \frac{2(c(|\xi|+\eta)(-2c\eta|\xi| + |\xi| + \eta(-2c\eta+3)) - \eta)}{-|\xi|(-2c\eta(-4c\eta+5)+1) - \eta(4c - \eta(-c\eta+2)+5) + 2c\xi^2(-2c\eta+1)} = \frac{11}{18}. \quad (56)
$$

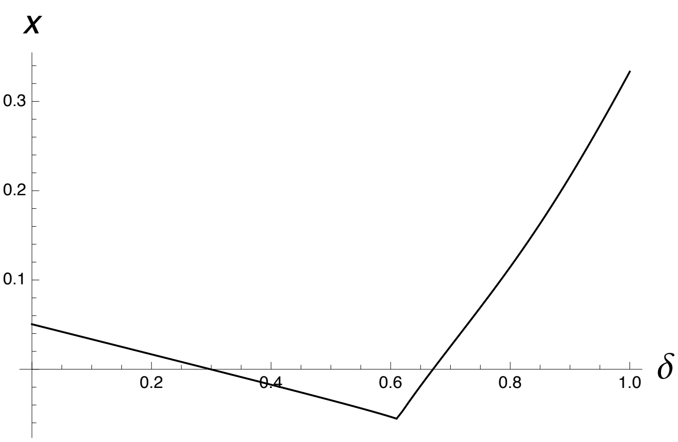

Figure 4: Parameter x that certifies an optimal information structure plotted as a function of the consumer surplus weight δ.