

# Quadratic Games\*

Nicolas S. Lambert<sup>†</sup>      Giorgio Martini<sup>‡</sup>      Michael Ostrovsky<sup>§</sup>

First version: February 24, 2017

This version: December 12, 2017

## Abstract

We study general quadratic games with multi-dimensional actions, stochastic payoff interactions, and rich information structures. We first consider games with arbitrary finite information structures. In such games, we show that there generically exists a unique equilibrium. We then extend the result to games with possibly infinite information structures, under an additional assumption of linearity of certain conditional expectations. In that case, there generically exists a unique linear equilibrium. In both cases, the equilibria can be explicitly characterized in closed form. We illustrate our framework by studying information aggregation in large asymmetric Cournot markets and by considering the effects of stochastic payoff interactions in beauty contest games. Our results apply to general games with linear best responses, and also allow us to characterize the effects of small perturbations in arbitrary Bayesian games with finite information structures and smooth payoffs.

---

\*We thank Kostas Bimpikis, Johannes Hörner, Matt Jackson, David Myatt, Alessandro Pavan, Andy Skrzypacz, Xavier Vives, Bob Wilson, Anthony Lee Zhang, and seminar participants at Stanford, UIUC, the 22nd Coalition Theory Network Workshop, and the 2017 Workshop on Markets with Information Asymmetries at Collegio Carlo Alberto for helpful comments and suggestions. Lambert is grateful to Microsoft Research New York and the Cowles Foundation at Yale University for their hospitality and financial support.

<sup>†</sup>Stanford Graduate School of Business; nlambert@stanford.edu.

<sup>‡</sup>Stanford Graduate School of Business; gmartini@stanford.edu.

<sup>§</sup>Stanford Graduate School of Business and NBER; ostrovsky@stanford.edu.

# 1 Introduction

We study general quadratic games with multidimensional actions, stochastic payoff interactions, and rich information structures. After privately observing their signals, players simultaneously choose vectors of real-valued actions to maximize quadratic payoff functions whose terms may be stochastic.

We consider two types of information structures. In the first, any finite information structure is allowed: in particular, the signals of different players may be arbitrarily correlated, may overlap, may contain information about the signals of other players, and so on—we allow for the same level of generality as the framework of Aumann (1976), subject only to the finiteness of the signal space. In that setting, we show that generically, there exists a unique Bayesian Nash equilibrium, and give a closed-form characterization of equilibrium strategies.

We then consider information structures that combine finite and (potentially) infinite signals. The joint distribution of the finite components of the signals is arbitrary, as in the previous case. The infinite components have the restriction that certain conditional expectations must be linear. In particular, this setup includes, as a special case, the Linear-Quadratic Gaussian setup, where (infinite) signals are distributed according to a multivariate normal distribution. By also allowing for arbitrary finite signal components, we can capture various stochastic interactions and heterogeneous information about the infinite signals (for example, about their correlations and observability across players). We show that generically, there exists a unique linear Bayesian Nash equilibrium, and give a closed-form characterization of equilibrium strategies.

To illustrate our framework and results, we present two sets of applications.

First, we consider Cournot competition with quadratic production costs and linear demand functions, both of which may be only partially known by firms. This model fits our framework with finite signals, and as in that framework, we allow for arbitrary information structures about the costs and about the demand function. In particular, we allow not just the intercept but also the slope of the demand function to be only partially known to the firms. We use our general results to show that in this setting, an equilibrium of any Cournot competition game always exists and is unique, and characterize it in closed form. We then prove two results about decentralized efficiency in large asymmetric markets. First, as the market grows large, equilibrium becomes ex-ante efficient. Second, when costs are linear, equilibrium forces act as an “invisible statistician”: the equilibrium quantity converges to the best linear estimate of the ex-post socially efficient quantity, given the information available to the firms.

Second, we analyze beauty contest games in which players may target different state variables and are allowed to have only incomplete information about the payoff structure of others. This generalization makes it possible to consider rich versions of beauty contest and coordination games: e.g., one can consider a coordination game on a network in which each player knows who her friends are, but has only partial information about the rest of the network, the strengths of the connections there, etc. In this framework, we show that there always exists a unique linear equilibrium, and characterize it in closed form. We illustrate the framework with an example that shows some of

the novel effects that arise when the strengths of payoff interactions in a beauty contest game are stochastic.

We also present several additional findings. First, we discuss how our closed-form characterizations (both for the finite and infinite information cases) further simplify if payoff interactions between players' actions are deterministic rather than stochastic (but there is still uncertainty about the random variables that directly affect the players' payoffs, as in, e.g., the conventional beauty contest games or in Cournot competition games in which the slope of the demand function is known). Second, we discuss the applicability of our results to general games with linear best responses. Finally, we show how our framework can be used to characterize the effects of small perturbations in *arbitrary* Bayesian games with finite information structures and smooth payoffs. For the last application, it is essential that the framework allows for stochastic payoff interaction terms; without this feature, such characterization would not be possible.

The paper is organized as follows. In Section 1.1, we briefly review the related literature on general Linear-Quadratic Gaussian games (the literatures on Cournot competition and beauty contests are briefly reviewed in the corresponding sections). We present the model for arbitrary finite information structures in Section 2, and prove the existence and uniqueness of Bayesian Nash equilibrium. We apply this framework to the questions of equilibrium characterization and information aggregation in Cournot markets in Section 3. In Section 4, we present the richer model with both finite and potentially infinite signals. Applications to beauty contests are presented in Section 5. In Section 6, we present the simplified closed-form characterization for the case in which payoff interaction coefficients are deterministic. Finally, in Section 7, we discuss the applications of our framework to general games with linear best responses and to the characterization of the effects of small perturbations in general Bayesian games.

## 1.1 Related Literature on General Linear-Quadratic Gaussian Games

A large literature spanning multiple fields has investigated general Linear Quadratic Gaussian (LQG) games. In this section, we briefly review theoretical work on the properties of LQG games, highlighting the relative contributions of our paper. We review the literatures related to our applications to Cournot competition and to beauty contests in Sections 3 and 5, respectively.

The foundational paper of Radner (1962) formalizes team problems as multi-agent decision problems where agents share a common goal but information is decentralized. Using Hilbert space methods, he proves that under the quadratic payoff specification the person-by-person optimal decision policies are linear in signals. Ui (2009) introduces Bayesian potential games in order to apply Radner's methods to games where the common goal assumption does not hold. The drawback of this approach is that stringent symmetry assumptions are required for the games to admit a potential function. In contrast, our paper allows for possibly asymmetric payoff structures. Pachter and Pham (2013) also revisit Radner (1962), prove necessary and sufficient conditions for the existence and uniqueness of the linear equilibrium in a two-player environment, and relate the Bayesian Nash equilibrium to the centralized optimum.

The most general model present in the literature is Basar (1978a), which generalizes the earlier work in Basar (1975) from 2 to  $n$  players. Like us, he allows for an arbitrary number of players, potential asymmetry in the payoff structure, and multidimensional actions and signals. He employs a contraction-mapping approach that gives sufficient conditions for equilibrium uniqueness in general. However, this approach does not give necessary and sufficient conditions under which the linear equilibrium exists and is unique. Furthermore, in contrast to Basar (1978a), we provide closed-form solutions for the unique linear equilibrium. Basar (1978b) studies 2-player quadratic games in which payoff interaction terms are allowed to be stochastic. Like Basar (1978a), Basar (1978b) only gives sufficient conditions for equilibrium existence, and does not provide closed-form solutions.

A recent paper that also studies general quadratic games with rich asymmetric information structures is Bergemann, Heumann, and Morris (2017). The focus of that paper, however, is different from ours. Bergemann, Heumann, and Morris (2017) focus on games with single-dimensional actions, fixed payoff interaction terms, and Gaussian information structures, and explore such topics as identification in quadratic games; the range of possible outcomes as the information structure varies; and the connections to Bayes Correlated Equilibrium (Bergemann and Morris, 2013, 2016). By contrast, we provide existence, uniqueness, and closed-form characterization results for games with multi-dimensional actions, stochastic payoff interaction terms, and general information structures with linear conditional expectations. This level of generality allows us to explore such applications as Cournot competition games with uncertain slope of demand, beauty contest games in which the weights that players put on matching others' actions are stochastic, and the characterization of the effects of small perturbations in arbitrary Bayesian games with finite information structures and smooth payoffs.

## 2 Games with Finite Information Structures

In this section, we state and prove our first main result: generic equilibrium existence and uniqueness for quadratic games with arbitrary finite information structures. We impose no restrictions on the information structure (beyond finiteness), which enables us to represent arbitrary informational asymmetries and interdependencies.<sup>1</sup> In Section 4, we present parallel results for games with potentially infinite state and signal spaces, under some additional assumptions.

### 2.1 Model

There are  $n$  players,  $i = 1, \dots, n$ . Prior to taking an action, each player  $i$  privately observes a signal  $s_i$  whose set  $S_i$  of possible realizations is finite and contains  $k_i \geq 1$  elements. We denote by  $s = (s_1; s_2; \dots; s_n)$  the random vector summarizing the signals of all  $n$  players. Players share a common prior belief  $P$  about the joint probability distribution over signal vectors  $s$ . Without loss of generality, we assume that for every player  $i$  and realization  $\tilde{s}_i \in S_i$ , the probability of player  $i$

---

<sup>1</sup>Our informational framework is equivalent to the canonical “partitional” information structure of Aumann (1976) with a finite underlying state space, although its mathematical description is slightly different.

observing signal  $\tilde{s}_i \in S_i$  is positive. We do not impose any other restrictions on distribution  $P$ . As in Aumann (1976), this lack of restrictions is the key feature that allows our framework to accommodate arbitrary information asymmetries and interdependencies, by appropriately choosing the joint distribution  $P$  over signal vectors  $s$ .

## 2.2 Actions and Payoffs

After observing his signal  $s_i$ , each player  $i$  chooses action  $a_i \in \mathbb{R}^{m_i}$ , where  $m_i \geq 1$  is the dimension of player  $i$ 's action. The payoff of player  $i$  depends on his own action  $a_i$ , the actions of other players (denoted by vector  $a_{-i}$ ), and the vector of signals  $s$ . Specifically,

$$\pi_i(a_i, a_{-i}, s) = \frac{1}{2}a_i^T \Gamma_{ii}(s)a_i + a_i^T \Gamma_{i,-i}(s)a_{-i} + a_i^T g_i(s) + h_i(a_{-i}, s), \quad (1)$$

where  $\Gamma_{ii}(s)$  is a matrix of size  $m_i \times m_i$ ,  $\Gamma_{i,-i}(s)$  is a matrix of size  $m_i \times (\sum_{j \neq i} m_j)$ ,  $g_i(s)$  is a vector of size  $m_i$ , and  $h_i$  is an arbitrary function of  $a_{-i}$  and  $s$ . Note that the presence of function  $h_i$  does not affect the incentives of player  $i$  (and thus the equilibria of the game), but does in general affect the efficiency and welfare properties of various strategy profiles.

The only assumption we impose on the parameters of payoff functions  $\pi_i$  is that for each  $i$  and each signal realization  $\tilde{s}_i \in S_i$ , the conditional expectation  $E[\Gamma_{ii}(s)|s_i = \tilde{s}_i]$  is a symmetric and negative definite matrix. The symmetry assumption is made for notational convenience and is without loss of generality. The assumption that the matrix is negative definite is substantive: it ensures that player  $i$ 's optimization problem is well-defined and has a finite solution. In the case of single-dimensional actions, this assumption reduces to the conditional expectation of the quadratic coefficient in equation (1) being negative.

We impose no other assumptions on the parameters of payoff functions  $\pi_i$ . In particular, the dimensions of different players' action spaces may be different; the interaction terms  $\Gamma_{i,-i}$  between different players' actions can be different and asymmetric, and can vary arbitrarily as functions of the entire vector of signals  $s$ ; and functions  $g_i$ , which determine the direct payoff interactions between the vector of signals  $s$  and the players' actions, can be arbitrary as well.

## 2.3 Bayesian Nash Equilibrium

A profile of pure strategies  $a_i(\cdot)$  of all players is a Bayesian Nash equilibrium if for every player  $i$ , for every signal realization  $\tilde{s}_i$ , action  $a_i(\tilde{s}_i)$  maximizes player  $i$ 's expected payoff, conditional on the realization of the signal, given the primitives of the game and the strategies of other players.<sup>2</sup>

## 2.4 Equilibrium Existence, Uniqueness, and Characterization

We can now state and prove the first main result of the paper.

---

<sup>2</sup>In general, Bayesian games may of course have mixed-strategy equilibria in addition to pure-strategy ones, but as we will note in the proof of Theorem 1, in our setting best responses are always unique. Thus, there cannot be any mixed-strategy equilibria, and so to simplify notation we only talk about pure strategy profiles and equilibria.

**Theorem 1** *Generically, there exists a unique Bayesian Nash equilibrium.*

The remainder of this section presents the proof of Theorem 1 and the closed-form characterization of the unique Bayesian Nash equilibrium. The proof consists of two parts. Lemma 1 provides a condition that guarantees equilibrium existence and uniqueness. It also provides closed-form expressions for the equilibrium strategies when the condition is satisfied. Lemma 2 shows that the condition for existence and uniqueness required in Lemma 1 holds generically.

To state Lemma 1, we need to introduce additional notation.

Let  $\Gamma_{ij}(s)$  be the matrix of size  $m_i \times m_j$  corresponding to the terms multiplying actions  $a_i^T$  and  $a_j$  of players  $i$  and  $j$  in equation (1), for a particular vector of signals  $s$ .

Let  $\Phi$  be a block matrix defined as follows. Enumerate the  $k_i$  possible signal realizations of player  $i$  as  $1, 2, \dots, k_i$ , and let  $K = k_1 + \dots + k_n$ . Matrix  $\Phi$  consists of  $K \times K$  blocks: each block row corresponds to one player  $i$  and one possible signal realization of that player,  $\tilde{s}_i$ ,<sup>3</sup> and analogously, each block column corresponds to one player  $j$  and one possible signal realization of that player,  $\tilde{s}_j$ . The block of matrix  $\Phi$  in block row  $(i, \tilde{s}_i)$  and block column  $(j, \tilde{s}_j)$  is given by

$$\Phi_{(i, \tilde{s}_i), (j, \tilde{s}_j)} = \mathbb{E}[\Gamma_{ij}(s) | s_i = \tilde{s}_i, s_j = \tilde{s}_j] \cdot \mathbb{P}[s_j = \tilde{s}_j | s_i = \tilde{s}_i].$$

Note that the size of block  $\Phi_{(i, \tilde{s}_i), (j, \tilde{s}_j)}$  is the same as the size of matrix  $\Gamma_{ij}(s)$ :  $m_i \times m_j$ . Also, when  $i = j$ , then  $\Phi_{(i, \tilde{s}_i), (i, \tilde{s}_i)} = \mathbb{E}[\Gamma_{ii}(s) | s_i = \tilde{s}_i]$ , and when  $\tilde{s}_i \neq \tilde{s}_i$ , then  $\Phi_{(i, \tilde{s}_i), (i, \tilde{s}_i)} = 0$ .

**Condition 1** *Matrix  $\Phi$  is invertible.*

**Lemma 1** *The game has a unique Bayesian Nash equilibrium if and only if Condition 1 is satisfied.*

The proof of Lemma 1 is in Appendix A.1. In addition to showing the result, the proof also shows that when Condition 1 is not satisfied, then either the game has no Bayesian Nash equilibrium, or it has infinitely many. Also, the proof provides a convenient closed-form solution for the equilibrium when the condition is satisfied. Specifically, slightly abusing notation, let  $a_i \in \mathbb{R}^{m_i k_i}$  denote a “stacked” vector of player  $i$ ’s actions, one for each possible realization of signal  $s_i$  (i.e., the first  $m_i$  elements of vector  $a_i$  are player  $i$ ’s action after observing  $s_i = 1$ , the next  $m_i$  elements of  $a_i$  are the action after observing  $s_i = 2$ , and so on). Let vector  $a = (a_1; a_2; \dots; a_n)$  denote the combined profile of all players’ actions. Similarly, let  $g_i \in \mathbb{R}^{m_i k_i}$  denote a stacked vector of player  $i$ ’s conditional expectations of  $g_i(s)$ , one for each realization of signal  $s_i$  (i.e., the first  $m_i$  elements of vector  $g_i$  are equal to  $\mathbb{E}[g_i(s) | s_i = 1]$ , the next  $m_i$  elements of  $g_i$  are equal to  $\mathbb{E}[g_i(s) | s_i = 2]$ , and so on). Let  $g = (g_1; g_2; \dots; g_n)$  denote the vector combining the vectors  $g_i$  of all individual players  $i$ .

When Condition 1 is satisfied, the unique Bayesian Nash Equilibrium is given by

$$a = -\Phi^{-1}g. \tag{2}$$

---

<sup>3</sup>I.e., block row 1 corresponds to the first signal realization of player 1, block row 2 corresponds to the second signal realization of player 1,  $\dots$ , block row  $k_1$  corresponds to the last signal realization of player 1, block row  $(k_1 + 1)$  corresponds to the first signal realization of player 2, and so on.

The second step of the proof of Theorem 1 shows that Condition 1 holds generically. Formally, consider the following one-dimensional collection of quadratic games, parameterized by  $\gamma \in \mathbb{R}$ . The information structure is the same for all the games in the collection, and payoffs are given by

$$\pi_i(a_i, a_{-i}, s) = \frac{1}{2}a_i^T \Gamma_{ii}(s)a_i + a_i^T (\gamma \Gamma_{i,-i}(s)) a_{-i} + a_i^T g_i(s) + h_i(a_{-i}, s). \quad (3)$$

**Lemma 2** *Condition 1 holds for all  $\gamma \in \mathbb{R}$ , except for at most finitely many values.*

The proof of Lemma 2 is in Appendix A.2. Together with Lemma 1, this concludes the proof of Theorem 1.

### 3 Application: Cournot Competition

In this section, we apply the results of Section 2 to Cournot markets with incomplete information. Our model uses the classical linear demand functions and linear-quadratic costs that are used in Palfrey (1985) and Vives (1988), also in the context of incomplete information. While these two papers use symmetric information structures in which firms receive i.i.d. signals (conditionally on the state of the world), we leverage the framework of the previous section to derive implications for the case of arbitrary finite information structures.

Rich information structures in Cournot markets are also considered by Lambert, Ostrovsky, and Panov (2017a), who study information aggregation properties of large asymmetric markets. They assume that the firms' information is jointly normally distributed, and show that as the market grows large, information gets aggregated. As our results in Section 3.3 demonstrate, this result depends on the property of jointly normal distributions that one random variable can be “decomposed” as a linear combination of other jointly normal random variables and idiosyncratic noise. Without this property, under general information structures, full information aggregation does not occur. Instead, the decentralized market serves, in effect, as an “invisible statistician”, with equilibrium market quantities and prices providing optimal linear estimators of social welfare-maximizing quantities and prices.

We are not the first to step away from the assumption of joint normality of signals and parameters in Cournot competition with incomplete information. Li et al. (1987) point out that to obtain tractable results in a Cournot setting, it is sufficient to consider distributions for which certain conditional expectations are linear. They also note the advantages of certain non-Gaussian distributions, which make it possible to circumvent the issues with negative quantities and prices that can arise under the Gaussian specification (Hurkens, 2014). However, to the best of our knowledge, ours is the first specification that allows tractable analysis of linear-demand settings in which firms are uncertain not only about the intercept of the demand function, but also about its slope. Note also that analogously to the current setting, the framework of Section 2 can also be used to study models of differentiated Bertrand and differential Cournot competition with multi-product firms and rich asymmetric information structures about various demand and cost parameters.

### 3.1 Model of Cournot Competition

There is a market with  $n$  firms,  $i = 1, \dots, n$ . The firms produce and sell a good to a population of consumers. Each firm observes a finite signal  $s_i \in S_i$ , using the notation of Section 2.1. In addition, there is a state of the world  $\omega$  that takes values in a finite set  $\Omega$ . For each firm, the marginal probability distribution over signals has full support. The joint distribution of signals and state of the world is otherwise arbitrary.

After each firm  $i$  has observed its signal  $s_i$ , it decides on a quantity to produce,  $q_i$ ; this is the player's action in our general framework. Each firm incurs cost of production  $c_i(\omega)q_i + d_i(\omega)q_i^2$ , with  $d_i(\cdot) \geq 0$  and  $c_i(\cdot) \geq 0$ . Consumer demand is linear, and the inverse demand is

$$P = \alpha(\omega) - \beta(\omega)Q,$$

where  $\alpha(\cdot) > 0$  and  $\beta(\cdot) > 0$ . For a given profile of quantities  $q_1, \dots, q_n$  supplied by the firms, and a given state of the world, the realized profit of firm  $i$  is

$$\pi_i = (\alpha(\omega) - \beta(\omega)Q)q_i - c_i(\omega)q_i - d_i(\omega)q_i^2$$

where  $Q = \sum_j q_j$  is the total quantity supplied by the firms. This model of Cournot competition is therefore a special case of our framework of Section 2.<sup>4</sup> We refer to the game just described as a *Cournot game*.

The following proposition shows that there always exists a unique equilibrium in a Cournot game. Note that this statement is stronger than the statement of Theorem 1, because it does not rely on genericity.

**Proposition 1** *There exists a unique Bayesian Nash equilibrium of the Cournot game.*

The proof of Proposition 1 is in the appendix. The idea of the proof is as follows. First, we show that in the special case  $\alpha(\cdot) = c_i(\cdot) = 0$  for every  $i$ , all firms producing zero (for all signal realizations) is the unique Bayesian Nash equilibrium. Thus, by Lemma 1, Condition 1 is satisfied. Second, we observe that in the general case, with no conditions on  $\alpha(\cdot)$  and  $c_i(\cdot)$ , matrix  $\Phi$  (defined in Section 2.4), which determines existence and uniqueness of equilibrium via Condition 1, does not depend on  $\alpha(\cdot)$  and  $c_i(\cdot)$ . Hence, since Condition 1 is satisfied in the special case  $\alpha(\cdot) = c_i(\cdot) = 0$  for every  $i$ , it is also satisfied in the general case. We conclude by Lemma 1 that there exists a unique Bayesian Nash equilibrium in the general case.

---

<sup>4</sup> Strictly speaking, our Cournot competition model does not fit the framework of Section 2.1, as it involves a state of the world that is not part of the finite signals of the firms. This issue, however, is purely notational. To make the model fit the framework of Section 2.1, one can introduce an additional player, Nature, whose signal space is  $\Omega$  and whose payoff is given by, for example,  $\pi_N = -a_N^2$ , so that it always plays  $a_N = 0$ . Alternatively, without introducing an artificial player, one can replace the payoff functions  $\pi_i(a_i, a_{-i}, s, \omega)$  with their expectations conditional on the vector of firms' signals  $s$ :  $\pi_i(a_i, a_{-i}, s) = E[\pi_i(a_i, a_{-i}, s, \omega)|s]$ . Either approach brings the model into the formal framework of Section 2.1.



### 3.2 On the Efficiency of Large Markets

We now show that a property of invisible hand of competitive markets holds under incomplete information and arbitrary finite information structure, generalizing the work of Vives (1988) to arbitrary finite information structures. Our result also relates to Palfrey (1985), who gives conditions on the information structures for the first-best market outcome to be achieved in large Cournot markets for firms with constant marginal costs.

To establish the results of this subsection, we consider a sequence of markets of the type just described. The markets of the sequence are indexed  $k = 1, 2, \dots$ . Each market  $k$  corresponds to a market of Section 3.1 in which both the number of firms and consumer demand are scaled by a factor  $k$ . Specifically, market  $k$  has  $N$  groups of firms, and each group includes  $k$  firms. There continues to be a finite state of the world  $\omega$  taking values in  $\Omega$ . Every firm of group  $i$  observes the same finite signal  $s_i \in S_i$ , and incurs cost  $c_i(\omega)q_i + d_i(\omega)q_i^2$  when producing  $q_i$  units. Inverse demand in market  $k$  is

$$P = \alpha(\omega) - \beta(\omega) \frac{Q}{k}.$$

By Proposition 1, there exists a unique Bayesian Nash equilibrium in every market  $k$ .

In a Cournot environment with incomplete information, one must distinguish between two relevant notion of efficiency, which we refer to as ex-ante versus ex-post efficiency. The notion of ex-post efficiency assumes all uncertainty has resolved. Specifically, fixing a given state of the world  $\omega$ , a vector of production quantities is *ex-post efficient* when the total surplus is maximized among all possible vectors of production quantities. Thus, ex-post efficiency corresponds to a social planner who has full knowledge of the state of the world, and recommends to each firm an output level so as to maximize total surplus. In contrast, ex-ante efficiency captures the idea that the state of the world is not known by the firms at the time they must decide on production quantities. Furthermore, information is decentralized: each firm's decision can only depend on its own signal. Formally, we say that a strategy profile is *ex-ante efficient* when the expected total surplus is maximized among all firm strategy profiles. Hence, ex-ante efficiency corresponds to a social planner who can dictate the strategies to be followed by each firm, but does not have any particular information on the state of the world. When firms observe  $\omega$  directly then both ex-ante and ex-post efficiency reduce to the classical notion of efficiency under complete information. Our notion of ex-ante efficiency coincides with the notion of second-best decentralized efficiency of Vives (1988), also for Cournot games, and with the efficiency notion of Angeletos and Pavan (2007) for a broader class of quadratic games.

We let  $\text{TS}_k^a$  be the maximum expected total surplus that can be achieved in market  $k$  over all possible strategy profiles. Either  $\text{TS}_k^a$  is infinite for all  $k$ , or  $\text{TS}_k^a$  is proportional to the number of firms.<sup>5</sup> We restrict attention to the interesting case of finite maximum total surplus, and we let  $\overline{\text{TS}}^a$  be the normalized maximum expected total surplus per firm that can be achieved in any market  $k$ . Analogously, we let  $\text{TS}_k^*$  be the expected total surplus of the unique equilibrium in market  $k$ , and

---

<sup>5</sup>The possibility of unbounded total surplus is an artifact of the general quadratic framework from which our model of Cournot competition derives, and which does not restrict firms to nonnegative output levels. Such situation may occur when the costs are linear and the marginal costs vary across firms.

$\overline{\text{TS}}_k^*$  the expected total surplus per firm.

Under incomplete information, it is well-known that it is generally not possible for the market equilibrium to be ex-post efficient<sup>6</sup> (see, for example, Proposition 1 of Vives (1988)). Ex-ante efficiency is thus the natural benchmark for models of Cournot competition with incomplete information, as a second-best welfare property. Our next result shows that as markets grow large and many firms share the same information, the market equilibrium becomes ex-ante efficient.

**Proposition 2** *In the limit as the market size grows, the market equilibrium becomes ex-ante efficient:  $\lim_{k \rightarrow \infty} \overline{\text{TS}}_k^* = \overline{\text{TS}}^a$ .*

The decentralized efficiency of large Cournot markets was observed by Vives (1988) in the case of i.i.d. signals. Our result implies that this efficiency result does not rely on the symmetry of information.

The proof of Proposition 2 is in the appendix. It relies on the observation that the strategies which obtain in the Bayesian Nash equilibrium correspond to the strategies that a planner would want to enforce to maximize a biased expected total surplus, where the bias is additive and inversely proportional to market size. As the market size grows, the objective of the planner becomes close to the unbiased expected total surplus, and the equilibrium strategies yield a expected total surplus that, in turn, becomes close to the surplus obtained under an ex-ante efficient strategy profile.

### 3.3 The Invisible Statistician

Because information is incomplete, in general as the market grows large the equilibrium market quantity cannot converge to the ex-post efficient quantity. However, these two quantities are related in a statistical sense. In this section, we show that the equilibrium market quantity converges to the best linear estimate of the ex-post efficient quantity.

Throughout this section, marginal costs are constant and identical across firms:  $d_i(\cdot) = 0$  and  $c_i(\omega) = c(\omega)$  for every  $i$ . Let  $\mathcal{Q}$  denote the set of all real-valued mappings with vectors of firms' signals and states of the world as input. Elements of  $\mathcal{Q}$  are market quantities determined by the realization of the firms' signals and the state of the world; they can be interpreted as random variables. For the cost structures under consideration, both the ex-ante and ex-post efficient market quantities exist in every market  $k$ . Let us denote them by  $Q_k^a$  and  $Q_k^p$ , respectively. Note that  $Q_k^a, Q_k^p \in \mathcal{Q}$ . We continue to make use of the bar notation to denote per-firm averages, that is,  $\overline{Q}^p = Q_k^p/(Nk)$  and  $\overline{Q}^a = Q_k^a/(Nk)$ . We omit subscript  $k$  because per-firm average efficient quantities do not depend on the market  $k$ .

To state our result, we introduce the collection  $\mathcal{C} \subset \mathcal{Q}$  of additively separable mappings:

$$\mathcal{C} = \{Q \in \mathcal{Q} \text{ with } Q(s_1, \dots, s_n, \omega) = q_1(s_1) + \dots + q_N(s_N)\}.$$

---

<sup>6</sup>Palfrey (1985) shows that if marginal costs are constant and identical across firms, and if firms receive i.i.d. signals on the state of the world that determines the firms' costs and the intercept of the demand function, ex-post efficiency is achieved in large markets under certain general conditions on the distributions of signals and states.

Note that  $\bar{Q}^a \in \mathcal{C}$ , but that in general,  $\bar{Q}^p \notin \mathcal{C}$ . In the following, the variables of  $\mathcal{C}$  can be interpreted as possible statistical estimates of  $\bar{Q}^p$ . These estimates are additively separable in the signal values, and so correspond to classical statistical linear estimates in which each firm signal is a dummy variable. We refer to the elements of  $\mathcal{C}$  as *linear estimates*.

Consider the statistical question of finding a linear estimate  $\bar{Q}$  that minimizes the  $\beta$ -weighted quadratic error

$$E[\beta(\omega)(\bar{Q} - \bar{Q}^p)^2].$$

When such a linear estimate exists, we refer to it as *best linear estimate of  $\bar{Q}^p$  with respect to the  $\beta$ -weighted squared error*.

The following proposition shows how this statistical question relates to Cournot markets. It consists of two parts. The first part of the proposition shows that the ex-ante efficient production is the unique best linear estimate of the ex-post efficient production for the  $\beta$ -weighted quadratic error. The second part of the proposition shows that, in large markets, the equilibrium production per firm approximates the best linear estimate of the ex-post efficient production level. In the following,  $Q_k^*$  denotes the equilibrium quantity in market  $k$ , and let  $\bar{Q}_k^*$  denote the output per firm. Both are a function of the firms' signals.

**Proposition 3** *There exists a unique best linear estimate  $\bar{Q}^\dagger$  of  $\bar{Q}^p$  with respect to the  $\beta$ -weighted squared error, and the following obtains:*

1.  $\bar{Q}^a = \bar{Q}^\dagger$ .
2. For every realization of joint signals,  $\lim_{k \rightarrow \infty} \bar{Q}_k^* = \bar{Q}^\dagger$ .

As opposed to the previous result of Section 3.2 on the ex-ante efficiency of large markets, Proposition 3 does not apply when the costs are quadratic.

One direct implication of the above result is that, if (and only if)  $\bar{Q}^p \in \mathcal{C}$ , then at the limit, the large Cournot market becomes ex-post efficient. For the case of i.i.d. signals, and assuming constant and identical marginal costs, Palfrey (1985) shows that under fairly general conditions on distributions, ex-post efficiency of large market obtains. Our result suggests that, with asymmetric information, ex-post efficiency requires stronger conditions: ex-post efficiency only obtains under additive separability of the first-best output.

The proof of Proposition 3 is in the appendix. Uniqueness of the best linear estimate owes to the convexity of the squared error, and to the convexity of the set of linear estimates. The key to the first statement of the proposition is the observation that, when there are no quadratic costs, the problem of minimizing the  $\beta$ -weighted error is identical to the problem of maximizing the expected surplus, so that if the problems have a unique maximizer, the maximizers must be identical. The second statement of the proposition uses an argument related to the proof of Proposition 2. The equilibrium market quantity is the unique solution of an optimization problem that can be interpreted as the minimization of the  $\beta$ -weighted squared error with a bias. The magnitude of the bias vanishes as the market size grows, consequently the problem of finding the equilibrium market

quantity becomes very similar to the problem of finding the quantity that minimizes the  $\beta$ -weighted squared error, and the two maximizers become arbitrarily close.

## 4 Games with Infinite Information Structures

In some applications, using potentially infinite signal spaces may be more convenient than restricting attention to finite signal spaces. For example, it is common to assume that signals of players are jointly normally distributed. In this section, we consider a setting with potentially infinite signals spaces, with some additional assumptions on the distribution of signals and the payoff structure. We restrict attention to linear equilibria (in which the actions of players are linear functions of some of the signals). Our second main result shows that generically, there exists a unique linear equilibrium. This equilibrium can be characterized in closed form.

### 4.1 Model

There are  $n$  players,  $i = 1, \dots, n$ . Prior to taking an action, each player  $i$  privately observes a pair of signals:  $s_i \in S_i$  (where, as in Section 2, set  $S_i$  is finite and contains  $k_i$  elements) and  $\theta_i \in \mathbb{R}^{\ell_i}$  (where  $\ell_i \geq 1$  is the dimension of signal  $\theta_i$ ). We denote by  $s = (s_1; s_2; \dots; s_n)$  and  $\theta = (\theta_1; \theta_2; \dots; \theta_n)$  the vectors summarizing the signals of all  $n$  players. For convenience, we refer to  $s_i$ 's as "finite signals" of the players and  $\theta_i$ 's as "infinite signals" (although the latter are not required to have an infinite number of possible realizations).

We make the following assumptions about the joint distribution of signals  $(s, \theta)$ . First, for every player  $i$  and realization  $\tilde{s}_i \in S_i$ , the probability of player  $i$  observing  $\tilde{s}_i$  is positive. As in Section 2, this is without loss of generality. Also as in Section 2, we do not impose any other restrictions on the joint distribution of finite signals  $s$ .

Second, for every player  $i$  and realization  $\tilde{s}_i \in S_i$ , the conditional distribution of  $\theta_i$  has finite first and second moments, and matrix  $\text{Var}(\theta_i | s_i = \tilde{s}_i)$  has full rank. The latter assumption is also without loss of generality, as redundant signals can always be replaced with independent noise.

Third, conditional on the finite signals of a subset of players, their infinite signals do not contain any additional information about other players' finite signals. Formally, take any subset  $T$  of players, with  $|T| \geq 1$  and  $|T| < n$ . Take any profile of finite and infinite signal realizations  $(\tilde{s}_T, \tilde{\theta}_T)$  of players in set  $T$ . Then the probability of any realization  $\tilde{s}_{-T}$  of finite signals of players outside of  $T$ , conditional on the entire profile of signals  $(\tilde{s}_T, \tilde{\theta}_T)$ , is equal to the probability of realization  $\tilde{s}_{-T}$  conditional only on the profile of finite signal realizations  $\tilde{s}_T$ . One immediate case of distributions satisfying this condition is vector  $\theta$  being independent of vector  $s$ . However, more interesting dependencies are also allowed. For example,  $n - 1$  players may observe independent identically distributed signals  $\theta_i$ , and the  $n$ -th player may observe a subset of those signals, with his signal  $s_n$  determining which subset of the  $n - 1$  independent signals he observes.<sup>7</sup> Another example is two

<sup>7</sup>Formally, for the first  $n - 1$  players,  $|S_i| = 1$ , and  $\theta_i \in \mathbb{R}^k$ . For the  $n$ -th player,  $S_n = \{0, 1\}^{n-1}$  and  $\theta_n \in \mathbb{R}^{(n-1)k}$ . When the  $j$ -th element of  $s_n$  is equal to 1, components  $(k(j - 1) + 1, \dots, kj)$  of  $\theta_n$  are equal to  $\theta_j$ . When the  $j$ -th

players observing signals  $\theta_i$  drawn from two potentially correlated distributions (with several possible correlation coefficients), and also observing partially informative signals  $s_i$  about the correlation coefficient.

Next, for convenience, we assume that for every player  $i$ , for every realization  $\tilde{s}_i$  of his finite signal, the expected value of his infinite signal is zero:  $E[\theta_i | s_i = \tilde{s}_i] = 0$ . This is without loss of generality. Note, however, that in conjunction with the previous assumption, this renormalization implies a stronger property: for every subset  $T$  of players, with  $|T| \geq 0$  and  $|T| \leq n$ , for any profile of their finite signal realizations  $\tilde{s}_T$  that has a positive probability, for any player  $i$  (who may or may not be a part of  $T$ ), we have  $E[\theta_i | s_T = \tilde{s}_T] = 0$ .<sup>8</sup>

Finally, we assume that for every player  $i$  and every player  $j \neq i$ , the conditional expectation  $E[\theta_j | s_i = \tilde{s}_i, s_j = \tilde{s}_j, \theta_i = \tilde{\theta}_i]$  is a linear function of  $\tilde{\theta}_i$ . This is a substantive assumption, which we discuss in more detail in Section 4.5. Note that while the expectation has to be a linear function of  $\tilde{\theta}_i$ , the coefficients in this linear function are allowed to depend on  $\tilde{s}_i$  and  $\tilde{s}_j$  (although of course the case when they do not depend on one or both of these finite signals is also allowed). In particular, these assumptions allow for the case of jointly normal signals  $\theta$  whose covariance matrix may depend on the profile of finite signals  $s$ .

## 4.2 Actions and Payoffs

After privately observing his signal, each player  $i$  chooses action  $a_i \in \mathbb{R}^{m_i}$ , where  $m_i \geq 1$  is the dimension of player  $i$ 's action. Note that there is in general no relation between the size  $k_i$  of player  $i$ 's finite signal space  $S_i$ , the dimension  $\ell_i$  of his infinite signal  $\theta_i$ , and the dimension  $m_i$  of his action  $a_i$ . The payoff of player  $i$  depends on his own action  $a_i$ , the actions of other players (denoted by vector  $a_{-i}$ ), and the entire vectors of signals  $s$  and  $\theta$  (which include both the signals of player  $i$  and the signals of other players):

$$\pi_i(a_i, a_{-i}, s, \theta) = \frac{1}{2} a_i^T \Gamma_{ii}(s) a_i + a_i^T \Gamma_{i,-i}(s) a_{-i} + a_i^T g_i(s, \theta) + h_i(a_{-i}, s, \theta), \quad (4)$$

where  $\Gamma_{ii}(s)$  is a matrix of size  $m_i \times m_i$ ,  $\Gamma_{i,-i}(s)$  is a matrix of size  $m_i \times (\sum_{j \neq i} m_j)$ ,  $g_i(s, \theta)$  is a vector of size  $m_i$ , and  $h_i$  is an arbitrary function of  $a_{-i}$ ,  $s$ , and  $\theta$ .<sup>9</sup> As before, the presence of function  $h_i$  does not affect the incentives of player  $i$  (and thus the equilibria of the game), but does in general affect the efficiency and welfare properties of various strategy profiles.

We assume that for each  $i$  and each signal realization  $\tilde{s}_i \in S_i$ , the conditional expectation  $E[\Gamma_{ii}(s) | s_i = \tilde{s}_i]$  is a symmetric negative definite matrix. In addition, we assume that for every player  $i$ , the expectation  $E[g_i(s, \theta) | s_i = \tilde{s}_i, \theta_i = \tilde{\theta}_i]$  is a linear function of  $\tilde{\theta}_i$ . Note that while the expectation has to be a linear function of  $\tilde{\theta}_i$ , the coefficients in this linear function are allowed to depend on  $\tilde{s}_i$ . We impose no other restrictions on payoff functions  $\pi_i$ .

---

element of  $s_n$  is equal to 0, these components are random noise.

<sup>8</sup>The proof of this statement is given in Appendix A.6.

<sup>9</sup>The only assumption we need to impose on function  $h_i$  is that the expected value of  $h_i(a_{-i}, s, \theta)$  is finite for every profile of linear strategies of players other than  $i$ , conditional on every possible realization  $(\tilde{s}_i, \tilde{\theta}_i)$  of player  $i$ 's signals.

### 4.3 Linear Equilibrium

We say that a strategy of player  $i$  is linear if it can be represented as  $a_i(s_i, \theta_i) = \kappa_i(s_i) + \Lambda_i(s_i)\theta_i$ , where for every realization  $\tilde{s}_i$  of signal  $s_i$ ,  $\kappa_i(\tilde{s}_i)$  is a vector in  $\mathbb{R}^{m_i}$  and  $\Lambda_i(\tilde{s}_i)$  is a matrix in  $\mathbb{R}^{m_i \times \ell_i}$ . Note, in particular, that the sensitivity of player  $i$ 's action to his infinite signal  $\theta_i$  may depend on the realization of his finite signal  $s_i$ .

We say that a profile of strategies of all players is a *linear equilibrium* if (1) every strategy  $a_i(\cdot)$  in the profile is linear, and (2) the profile of strategies is a Bayesian Nash equilibrium, i.e., for every player  $i$ , for every realization  $(\tilde{s}_i, \tilde{\theta}_i)$  of signals  $(s_i, \theta_i)$ , action  $a_i(\tilde{s}_i, \tilde{\theta}_i)$  maximizes player  $i$ 's expected payoff, given the primitives of the game and the strategies of other players.

### 4.4 Equilibrium Existence and Uniqueness

We can now state and prove the second main result of the paper.

**Theorem 2** *Generically, there exists a unique linear equilibrium.*

The remainder of this section presents the proof of Theorem 2 and the closed-form characterization of the unique linear equilibrium. As in the proof of Theorem 1, the argument consists of two parts. Lemma 3 provides conditions that guarantee the existence and uniqueness of linear equilibrium. It also provides closed-form expressions for the equilibrium strategies when these conditions are satisfied. Lemma 4 shows that the conditions for existence and uniqueness required in Lemma 3 hold generically.

The first condition in Lemma 3 is the same as Condition 1 stated earlier (in Section 2). To give the second condition, we need to introduce some additional notation. By assumption, the expectation  $E[\theta_j | s_i = \tilde{s}_i, s_j = \tilde{s}_j, \theta_i = \tilde{\theta}_i]$  is a linear function of  $\tilde{\theta}_i$ . Denote this linear function by  $E[\theta_j | s_i = \tilde{s}_i, s_j = \tilde{s}_j, \theta_i = \tilde{\theta}_i] = Q_{ij}^T(\tilde{s}_i, \tilde{s}_j)\tilde{\theta}_i$ .<sup>10</sup>

Let  $\Psi$  be a block matrix defined as follows. As in Section 2, enumerate the  $k_i$  possible signal realizations  $\tilde{s}_i$  of player  $i$  as  $1, 2, \dots, k_i$ , and let  $K = k_1 + \dots + k_n$ . Matrix  $\Psi$  consists of  $K \times K$  blocks: each block row corresponds to one player  $i$  and one possible signal realization of that player,  $\tilde{s}_i$ , and analogously, each block column corresponds to one player  $j$  and one possible signal realization of that player,  $\tilde{s}_j$ . The block of matrix  $\Psi$  in block row  $(i, \tilde{s}_i)$  and block column  $(j, \tilde{s}_j)$  is given by

$$\Psi_{(i, \tilde{s}_i), (j, \tilde{s}_j)} = \left( Q_{ij}(\tilde{s}_i, \tilde{s}_j) \otimes E[\Gamma_{ij}(s) | s_i = \tilde{s}_i, s_j = \tilde{s}_j] \right) \cdot P[s_j = \tilde{s}_j | s_i = \tilde{s}_i]$$

where “ $\otimes$ ” denotes the Kronecker product of two matrices.<sup>11</sup> When actions  $a$  or infinite signals  $\theta$  of all players are single-dimensional, the Kronecker product reduces to regular multiplication by a scalar. When both actions  $a$  and infinite signals  $\theta$  of all players are single-dimensional, the

<sup>10</sup>Note that this linear function does not have a constant term. As we show in Appendix A.6, this lack of the constant term is implied by our normalization of the signal distributions.

<sup>11</sup>If matrix  $A$  is of size  $k \times m$  with elements  $a_{ij}$  and matrix  $B$  is another matrix, of any size, then  $A \otimes B$  is defined as a block matrix with  $k \times m$  blocks, in which the size of each block is equal to the size of  $B$  and each block  $(i, j)$  is equal to  $a_{ij}B$ . See [http://en.wikipedia.org/wiki/Kronecker\\_product](http://en.wikipedia.org/wiki/Kronecker_product) for details.

expressions simplify further and each block  $\Psi_{(i,\tilde{s}_i),(j,\tilde{s}_j)}$  is just a real number. In the general case, the size of block  $\Psi_{(i,\tilde{s}_i),(j,\tilde{s}_j)}$  is  $m_i \ell_i \times m_j \ell_j$ .

**Condition 2** *Matrix  $\Psi$  is invertible.*

**Lemma 3** *The game has a unique linear equilibrium if and only if Conditions 1 and 2 are satisfied.*

The proof of Lemma 3 is in Appendix A.7. In addition to showing the result, the proof also shows that when Condition 1 or Condition 2 are not satisfied, then either the game has no linear equilibrium, or it has infinitely many. Also, the proof provides a closed-form solution for the equilibrium when both conditions are satisfied. Specifically, recall our representation of linear strategies as  $a_i(s_i, \theta_i) = \kappa_i(s_i) + \Lambda_i(s_i)\theta_i$ . Slightly abusing notation, let  $\kappa_i \in \mathbb{R}^{m_i k_i}$  be the stacked vector of player  $i$ 's constant terms  $\kappa_i(s_i)$ , one for each possible realization of signal  $s_i$  (i.e., the first  $m_i$  elements of  $\kappa_i$  are  $\kappa_i(1)$ , the next  $m_i$  elements are  $\kappa_i(2)$ , and so on). Let vector  $\kappa = (\kappa_1; \dots; \kappa_n)$  denote the combined profile of all players' constant terms. Next, consider the vectorization  $\text{vec } \Lambda_i(\tilde{s}_i)$ , i.e., the column vector of size  $m_i \ell_i$  in which the columns of matrix  $\Lambda_i(\tilde{s}_i)$  are stacked on top of each other.<sup>12</sup> Let  $\Lambda_i \in \mathbb{R}^{m_i \ell_i k_i}$  be the stacked vector of these vectorizations  $\text{vec } \Lambda_i(\tilde{s}_i)$ , one for each realization  $\tilde{s}_i$ . Finally, let vector  $\Lambda = (\Lambda_1; \dots; \Lambda_n)$  combine the vectors  $\Lambda_i$  of individual players.

Next, by assumption, the expectation  $E[g_i(s, \theta) | s_i = \tilde{s}_i, \theta_i = \tilde{\theta}_i]$  is a linear function of  $\tilde{\theta}_i$ . Denote this function by  $E[g_i(s, \theta) | s_i = \tilde{s}_i, \theta_i = \tilde{\theta}_i] = G_i(\tilde{s}_i) + F_i(\tilde{s}_i)\tilde{\theta}_i$ . Let  $g_i \in \mathbb{R}^{m_i k_i}$  be the stacked vector of individual vectors  $G_i(\tilde{s}_i)$ , one for each possible realization of  $s_i$  (i.e.,  $g_i = (G_i(1); \dots; G_i(k_i))$ ), and let  $g = (g_1; \dots; g_n)$  denote the vector combining these terms for all players. Finally, let  $f_i = (\text{vec } F_i(1); \dots; \text{vec } F_i(k_i)) \in \mathbb{R}^{m_i \ell_i k_i}$ , and let  $f = (f_1; f_2; \dots; f_n)$ .

When Conditions 1 and 2 are satisfied, the unique linear equilibrium is given by

$$\kappa = -\Phi^{-1}g \tag{5}$$

$$\Lambda = -\Psi^{-1}f. \tag{6}$$

The second step of the proof of Theorem 2 shows that Conditions 1 and 2 hold generically. Formally, consider the following one-dimensional collection of quadratic games, parameterized by  $\gamma \in \mathbb{R}$ . The information structure is the same for all the games in the collection, and payoffs are given by

$$\pi_i(a_i, a_{-i}, s, \theta) = \frac{1}{2}a_i^T \Gamma_{ii}(s)a_i + a_i^T (\gamma \Gamma_{i,-i}(s)) a_{-i} + a_i^T g_i(s, \theta) + h_i(a_{-i}, s, \theta) \tag{7}$$

**Lemma 4** *Conditions 1 and 2 hold for all  $\gamma \in \mathbb{R}$ , except for at most finitely many values.*

The proof of Lemma 4 is in Appendix A.8. Together with Lemma 3, this concludes the proof of Theorem 2.

<sup>12</sup>See [https://en.wikipedia.org/wiki/Vectorization\\_\(mathematics\)](https://en.wikipedia.org/wiki/Vectorization_(mathematics)) for details.

## 4.5 Linear Conditional Expectations

As we mentioned in Section 4.1, the assumption of the linearity of conditional expectations (i.e., the assumption that for every player  $i$  and every player  $j \neq i$ , the conditional expectation  $E[\theta_j | s_i = \tilde{s}_i, s_j = \tilde{s}_j, \theta_i = \tilde{\theta}_i]$  is a linear function of  $\tilde{\theta}_i$ ) is substantive. However, there are many interesting and common examples of joint distributions that satisfy this restriction.

The most common example is the multivariate normal distribution. For example, if vector  $\theta$  is distributed normally, with any variance-covariance matrix, and is independent of vector  $s$ , then  $E[\theta_j | s_i = \tilde{s}_i, s_j = \tilde{s}_j, \theta_i = \tilde{\theta}_i]$  is a linear function of  $\tilde{\theta}_i$ , for any  $i$  and  $j$ . More complex examples “built from” normal distributions are also possible. For example, finite signals  $s$  may determine correlations between the infinite signals of different players.

For a very different example of joint distributions satisfying the conditional linearity assumption, suppose random variables  $\alpha$  and  $\beta$  are independent (and other than that, come from arbitrary distributions), and suppose player  $i$  observes both of them (i.e.,  $\theta_i = (\alpha; \beta)$ ), while player  $j \neq i$  observes only one (e.g.,  $\theta_j = (\alpha)$ ). Then both  $E[\theta_j | \theta_i]$  and  $E[\theta_i | \theta_j]$  are linear functions. More generally, if there are several independent random variables, and each agent observes a subset of them, the resulting information structure satisfies the linearity condition. For example, suppose there are no finite signals  $s$  (formally, a fixed vector  $s$  is observed with probability 1), and consider a network of agents in which the infinite signal (type) of each agent is a random variable drawn from some distribution (possibly a different one for each agent). The types of agents are independent of each other. Each agent observes their own type, and the types of their neighbors in the network. Then, for all  $i$  and  $j$ ,  $E[\theta_j | \theta_i = \tilde{\theta}_i]$  is a linear function of  $\tilde{\theta}_i$ .

For yet another example, suppose random variables  $\alpha$  and  $\beta$  are independent and identically distributed, and player  $i$  observes each of them separately, while player  $j$  only observes their sum. Again, both  $E[\theta_j | \theta_i]$  and  $E[\theta_i | \theta_j]$  are linear functions. As before, we can use this observation to explore richer models in which some players observe specific variables while others only observe their aggregates (and in which who observes what may depend on finite signals  $s$ ).

Finally, note that if a player’s signal  $\theta_i$  is binary, then for *any* joint distribution of  $\theta_i$  and  $\theta_j$ , the expectation  $E[\theta_j | \theta_i]$  is a linear function of  $\theta_i$ . We return to this observation in Section 6.2, where it serves as a basis for a connection between two seemingly distinct models.

The list above is by no means exhaustive—there are many other types of joint distributions of signals for which our assumption of linear conditional expectations is satisfied.<sup>13</sup> So while the assumption is certainly substantive, it nevertheless allows for a wide variety of interesting cases.

## 5 Application: Beauty Contest Games

In this section, we illustrate the framework and results of Section 4 with an application: beauty contest games with potentially uncertain relative weights that players put on coordinating with others. In beauty contest games (Morris and Shin, 2002), players receive signals about an uncertain

---

<sup>13</sup>See, e.g., Li et al. (1987).



variable, and their optimal action is a weighted average of their estimate of the value of the variable and their estimate of the average action of other players. The standard BC framework assumes that all players are ex ante identical, and in particular, receive identically distributed signals and put the same weights on matching the actions of all other players. In their analyses of beauty contest games, Bergemann et al. (2017) and Lambert et al. (2017b) allow for rich asymmetric information structures, but maintain the assumption that each agent cares equally about matching all other agents’ actions (i.e., each player tries to match a weighted average of the true state of the world and the average action of other players).

Golub and Morris (2017), Leister (2016), and Myatt and Wallace (2017) relax this assumption, and allow players to put more weight on matching some agents’ actions and less (or none) on matching other agents’ actions, examining the interplay between beauty contests and network settings. In such environments, Golub and Morris (2017) consider common and heterogeneous priors, and focus on the characterization of higher-order expectations and their limit as the weight of the coordination component in payoff functions goes to 1. Leister (2016) and Myatt and Wallace (2017) explore endogenous information acquisition.<sup>14</sup>

The contribution of the current section of our paper is to not only allow players to put different weights on matching different agents’ actions, but to also allow these weights to be stochastic, with players potentially having rich information both about these weights and about the “states of the world” (and about other players’ knowledge of those parameters). We also allow players to target different state variables, allowing for many other interpretations of the framework (e.g., the framework can be viewed as a model of a coordination game, in which each player has some (individual) “bliss” action that he would pick in the absence of interactions with other players, but also puts some weight on picking an action close to the actions of some other players), and to receive multidimensional information (so that, for example, one dimension of the signal is informative about a player’s own optimal bliss action and another dimension is informative about another player’s bliss action; or one dimension is a player’s own signal, another dimension is a signal shared with some other players, and the third, finite dimension, contains information about other players’ payoff functions).

In Section 5.1, we describe our general model of beauty contest games. In Section 5.2, we show that in this model, there is always a unique linear equilibrium (unlike Theorem 2, this statement is always true, not just generically). In Section 5.3, we consider a particular example of a beauty contest game with uncertain interactions and show that allowing for uncertain interaction terms in players’ payoff functions leads to economically novel predictions.

---

<sup>14</sup>With the exception of Golub and Morris (2017), all papers on beauty contests mentioned above consider games with jointly normal information structures. Golub and Morris (2017) consider beauty contest games with finite information structures, which (in the case of common priors) fit the model of Section 6.1 below. As we explain in Section 6.2, all these models are special cases of a unified model of quadratic games with fixed interaction coefficients and information structures that satisfy the “linear conditional expectations” property.

## 5.1 Model of Beauty Contest Games

The model is a special case of the setting of Section 4. There are  $n$  players,  $i = 1, \dots, n$ . Each player observes a finite signal  $s_i \in S_i$  and an infinite signal  $\theta_i \in \mathbb{R}^{\ell_i}$ . Each player also has a bliss point  $b_i^* \in \mathbb{R}$ , which is a random variable. Denote by  $b = (b_1^*; \dots; b_n^*)$  the random vector of bliss points. As in Section 4, we do not impose any restrictions on the distribution of finite signals  $s$ , apart from assuming that for each player  $i$  and each signal realization  $\tilde{s}_i \in S_i$ , the probability of player  $i$  observing  $\tilde{s}_i$  is positive. For the profile of infinite signals  $\theta$  and the vector of bliss points  $b$ , we assume that they are jointly normally distributed, with the mean of  $\theta$  being equal to zero, the mean of  $b$  being equal to  $\bar{b}$ , and the variance-covariance matrix of  $(\theta; b)$  being arbitrary, subject only to the constraint that for each  $i$ , matrix  $Var(\theta_i)$  has full rank (as before, this assumption is without loss of generality). Note that different bliss points  $b_i^*$  can be assumed to be identical, but are not required to. We assume that  $s$  is independent of  $\theta$  and  $b$ .<sup>15</sup>

Each player's action is a real number:  $a_i \in \mathbb{R}$ . The payoff of each player is a function of his bliss point, his action, the actions of other players, and the profile of finite signals  $s$ , and is equal to

$$\pi_i(a_i, a_{-i}, b_i^*, s) = -\gamma_{ii}(s)(a_i - b_i^*)^2 - \sum_{j \neq i} \gamma_{ij}(s)(a_i - a_j)^2, \quad (8)$$

where for every  $s$ ,  $\gamma_{ii}(s) > 0$  and for each  $j \neq i$ ,  $\gamma_{ij}(s) \geq 0$ .<sup>16</sup> In the classical models of beauty contests, players have an incentive to choose an action that is close to a common random bliss point, and at the same time that is close to the average action of the other players. The weights that players put on each component can vary across players, but are fixed and commonly known. Our model of beauty contests allows for more flexibility along three dimensions. First, different players may have different bliss points. Second, a player may put different weights on matching the actions of different other players. Finally, the weights that players put on matching the bliss points and on matching various other players' actions may be stochastic, with complex information structures about these weights (e.g., on a network, a player may know the preferences of her neighbors, but only have probabilistic information about the preferences of agents who are further away).

## 5.2 Equilibrium Existence and Uniqueness

A linear equilibrium in our model is defined as in Section 4.3. The following proposition shows that there always exists a unique equilibrium. Note that this statement is stronger than the corresponding statement in Theorem 2, since it does not rely on genericity.

---

<sup>15</sup>The results of this section can be extended to more general cases in which the distribution of  $\theta$  is not necessarily normal, and in which  $s$  and  $\theta$  are not independent, while still remaining within the framework of Section 4. We restrict attention to the current set of assumptions for expositional convenience.

<sup>16</sup>As in the Cournot competition model of Section 3, our beauty contests model does not formally fit the framework of Section 4, because the bliss points are not directly parts of either finite signals  $s$  or infinite signals  $\theta$ . Similarly to the Cournot competition model, to fit formally the framework of Section 4.1, one can introduce a dummy player Nature whose finite signal space consists of only one element and whose infinite signal is equal to vector  $b$  and whose payoff  $\pi_N$  is given by  $\pi_N = -a_N^2$ , so that it always plays  $a_N = 0$ .

**Proposition 4** *In a model of beauty contest games, there always exists a unique linear equilibrium.*

The proof of Proposition 4 is in the appendix. The idea of the proof is to first observe that in a given game, Conditions 1 and 2 of Lemma 3 are satisfied if and only if they are satisfied in a modified game where all the entries  $b_i^*$  are replaced by zeros. The next step of the proof is to show that the modified game has a unique equilibrium (the one in which all players always play zero). This implies that Conditions 1 and 2 are satisfied for the modified game—which in turn implies that they are satisfied for the original game, and thus the original game has a unique linear equilibrium.

### 5.3 Example of a Beauty Contest Game with Stochastic Payoff Interactions

To illustrate the framework, we present a simple example in which players are uncertain about the relative weights other players put on the bliss point vs. matching the actions of other players. We find that the impact of this uncertainty on the equilibrium of the game is subtle, and is substantively different from an equilibrium that would arise if these uncertain terms were simply replaced by their expected values. Our framework also allows us to provide predictions on the signs of various comparative static effects as this uncertainty increases.

**Example 1** *There are  $n \geq 2$  players who receive signals  $\theta_i = \theta + \epsilon_i$ , where  $\theta \sim \mathcal{N}(0, \sigma^2)$  is the players' common "bliss point" and errors  $\epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2)$  are iid. We denote by  $\rho = \frac{\sigma^2}{\sigma^2 + \sigma_\epsilon^2}$  the informativeness of signal  $\theta_i$  about the state variable  $\theta$ . Each player  $i$  also observes his "payoff type"  $s_i \in \{H, L\}$ , with prior  $P(s_i = H) = P(s_i = L) = 1/2$ . Payoff types are correlated across players:  $P[s_j = x | s_i = x] = q \geq \frac{1}{2}$ , for  $x \in \{H, L\}$ , for all  $j \neq i$ . At the same time, the profile of players' payoff types  $s$  is independent of state variable  $\theta$  and signals  $\theta_i$ .*

After observing his signal  $\theta_i$  and payoff type  $s_i$ , each player  $i$  chooses action  $a_i \in \mathbb{R}$  to minimize a combination of two quadratic loss terms: one from the distance to the bliss point  $\theta$  and one from the distance to the average action of other players,  $\bar{a}_{-i}$ . The relative weights player  $i$  puts on each of the two terms are determined by his payoff type: when player  $i$  has type  $s_i = H$ , the weight he places on the coordination component of his payoff is  $\alpha_i = \alpha_H$ , and when his type is  $s_i = L$ , the weight is  $\alpha_L$ . We assume that  $0 \leq \alpha_L \leq \alpha_H < 1$ .<sup>17</sup>

Formally, player  $i$ 's payoff depends on his action  $a_i$ , the average action of other players  $\bar{a}_{-i}$ , his payoff type  $\alpha_i$ , and the state  $\theta$ :

$$\pi_i(a_i, \bar{a}_{-i}, \alpha_i, \theta) = -(1 - \alpha_i)(a_i - \theta)^2 - \alpha_i(a_i - \bar{a}_{-i})^2. \quad (9)$$

By Proposition 4, there exists a unique linear equilibrium of this game. In this equilibrium, the sensitivity of the player's action to his signal  $\theta_i$  depends on his payoff type  $s_i$ . Since the game is symmetric and the linear equilibrium is unique, the equilibrium is also symmetric. Equilibrium

---

<sup>17</sup>The case  $\alpha_H = 1$  would violate the assumption in the model of Section 5.1 that  $\gamma_{ii}(s)$  is always positive, and so we rule it out. Note, however, the existence and uniqueness result, and the explicit formulas below would continue to hold even in the case  $\alpha_H = 1$ , as long as  $q < 1$  and  $\alpha_L < 1$ , allowing for the possibility that some agents care only about matching the average action of other players and put zero weight on matching the fundamental value  $\theta$ .

strategies are described by two numbers: the sensitivity  $\beta_H$  when payoff type is  $H$ , and the sensitivity  $\beta_L$  when payoff type is  $L$  (the affine term is 0 for both payoff types). I.e., when a player of type  $H$  observes signal  $\theta_i$ , his action is  $a_i = \beta_H \theta_i$ , and when a player of type  $L$  observes signal  $\theta_j$ , his action is  $a_j = \beta_L \theta_j$ . Using our closed-form characterization, we find

$$\beta_H = \rho \frac{(1 - \alpha_H)(1 - \alpha_L q \rho) + \alpha_H(1 - \alpha_L)(1 - q)\rho}{(1 - \alpha_H q \rho)(1 - \alpha_L q \rho) - \alpha_H \alpha_L (1 - q)^2 \rho^2},$$

$$\beta_L = \rho \frac{(1 - \alpha_L)(1 - \alpha_H q \rho) + \alpha_L(1 - \alpha_H)(1 - q)\rho}{(1 - \alpha_H q \rho)(1 - \alpha_L q \rho) - \alpha_H \alpha_L (1 - q)^2 \rho^2}.$$

For the discussion of the properties of the equilibrium, it is convenient to introduce two new variables. Let  $\alpha = \frac{\alpha_H + \alpha_L}{2}$  and  $\Delta = \frac{\alpha_H - \alpha_L}{2}$ , so that  $\alpha$  represents the average weight that each player puts on the square of the difference between his action and the average action of other players (recall that types  $H$  and  $L$  are equally likely) and  $\Delta$  represents the absolute amount by which the actual weights  $\alpha_H$  and  $\alpha_L$  differ from that average. Note first that our model naturally nests the standard beauty contest case without stochastic interactions: if  $\alpha_H = \alpha_L = \alpha$ , the equilibrium reduces to  $\beta_H = \beta_L = \rho \frac{1 - \alpha}{1 - \rho \alpha}$ , and if  $q = 1$ , the game “decomposes” into either all players being of type  $H$  (and knowing that) or all players being of type  $L$ , with the corresponding equilibrium strategies reducing to  $\beta_H = \rho \frac{1 - \alpha_H}{1 - \rho \alpha_H}$  and  $\beta_L = \rho \frac{1 - \alpha_L}{1 - \rho \alpha_L}$ .<sup>18</sup>

Another “corner case” is when the type of player  $i$  is not informative about the types of other players, i.e.,  $q = \frac{1}{2}$ . In that case, we get  $\beta_H = \rho \frac{1 - \alpha - (1 - \rho)\Delta}{1 - \rho \alpha}$  and  $\beta_L = \rho \frac{1 - \alpha + (1 - \rho)\Delta}{1 - \rho \alpha}$ , and the average sensitivity of a player to his signal,  $\frac{\beta_H + \beta_L}{2}$ , is equal to  $\rho \frac{1 - \alpha}{1 - \rho \alpha}$ . I.e., the average sensitivity of a player to his signal is the same as in the standard beauty contest game with constant interaction term  $\alpha$ , and each individual player behaves as if he is simply best responding to a population of such “average” players.

Away from these corner cases, the relationship between the parameters of the game and the equilibrium strategy profiles becomes more subtle. To illustrate this relationship, we look at the comparative statics of the equilibria, holding fixed the average weight  $\alpha$  and varying the difference between the weights in the two states ( $2\Delta$ ) as well as the informativeness of a player’s state about the states of other players ( $q$ ). Proposition 5 summarizes our results.

**Proposition 5** *The following comparative statics hold:*

1. *The average sensitivity  $\frac{\beta_H + \beta_L}{2}$  decreases in  $\Delta$  and decreases in  $q$ .*
2. *The sensitivity  $\beta_H$  decreases in  $\Delta$  and decreases in  $q$ .*
3. *The sensitivity  $\beta_L$  increases in  $q$ , but is not monotone in  $\Delta$ .*

<sup>18</sup>Note that the number of players  $n$  has no impact on the equilibrium behavior of a given player in the game, as long as  $n \geq 2$ . In particular, we could consider a version of the beauty contest game with a continuum of players, as is common in the literature. All of the formulas and results would remain the same. More generally, we could consider a model of Section 5.1 with continuous masses of players of different types (or even with continuous masses of “small” players of some types and discrete “large” players of other types). Aside from minor notational adjustments, none of our results would change.

Intuition for the non-monotonicity in part 3 is as follows: as  $\Delta$  increases, the weight  $\alpha_L$  (that a player of type  $L$  puts on matching the actions of other players) decreases, which in the absence of other effects would lead such a player to increase the weight he puts on his own signal. However, the weight that players of type  $H$  put on their signals decreases, which in equilibrium also gives an incentive to players of type  $L$  to decrease the sensitivity of their action to their signals. For certain parameter values (for example  $\alpha = \frac{7}{8}$ ,  $q = \frac{15}{16}$ , and  $\rho = \frac{95}{96}$ ), the latter effect outweighs the former.

## 6 Models with Constant Interaction Terms

The models of Sections 2 and 4 allow the interaction terms  $\Gamma_{ij}$  to vary stochastically as a function of the profile of players' finite signals  $s$ . While the results of those sections provide tractable closed-form solutions for those general models, the solutions are further simplified in applications in which interaction terms  $\Gamma_{ij}$  are constant. In this section, we present these simplified solutions in two special cases. First, we consider the case of constant interaction terms and finite signals. Second, we consider the case of constant interaction terms and potentially infinite signals.

### 6.1 Constant Interaction Terms and Finite Signals

Consider the setting of Section 2, and suppose that for all  $i$  and  $j$ , interaction terms  $\Gamma_{ii}(s)$  and  $\Gamma_{ij}(s)$  do not depend on the profile of signals  $s$ , and are simply always equal to some matrices  $\Gamma_{ii}$  and  $\Gamma_{ij}$  (where matrix  $\Gamma_{ii}$  is negative definite). The payoff term  $g_i(s)$  is still allowed to depend on signal  $s$  in an arbitrary way.

Blocks of matrix  $\Phi$  then simplify to  $\Phi_{(i,\tilde{s}_i),(j,\tilde{s}_j)} = \Gamma_{ij} \cdot P[s_j = \tilde{s}_j | s_i = \tilde{s}_i]$ . Furthermore, for every pair of players  $i$  and  $j$ , all blocks  $\Phi_{(i,\tilde{s}_i),(j,\tilde{s}_j)}$  (across all possible realizations of signals  $\tilde{s}_i$  and  $\tilde{s}_j$ ) can be joined together, and expressed more compactly as

$$\Phi_{ij} = \Gamma_{ij} \otimes M_{ij},$$

where  $\otimes$  denotes the Kronecker product and  $M_{ij}$  is a matrix of size  $k_i \times k_j$  given by

$$M_{ij} = \begin{pmatrix} P[s_j = 1 | \tilde{s}_i = 1] & \cdots & P[s_j = k_j | \tilde{s}_i = 1] \\ \vdots & \ddots & \vdots \\ P[s_j = 1 | \tilde{s}_i = k_i] & \cdots & P[s_j = k_j | \tilde{s}_i = k_i] \end{pmatrix}.$$

Lemma 1, of course, continues to hold: there is a unique Bayesian Nash equilibrium if and only if matrix  $\Phi$  is invertible, and when it is, equilibrium strategies are given by  $a = -\Phi^{-1}g$ . Lemma 2 also applies directly, showing that in the model with fixed interaction terms, generically, there exists a unique Bayesian Nash equilibrium.

## 6.2 Constant Interaction Terms and Potentially Infinite Signals

The second case we consider is the setting of Section 4, in which we assume that the players only receive potentially infinite signals  $\theta_i$ , and do not receive any finite signals  $s_i$  (a formal way to incorporate this into the setting of Section 4 is to simply assume that for each  $i$ , set  $S_i$  contains only one element, which player  $i$  observes with probability 1 and which therefore contains no information). This assumption leads to several simplifications. First, as in Section 6.1, interaction terms are constant, given by matrices  $\Gamma_{ii}$  and  $\Gamma_{ij}$ . Second, the joint distribution of signals  $\theta_i$  has to satisfy only the following restrictions: (i) for each  $i$ , matrix  $\text{Var}(\theta_i)$  has full rank; (ii) for each  $i$ ,  $E[\theta_i] = 0$ ; and (iii) for each  $i$  and  $j \neq i$ , the conditional expectation  $E[\theta_j | \theta_i = \tilde{\theta}_i]$  is a linear function of  $\tilde{\theta}_i$ . Finally, the payoff terms  $g_i(\theta)$  also have to satisfy a linearity property:  $E[g_i(\theta) | \theta_i = \tilde{\theta}_i]$  is a linear function of  $\tilde{\theta}_i$ . Note that only the last two assumptions, on the linearity of conditional expectations, are substantive; assumptions (i) and (ii) are made only for convenience, without loss of generality.

In this setting, the expression for the blocks of matrix  $\Phi$  simplify even further, and we can drop the realizations of finite signals  $s_i$  from the subindices:

$$\Phi_{ij} = \Gamma_{ij}.$$

Thus, matrix  $\Phi$  is simply equal to the matrix  $\Gamma$  that summarizes the interaction terms in players' payoff functions.

The expression for the blocks of matrix  $\Psi$  is also substantially simplified, and also no longer requires the realizations of finite signals  $s_i$  in the subindices:

$$\Psi_{ij} = Q_{ij} \otimes \Gamma_{ij}.$$

Of course, Lemma 3 continues to hold, and the game has a unique linear equilibrium if and only if matrices  $\Phi$  and  $\Psi$  are invertible. When the matrices are invertible, the equilibrium strategies continue to be given by  $\kappa = -\Phi^{-1}g$  and  $\Lambda = -\Psi^{-1}f$ . And as before, Lemma 4 applies directly, showing that in this model, generically, there exists a unique linear equilibrium.

The models of Sections 6.1 and 6.2 are seemingly quite different: the former allows arbitrary finite signals without imposing any restrictions on their joint distribution or on functions  $g_i(s)$ , while the latter allows infinite signals but imposes the restrictions of linearity on conditional expectations  $E[\theta_j | \theta_i = \tilde{\theta}_i]$  and  $E[g_i(\theta) | \theta_i = \tilde{\theta}_i]$ . The two models, however, are closely related: the model of Section 6.1 is a special case of the model of Section 6.2. To see that, recall the observation made in Section 4.5 that if a player's signal  $\theta_i$  is binary, then for any joint distribution of  $\theta_i$  and  $\theta_j$ , the expectation  $E[\theta_j | \theta_i = \tilde{\theta}_i]$  is a linear function of  $\tilde{\theta}_i$ . Similarly, if player  $i$ 's signal takes  $k_i$  possible values (as in Section 6.1), we can “project” it onto a  $k_i$ -dimensional space  $\mathbb{R}^{k_i}$ , as follows: when the original signal realization is 1, the projection takes the value  $(1, 0, \dots, 0)$ ; when the original signal realization is 2, the projection takes the value  $(0, 1, 0, \dots, 0)$ ; and so on—when the original signal realization is  $k_i$ , the projection takes the value  $(0, \dots, 0, 1)$ . With this representation, any function of the projected signal is linear. Thus, if we represent the signals of all players in this

fashion, all conditional expectations  $E[\theta_j|\theta_i = \tilde{\theta}_i]$  and  $E[g_i(\theta)|\theta_i = \tilde{\theta}_i]$  become linear, and thus the model of Section 6.1 can be represented as a special case of the model of Section 6.2. (Formally, this representation violates both assumptions (i) and (ii) of the model of Section 6.2. However, as we noted, these assumptions are without loss of generality, and it is straightforward to further adjust the representation so that it fully fits the model.)

Thus, the model of Section 6.2 can be viewed as a unifying framework for a “linear conditional expectations” family of quadratic games with fixed interaction coefficients, including those with finite information structures, Gaussian information structures, and so on.

## 7 Beyond Quadratic Games

We end with a brief discussion on applying the main framework beyond quadratic payoff functions. In Section 7.1, we discuss how our results can be applied to other, not necessarily quadratic, games and economies with linear best responses. In Section 7.2, we provide an informal discussion on how our framework and results can be used to characterize changes in equilibrium behavior after small perturbations in general games with smooth payoffs.

### 7.1 Games with Linear Best Responses

In some applications, it is useful to abstract away from the game itself, and instead write down directly the best responses of each player as a function of the strategies of other players. These best responses continue to characterize the behavior of the agents in an economic system, without the need to specify agents’ payoffs. When the best response functions are a linear function of the strategies of other players, the “game” is one of linear best responses. Linear best response games can be interpreted as an alternative formulation to quadratic games. Which formulation to use is a matter of preferences, and our results apply to both formulations.

To illustrate, consider the game described in Section 2.1. The best response of player  $i$  satisfies

$$-\Gamma_{ii}a_i = E[g_i(s)|s_i] + \sum_{j \neq i} E[\Gamma_{ij}(s)a_j(s_j)|s_i] \quad (10)$$

When Condition 1 holds, the Bayesian Nash equilibrium is characterized by the unique solution to (10).

Now let us consider the case in which the game is not specified. Instead, assume that when player  $i$  conjectures that player  $j$  takes action  $\hat{a}_j(s_j)$  upon receiving his signal  $s_j$ , player  $i$  strictly prefers to play action

$$a_i(s_i) = E[g_i(s)|s_i] + \sum_{j \neq i} E[\Gamma_{ij}(s)\hat{a}_j(s_j)|s_i]$$

upon receiving signal  $s_i$ . In the above best response equation,  $g_i$  is arbitrary and  $\Gamma_{ij}(s)$  is a matrix of size  $m_i \times m_j$  whose elements can depend arbitrarily on the joint vector of signals  $s$ . Except for

payoff structure which is left unspecified, we continue to use the notation and information structure of Section 2.

An *equilibrium* in this context is defined as a strategy profile  $(a_1(\cdot), \dots, a_n(\cdot))$  such that for every  $i$ ,

$$a_i(s_i) = \mathbb{E}[g_i(s)|s_i] + \sum_{j \neq i} \mathbb{E}[\Gamma_{ij}(s)a_j(s_j)|s_i].$$

The results of Section 2 continue to apply in this alternative model specification, replacing  $\Gamma_{ii}$  by the negative of the identity matrix of size  $m_i$ . Analogous arguments apply to the model of Section 4.

## 7.2 Approximate Equilibrium in General Games

Quadratic games with stochastic interaction terms can be used to understand changes in player behavior in general games with smooth payoff functions when small changes are made to those games' payoff functions or information structures.

Specifically, consider a general game of incomplete information, whose payoff functions are smooth but not necessarily quadratic in players' actions. The signal structure is finite and the actions are real-valued or multi-dimensional. Suppose the game has a known pure-strategy equilibrium. How does the equilibrium strategy profile change after a small variation of the information structure or payoff functions? Quadratic games with stochastic interaction terms provide a convenient approach to answering these questions. We give an informal illustration of this approach below.

For concreteness, we focus on changes to payoff functions; small changes to the information structure can be treated analogously. Consider the following example. There are  $n$  players. Each player  $i$  takes a one-dimensional action  $a_i$ , taking real values (the argument extends directly to multi-dimensional actions). Each player  $i$  privately observes a random signal  $s_i$ , which takes finitely many possible values, as in Section 2.

However, unlike the framework of Section 2, player  $i$  now gets utility

$$u_i(a_1, \dots, a_n, s_1, \dots, s_n; x),$$

where  $x$  is a real-valued parameter, and where  $u_i$  is not necessarily quadratic, but is sufficiently smooth in the players' actions and parameter  $x$ .

Suppose there exists a pure-strategy Bayesian Nash equilibrium of the game just described for the case  $x = 0$ . The equilibrium strategy profile is denoted  $a^*(s) = (a_1^*(s_1); \dots; a_n^*(s_n))$ .

In general, when  $x$  is nonzero, the strategy profile  $a^*(s)$  is no longer an equilibrium. We now explain how the framework and the results of Section 2 help us find an approximate equilibrium profile for small values of  $x$ .

First, we perform a second-order expansion of player  $i$ 's utility around the initial equilibrium for  $x = 0$ , by computing the utility of a player  $i$  when every player  $j$  plays action  $a_j^*(s_j) + \Delta a_j$ , for



small values of  $\Delta a_j$ :

$$\begin{aligned}
u_i(a^*(s) + \Delta a, s, x) &\approx u_i(a^*(s), s, 0) + x \frac{\partial u_i}{\partial x}(a^*(s), s, 0) \\
&\quad + \sum_j \Delta a_j \frac{\partial u_i}{\partial a_j}(a^*(s), s, 0) + \sum_{j,k} \Delta a_j \Delta a_k \frac{\partial^2 u_i}{\partial a_j \partial a_k}(a^*(s), s, 0) \\
&\quad + \sum_j x \Delta a_j \frac{\partial^2 u_i}{\partial a_j \partial x}(a^*(s), s, 0).
\end{aligned}$$

Observe that the payoff of player  $i$  as a function of the action increments  $\Delta a$  is linear quadratic, when approximating the payoffs to the second order.

Thus, these payoffs define a linear-quadratic game in the action increments. Using the notation of Section 2, this game is defined by

$$\begin{aligned}
\Gamma_{ii}(s) &= 2 \frac{\partial^2 u_i}{\partial a_i^2}(a^*(s), s, 0) \\
\Gamma_{ij}(s) &= 2 \frac{\partial^2 u_i}{\partial a_i \partial a_j}(a^*(s), s, 0) \\
g_i(s) &= x \frac{\partial^2 u_i}{\partial a_i \partial x}(a^*(s), s, 0) \\
h_i(\Delta a_{-i}, s) &= u_i(a^*(s), s, 0) + x \frac{\partial u_i}{\partial x}(a^*(s), s, 0) \\
&\quad + \sum_{j \neq i} \Delta a_j \frac{\partial u_i}{\partial a_j}(a^*(s), s, 0) + \sum_{j \neq i, k \neq i} \Delta a_j \Delta a_k \frac{\partial^2 u_i}{\partial a_j \partial a_k}(a^*(s), s, 0) \\
&\quad + \sum_{j \neq i} x \Delta a_j \frac{\partial^2 u_i}{\partial a_j \partial x}(a^*(s), s, 0).
\end{aligned}$$

Note that the second-order approximation of utility is necessary. The first-order approximation is not enough because the first-order terms vanish in expectation, meaning that

$$\mathbb{E} \left[ \frac{\partial u_i}{\partial a_i}(a^*(s), s, 0) \middle| s_i \right] = 0,$$

which is the first order condition for  $a^*(s)$  to be an equilibrium profile for the case  $x = 0$ .

The last equality implies that  $\mathbb{E}[g_i(s)|s_i]$  is of order  $x$ , which in turn implies that the equilibrium actions of this linear-quadratic game are of order  $x$  as well. Hence, the second-order approximation of utility is also sufficient to obtain a first-order approximation of the equilibrium strategy following a small perturbation of the original game.

Finally, we remark that in this application, stochastic interactions play a major part. A model of quadratic games with deterministic payoff interaction terms does not allow to derive the approximate equilibrium of the perturbed game, because the partial derivatives of the players' utility functions, which define the interaction terms  $\Gamma_{ij}$  above, generally depend on the vector of signals.

## Appendix A: Proofs

### A.1 Proof of Lemma 1

Consider the optimization problem of player  $i$  who has observed realization  $\tilde{s}_i$  of signal  $s_i$ . If he chooses action  $a_i$ , his expected payoff (ignoring the term  $h_i(a_{-i}, s)$ , which does not affect incentives) is equal to

$$E[\pi_i|a_i, \tilde{s}_i] = \frac{1}{2} a_i^T E[\Gamma_{ii}(s)|\tilde{s}_i] a_i + a_i^T E[\Gamma_{i,-i}(s)|\tilde{s}_i] a_{-i} + a_i^T E[g_i(s)|\tilde{s}_i].$$

Since matrix  $E[\Gamma_{ii}(s)|\tilde{s}_i]$  is negative definite for all  $\tilde{s}_i$ , the unique best response of player  $i$  is to set  $a_i(\tilde{s}_i)$  to the value that satisfies the first-order condition

$$E[\Gamma_{ii}(s)|\tilde{s}_i] a_i(\tilde{s}_i) + \sum_{j \neq i} E[\Gamma_{ij}(s) a_j(s_j)|\tilde{s}_i] + E[g_i(s)|\tilde{s}_i] = 0, \quad (11)$$

and a necessary and sufficient condition for a profile of strategies  $a_i(\cdot)$  to be a Bayesian Nash equilibrium is that the first-order condition (11) is satisfied for every  $i$  and every signal realization  $\tilde{s}_i$ .

Bringing the first term under the summation sign, equation (11) becomes

$$\sum_{j=1}^n E[\Gamma_{ij}(s) a_j(s_j)|\tilde{s}_i] + E[g_i(s)|\tilde{s}_i] = 0.$$

By the law of iterated expectations, conditioning over all  $k_j$  possible realizations of signal  $\tilde{s}_j$ , the equation can be further rewritten as

$$\sum_{j=1}^n \sum_{\tilde{s}_j=1}^{k_j} E[\Gamma_{ij}(s)|\tilde{s}_i, \tilde{s}_j] P[s_j = \tilde{s}_j | s_i = \tilde{s}_i] a_j(\tilde{s}_j) + E[g_i(s)|\tilde{s}_i] = 0. \quad (12)$$

Using the notation introduced in Section 2.4, we can rewrite (12) as

$$\sum_{j, \tilde{s}_j} \Phi_{(i, \tilde{s}_i), (j, \tilde{s}_j)} a_j(\tilde{s}_j) + E[g_i(s)|\tilde{s}_i] = 0. \quad (13)$$

Stacking equations (13) for all  $i = 1, \dots, n$  and all  $\tilde{s}_i = 1, \dots, k_i$  we obtain

$$\Phi a + g = 0. \quad (14)$$

If  $\Phi$  is invertible, linear equation (14) admits the unique solution

$$a = -\Phi^{-1} g. \quad (15)$$

Conversely, if  $\Phi$  is not invertible, then (14) has either zero or infinitely many solutions, and hence either the equilibrium does not exist or there are infinitely many equilibria.

## A.2 Proof of Lemma 2

Slightly abusing notation, let  $\Phi(\gamma)$  denote the matrix  $\Phi$  corresponding to the quadratic game with parameter  $\gamma$ . Matrix  $\Phi(\gamma)$  is invertible if and only if its determinant  $\det \Phi(\gamma)$  is not equal to zero. Note that  $\det \Phi(\gamma)$  is a polynomial in  $\gamma$ , because each element of matrix  $\Phi(\gamma)$  is either independent of  $\gamma$  or is a linear function of it, and the determinant of any matrix is a polynomial function of its elements.

Consider the quadratic game with  $\gamma = 0$ , so that payoffs are  $\pi_i(a_i, a_{-i}, s) = \frac{1}{2}a_i^T \Gamma_{ii}(s)a_i + a_i^T g_i(s) + h_i(a_{-i}, s)$ . Since  $E[\Gamma_{ii}(s)|\tilde{s}_i]$  is negative definite for all  $\tilde{s}_i$ , there is a unique optimal action for each player  $i$ , and this optimal action is independent of the actions of other players. Therefore, this game has a unique Bayesian Nash equilibrium. Thus, by Lemma 1, Condition 1 holds, and so  $\det \Phi(\gamma) \neq 0$  at  $\gamma = 0$ .

Thus, the polynomial  $\det \Phi(\gamma)$  is not equal to the zero polynomial. Therefore, it has at most a finite number of roots, and so matrix  $\Phi(\gamma)$  is invertible for all  $\gamma \in \mathbb{R}$ , except for at most finitely many values.

## A.3 Proof of Proposition 1

To begin, let us consider the special case  $\alpha(\cdot) = c_i(\cdot) = 0$  for every  $i$ . We show that in this case, there exists a unique Bayesian Nash equilibrium. The profit of firm  $i$  becomes  $\pi_i = -\beta(\omega)q_i Q - d_i(\omega)q_i^2$ . For such a profit function, it is readily verified that all firms producing zero is an equilibrium: thus, there exists at least one equilibrium of this game. We also note that no matter what other firms produce, a firm can always guarantee itself a profit of zero by producing zero. Hence, in every equilibrium profile  $q_i(s_i)$ ,  $i = 1, \dots, n$ , the expected profit of each firm is nonnegative, and so is the sum of profits for all the players,

$$\sum_i E[-\beta(\omega)q_i(s_i)Q(s) - d_i(\omega)q_i(s_i)^2] = -E\left[\beta(\omega)Q(s)^2 + \sum d_i(\omega)q_i(s_i)^2\right],$$

where  $Q(s) = \sum_j q_j(s_j)$  is the total equilibrium quantity supplied by the firms. Note that the right hand side is also nonpositive, and hence has to be equal to zero. Therefore, in every equilibrium, the expected profit of every firm is equal to zero. Fixing the strategies of other firms, the best response of firm  $i$ , that maximizes its expected profit, conditional on its signal and conditional on the strategies of the other firms, is unique, because  $\beta(\cdot) > 0$  and  $d_i(\cdot) \geq 0$ . Hence, the only strategy for which the best response of firm  $i$  yields an expected profit of zero is to produce zero.

We have shown that there exists a unique Bayesian Nash equilibrium for the special case where  $\alpha(\cdot) = c_i(\cdot) = 0$  for every  $i$ . We now consider the general case in which  $\alpha(\cdot)$  and  $c_i(\cdot)$  are arbitrary. Note that, using the notation of Section 2, the term  $\Gamma_{ij}$  reduces to  $-\beta(\omega)$  if  $i \neq j$  and  $-2d_i(\omega) - 2\beta(\omega) < 0$  if  $i = j$ . Hence, the assumptions of Section 2.2 are satisfied, and Lemma 1 applies. Observe that the matrix  $\Phi$ , defined in Section 2.4, does not depend on the terms  $\alpha(\cdot)$  and  $c_i(\cdot)$  of the profit function. By Lemma 1, Condition 1 is satisfied for the special case  $\alpha(\cdot) = c_i(\cdot) = 0$ ,

and so by our observation Condition 1 is also satisfied for the general case. Therefore, by Lemma 1, there exists a unique Bayesian Nash equilibrium in this general case.

#### A.4 Proof of Proposition 2

To save on notation, we restrict attention to symmetric strategy profiles in which the firms of a given group use identical strategies. (By symmetry of the game structure, this restriction is without loss of generality.) A strategy profile then reduces to the specification of the strategies  $q_i(s_i)$  common to every firm of group  $i$ , for every group  $i$ .

Let us write the total surplus of a given strategy profile  $q_i(s_i)$  for every firm of group  $i$  in market  $k$ . Given market price  $P$  and market quantity  $Q$ , the ex-post consumer surplus is:

$$\frac{1}{2}(\alpha(\omega) - P)Q = \frac{1}{2k}\beta(\omega)Q^2,$$

and the ex-post producer surplus is

$$PQ - k \sum_i c_i(\omega)q_i(s_i) - k \sum_i d_i(\omega)q_i(s_i)^2 = \alpha(\omega)Q - \frac{1}{k}\beta(\omega)Q^2 - k \sum_i c_i(\omega)q_i(s_i) - k \sum_i d_i(\omega)q_i(s_i)^2,$$

so the ex-post total surplus is

$$\begin{aligned} \alpha(\omega)Q - \frac{1}{2k}\beta(\omega)Q^2 - k \sum_i c_i(\omega)q_i(s_i) - k \sum_i d_i(\omega)q_i(s_i)^2 \\ = k\alpha(\omega) \left( \sum_i q_i(s_i) \right) - \frac{k}{2}\beta(\omega) \left( \sum_i q_i(s_i) \right)^2 - k \sum_i c_i(\omega)q_i(s_i) - k \sum_i d_i(\omega)q_i(s_i)^2, \end{aligned}$$

and the ex-ante expected total surplus, denoted  $\text{TS}_k$ , is

$$\text{TS}_k = \text{E} \left[ k\alpha(\omega) \left( \sum_i q_i(s_i) \right) - \frac{k}{2}\beta(\omega) \left( \sum_i q_i(s_i) \right)^2 - k \sum_i c_i(\omega)q_i(s_i) - k \sum_i d_i(\omega)q_i(s_i)^2 \right].$$

In the remainder of the proof, it is useful to make explicit the dependence on the strategy profile. For a strategy profile summarized by  $\mathbf{q} = (q_1(\cdot), \dots, q_n(\cdot))$  in any market  $k$ , we denote by  $\text{TS}_k(\mathbf{q})$  the expected total surplus in that market. By assumption,  $\overline{\text{TS}}^a < \infty$ , so  $\sup_{\mathbf{q}} \text{TS}_k(\mathbf{q}) < \infty$ . The optimization problem that consists in finding the supremum is quadratic, which implies that the supremum is reached for some instance of strategy profile. Thus, there exists at least one ex-ante efficient strategy profile, which we denote  $\mathbf{q}^a$ . Since the total surplus is linear in the market size,  $\mathbf{q}^a$  does not depend on the market index  $k$ .

Next, let us consider the equilibrium in market  $k$ . Recall that, by Proposition 1, there exists a unique Bayesian Nash equilibrium profile  $(q_1(s_1), \dots, q_N(s_N))$ . The profit of any firm in group  $i$ ,

who produces  $q$  instead of its prescribed equilibrium quantity  $q_i(s_i)$ , is

$$\begin{aligned}
& Pq - c_i(\omega)q - d_i(\omega)q^2 \\
&= \left( \alpha(\omega) - \frac{\beta(\omega)}{k}Q \right) q - c_i(\omega)q - d_i(\omega)q^2 \\
&= \left( \alpha(\omega) - \frac{\beta(\omega)}{k}q - \frac{\beta(\omega)(k-1)}{k}q_i(s_i) - \beta(\omega) \sum_{j \neq i} q_j(s_j) \right) q - c_i(\omega)q - d_i(\omega)q^2,
\end{aligned}$$

where  $P$  is the market price and  $Q$  is the market quantity. The equilibrium profile is the unique solution to the first-order conditions

$$\mathbb{E} \left[ \alpha(\omega) - \frac{\beta(\omega)}{k}q_i(s_i) - \beta(\omega) \sum_{j=1}^N q_j(s_j) - c_i(\omega) - 2d_i(\omega)q_i(s_i) \middle| s_i \right] = 0 \quad (16)$$

for every  $i$ .

These first-order conditions are the same first-order conditions of the maximization of the expected total surplus with the addition of an extra term, a modified total surplus which we write  $\text{MTS}_k(\mathbf{q})$  and is expressed as  $\text{MTS}_k(\mathbf{q}) = \text{TS}_k(\mathbf{q}) - \text{R}_k(\mathbf{q})$  with

$$\text{R}_k(\mathbf{q}) = \mathbb{E} \left[ \frac{\beta(\omega)}{2} \sum_i q_i(s_i)^2 \right] \geq 0.$$

Denote by  $\mathbf{q}_k^*$  the unique equilibrium strategy profile in market  $k$ , and by  $\mathbf{q}^a$  a strategy profile that is ex-ante efficient. Since the total surplus is linear in  $k$ , ex-ante efficient strategy profiles do not depend on the market index  $k$ .

As  $\mathbf{q}_k^*$  maximizes the expected modified total surplus,  $\text{MTS}_k(\mathbf{q}_k^*) \geq \text{MTS}_k(\mathbf{q}^a)$ . Hence, we have

$$\text{TS}_k(\mathbf{q}_k^*) = \text{MTS}_k(\mathbf{q}_k^*) + \text{R}_k(\mathbf{q}_k^*) \geq \text{MTS}_k(\mathbf{q}_k^*) \geq \text{MTS}_k(\mathbf{q}^a) = \text{TS}_k(\mathbf{q}^a) - \text{R}_k(\mathbf{q}^a).$$

Besides,  $\mathbf{q}^a$  maximizes the expected total surplus, so  $\text{TS}_k(\mathbf{q}_k^*) \leq \text{TS}_k(\mathbf{q}^a)$ . Hence, letting  $\bar{\text{R}}_k(\mathbf{q}) = \text{R}_k(\mathbf{q})/(kN)$  and  $\bar{\text{TS}}_k(\mathbf{q}) = \text{TS}_k(\mathbf{q})/(kN)$ , noting that  $\bar{\text{TS}}_k(\mathbf{q}^a) = \bar{\text{TS}}^a$  and  $\bar{\text{TS}}_k(\mathbf{q}_k^*) = \bar{\text{TS}}_k^*$ , we have

$$\bar{\text{TS}}^a - \bar{\text{R}}_k(\mathbf{q}^a) \leq \bar{\text{TS}}_k^* \leq \bar{\text{TS}}^a.$$

For any fixed strategy profile  $\mathbf{q}$ ,  $\lim_{k \rightarrow \infty} \bar{\text{R}}_k(\mathbf{q}) = 0$ . Thus,

$$\lim_{k \rightarrow \infty} \bar{\text{TS}}_k^* = \bar{\text{TS}}^a,$$

which concludes the proof.

## A.5 Proof of Proposition 3

First, let us demonstrate that a best linear estimate exists. The problem of minimizing the  $\beta$ -weighted squared error can be formulated as a quadratic optimization problem, in which we minimize

$$\sum_{s_1, \dots, s_N, \omega} \beta(\omega)(q_{1,s_1} + \dots + q_{N,s_N} - \bar{Q}^p(\omega))^2 \mathbb{P}[s_1, \dots, s_N, \omega],$$

over the variables  $q_{1,1}, \dots, q_{1,k_1}, \dots, q_{N,1}, \dots, q_{N,k_N}$ . In this optimization problem, the objective is a quadratic function of  $k_1 + \dots + k_N$  variables. It is nonnegative, thus a minimum exists, and this minimum is reached in at least one instance of the variables. Hence, there exists at least one best linear estimate of  $\bar{Q}^p$ .

Next, we prove that the best linear estimate is unique. Suppose, by contradiction, that  $Q_1$  and  $Q_2$  are two different best linear estimates: there is at least one vector of signals for which  $Q_1$  and  $Q_2$  take different values. Observing that  $\mathcal{C}$  is convex, we claim that  $(Q_1 + Q_2)/2$  is a strictly better linear estimate of  $\bar{Q}^p$ . We have, for all realizations of joint signals,

$$\begin{aligned} \left( \frac{Q_1 + Q_2}{2} - \bar{Q}^p \right)^2 &= \frac{1}{4}(Q_1 - \bar{Q}^p)^2 + \frac{1}{4}(Q_2 - \bar{Q}^p)^2 + \frac{1}{2}(Q_1 - \bar{Q}^p)(Q_2 - \bar{Q}^p) \\ &\leq \frac{1}{4}(Q_1 - \bar{Q}^p)^2 + \frac{1}{4}(Q_2 - \bar{Q}^p)^2 + \frac{1}{4}(Q_1 - \bar{Q}^p)^2 + \frac{1}{4}(Q_2 - \bar{Q}^p)^2 \\ &= \frac{1}{2}(Q_1 - \bar{Q}^p)^2 + \frac{1}{2}(Q_2 - \bar{Q}^p)^2, \end{aligned}$$

and where the inequality is strict for the vectors of signals are such that  $Q_1 \neq Q_2$ . As all vectors of signals have positive probability, we get

$$\begin{aligned} \mathbb{E} \left[ \left( \frac{Q_1 + Q_2}{2} - \bar{Q}^p \right)^2 \right] &< \frac{1}{2} \mathbb{E} \left[ (Q_1 - \bar{Q}^p)^2 \right] + \frac{1}{2} \mathbb{E} \left[ (Q_2 - \bar{Q}^p)^2 \right] \\ &= \mathbb{E} \left[ (Q_1 - \bar{Q}^p)^2 \right] = \mathbb{E} \left[ (Q_2 - \bar{Q}^p)^2 \right] \end{aligned}$$

and hence  $(Q_1 + Q_2)/2$  is a strictly better linear estimate. Thus, we have established that there exists a unique best linear estimate  $Q^\dagger$ .

We now proceed to prove that  $\bar{Q}^\dagger = \bar{Q}^a$ . First, note that as  $\bar{Q}^\dagger$  is the unique minimizer of

$$\mathbb{E} [\beta(\omega)(Q - \bar{Q}^p)^2]$$

for  $Q \in \mathcal{C}$ , subtracting  $\beta(\omega)(\bar{Q}^p)^2$  from the objective and then multiplying by  $-N^2k/2$ , we get that  $\bar{Q}^\dagger$  is also the unique maximizer of

$$\mathbb{E} \left[ -\frac{1}{2}\beta(\omega)kN^2Q^2 + kN^2\beta(\omega)Q\bar{Q}^p \right] = \mathbb{E} \left[ -\frac{1}{2}\beta(\omega)kN^2Q^2 + kN^2Q\alpha(\omega) - kN^2Qc(\omega) \right] \quad (17)$$

for  $Q \in \mathcal{C}$ , where we used the fact that the ex-post efficient market quantity  $Q_k^p$  is equal to

$k(\alpha(\omega) - c(\omega))/\beta(\omega)$  in market  $k$ , and thus  $\bar{Q}^p = (\alpha(\omega) - c(\omega))/(N\beta(\omega))$ . Writing  $Q$  in the form  $Q = \frac{1}{N}(q_1(s_1) + \dots + q_N(s_N))$ , the value of (17) is equal to

$$\mathbb{E} \left[ k\alpha(\omega) \left( \sum_i q_i(s_i) \right) - \frac{k}{2}\beta(\omega) \left( \sum_i q_i(s_i) \right)^2 - k \sum_i c(\omega)q_i(s_i) \right],$$

which is the expected total surplus when each firm of group  $i$  produces  $q_i(s_i)$ , as shown in the proof of Proposition 2. The variable  $Q$  can then be interpreted as the per-firm average production. Hence, the value  $\bar{Q}^\dagger$  is the per-firm average production that makes the market ex-ante efficient, and  $\bar{Q}^\dagger = \bar{Q}^a$ .

Finally, we prove that the per-firm average equilibrium production converges to the ex-ante efficient level. Borrowing notation and terminology from the proof of Proposition 2, let us consider the problem of finding a profile  $\mathbf{q} = (q_1(s_1), \dots, q_N(s_N))$  that maximizes the value of

$$\mathbb{E} \left[ -\beta(\omega)(Q - \bar{Q}^p)^2 - \frac{\beta(\omega)}{kN} \sum_i q_i(s_i)^2 \right] \quad (18)$$

where we write  $Q = (q_1(s_1) + \dots + q_N(s_N))/N$ . Using that  $\bar{Q}^p = (\alpha(\omega) - c(\omega))/(N\beta(\omega))$ , the first-order conditions of this maximization problem are

$$\frac{2}{N} \mathbb{E} \left[ -\beta(\omega) \sum_i q_i(s_i) + \alpha(\omega) - c(\omega) - \frac{\beta(\omega)}{k} q_i(s_i) \middle| s_i \right] = 0 \quad (19)$$

for every  $i$ . Observe that these first-order conditions are identical to the first-order conditions associated with the unique equilibrium of market  $k$ . Indeed, recall from the proof of Proposition 2 that an equilibrium profile  $(q_1(s_1), \dots, q_N(s_N))$  in market  $k$  is the unique solution to (16) for every  $i$ , and equations (16) and (19) are identical for the cost structures of the environments being considered. Thus, there exists a unique profile that maximizes (18), and that profile is the equilibrium strategy profile of market  $k$ , denoted  $\mathbf{q}_k^*$ .

For a profile  $\mathbf{q}$ , and for  $Q \in \mathcal{C}$ , we let

$$\begin{aligned} A(Q) &= \mathbb{E} \left[ -\beta(\omega) (Q - \bar{Q}_k^p)^2 \right], \\ B_k(\mathbf{q}) &= \mathbb{E} \left[ \frac{\beta(\omega)}{kN} \sum_i q_i^2(s_i) \right], \\ F_k(\mathbf{q}) &= A \left( \frac{1}{N} \sum_i q_i(s_i) \right) - B_k(\mathbf{q}). \end{aligned}$$

Thus,  $\mathbf{q}_k^*$  maximizes  $F_k(\mathbf{q})$  over all profiles  $\mathbf{q}$ . Let  $\mathbf{q}^a = (q_1^a(s_1), \dots, q_N^a(s_N))$  be a profile such that  $\bar{Q}^a = (q_1^a(s_1) + \dots + q_N^a(s_N))/N$ . We have shown above that  $\bar{Q}^a = \bar{Q}^\dagger$ , and that  $\bar{Q}^a$  maximizes  $A(Q)$  over all profiles  $Q \in \mathcal{C}$ .

As  $\bar{Q}^a$  maximizes  $A(Q)$ , and  $\bar{Q}_k^* \in \mathcal{C}$ , we have

$$A(\bar{Q}_k^*) \leq A(\bar{Q}^a). \quad (20)$$

Noting that  $B_k(\mathbf{q}) \geq 0$ , and that  $\mathbf{q}_k^*$  maximizes  $F_k(\mathbf{q})$ , we also have

$$A(\bar{Q}_k^*) = F_k(\mathbf{q}_k^*) + B_k(\mathbf{q}_k^*) \geq F_k(\mathbf{q}_k^*) \geq F_k(\mathbf{q}^a) = A(\bar{Q}^a) - B_k(\mathbf{q}^a) \rightarrow A(\bar{Q}^a) \quad (21)$$

where the limit obtains as  $k \rightarrow \infty$  and is implied by the fact that for any fixed profile  $\mathbf{q}$ ,  $\lim_{k \rightarrow \infty} B_k(\mathbf{q}) = 0$ .

Hence, putting together the inequalities (20) and (21), we obtain the limit

$$\lim_{k \rightarrow \infty} A(\bar{Q}_k^*) = A(\bar{Q}^a). \quad (22)$$

To finish, we observe that the limit (22) implies that  $\bar{Q}_k^*$  converges to  $\bar{Q}^a$  for every realization of joint signals, because every realization of joint signals occurs with positive probability and  $\bar{Q}^a$  is the unique maximizer of  $A(Q)$  for  $Q \in \mathcal{C}$ . Hence,  $\bar{Q}_k^* \rightarrow \bar{Q}^a$  as  $k \rightarrow \infty$ , where the limit holds for every realization of joint signals.

## A.6 Additional Properties of Conditional Expectations

In this section, we state and prove a lemma that summarizes two additional properties of conditional expectations of variables  $\theta_i$  that follow from our assumptions.

**Lemma A.1** *For every  $i$ , conditional expectations of signal  $\theta_i$  satisfy the following properties:*

1. *For every subset  $T$  of players, for any profile of their finite signal realizations  $\tilde{s}_T$  that has a positive probability, we have  $\mathbb{E}[\theta_i | s_T = \tilde{s}_T] = 0$ .*
2. *For every  $j \neq i$ , for every pair of finite signal realizations  $\tilde{s}_i$  and  $\tilde{s}_j$ , there exists a matrix  $Q_{ji}(\tilde{s}_j, \tilde{s}_i)$  such that for every infinite signal realization  $\tilde{\theta}_j$ , the conditional expectation of  $\theta_i$  is given by  $\mathbb{E}[\theta_i | s_j = \tilde{s}_j, s_i = \tilde{s}_i, \theta_j = \tilde{\theta}_j] = Q_{ji}^T(\tilde{s}_j, \tilde{s}_i) \tilde{\theta}_j$ .*

**Proof.**

**Statement 1:** It is sufficient to prove the statement for the case  $i \in T$ : if the statement is true for all  $i$  and  $T$  such that  $i \in T$ , then for any  $T$  and  $i \notin T$ ,  $\mathbb{E}[\theta_i | s_T = \tilde{s}_T] = \sum_{\tilde{s}_i} \mathbb{E}[\theta_i | s_T = \tilde{s}_T, s_i = \tilde{s}_i] \mathbb{P}[s_i = \tilde{s}_i | s_T = \tilde{s}_T] = 0$ .

Now suppose  $i \in T$ . Consider a new random variable,  $\chi$ , which is equal to  $\theta_i$  when  $s_T = \tilde{s}_T$  and is equal to zero otherwise. Note that  $\mathbb{E}[\chi] = \mathbb{E}[\theta_i | s_T = \tilde{s}_T] P(\tilde{s}_T)$ , so it is sufficient to show that



$E[\chi] = 0$  (because by assumption,  $P(\tilde{s}_T) \neq 0$ ). Decompose  $\tilde{s}_T$  as  $\tilde{s}_T = (\tilde{s}_i, \tilde{s}_{T \setminus i})$ . Then

$$\begin{aligned}
E[\chi] &= P(\tilde{s}_i) E[\chi | s_i = \tilde{s}_i] \\
&= P(\tilde{s}_i) \int_{\tilde{\theta}_i} \tilde{\theta}_i P(s_{T \setminus i} = \tilde{s}_{T \setminus i} | s_i = \tilde{s}_i, \theta_i = \tilde{\theta}_i) dP(\tilde{\theta}_i | s_i = \tilde{s}_i) \\
&= P(\tilde{s}_i) \int_{\tilde{\theta}_i} \tilde{\theta}_i P(s_{T \setminus i} = \tilde{s}_{T \setminus i} | s_i = \tilde{s}_i) dP(\tilde{\theta}_i | s_i = \tilde{s}_i) \\
&= P(\tilde{s}_i) P(s_{T \setminus i} = \tilde{s}_{T \setminus i} | s_i = \tilde{s}_i) \int_{\tilde{\theta}_i} \tilde{\theta}_i dP(\tilde{\theta}_i | s_i = \tilde{s}_i) \\
&= P(\tilde{s}_i) P(s_{T \setminus i} = \tilde{s}_{T \setminus i} | s_i = \tilde{s}_i) E[\theta_i | s_i = \tilde{s}_i] \\
&= 0.
\end{aligned}$$

Note that a key step in the calculation is the substitution  $P(s_{T \setminus i} = \tilde{s}_{T \setminus i} | s_i = \tilde{s}_i, \theta_i = \tilde{\theta}_i) = P(s_{T \setminus i} = \tilde{s}_{T \setminus i} | s_i = \tilde{s}_i)$ , which is one of our assumptions on the joint distribution of finite and infinite signals.

**Statement 2:** By assumption, the conditional expectation of  $\theta_i$  is a linear function of  $\tilde{\theta}_j$ , and so  $E[\theta_i | s_j = \tilde{s}_j, s_i = \tilde{s}_i, \theta_j = \tilde{\theta}_j] = q_{ji}(\tilde{s}_j, \tilde{s}_i) + Q_{ji}^T(\tilde{s}_j, \tilde{s}_i) \tilde{\theta}_j$  for some vector  $q_{ji}(\tilde{s}_j, \tilde{s}_i)$  and matrix  $Q_{ji}^T(\tilde{s}_j, \tilde{s}_i)$ . To prove that  $q_{ji}(\tilde{s}_j, \tilde{s}_i) = 0$ , consider the conditional expectation  $E[\theta_i | s_j = \tilde{s}_j, s_i = \tilde{s}_i]$ , and note that on one hand, we have  $E[\theta_i | s_j = \tilde{s}_j, s_i = \tilde{s}_i] = 0$ , and on the other hand,  $E[\theta_i | s_j = \tilde{s}_j, s_i = \tilde{s}_i] = q_{ji}(\tilde{s}_j, \tilde{s}_i) + Q_{ji}^T(\tilde{s}_j, \tilde{s}_i) E[\theta_j | s_j = \tilde{s}_j, s_i = \tilde{s}_i] = q_{ji}(\tilde{s}_j, \tilde{s}_i)$ . ■

## A.7 Proof of Lemma 3

Consider the optimization problem of player  $i$  who has observed realization  $(\tilde{s}_i, \tilde{\theta}_i)$  of signal  $(s_i, \theta_i)$ . If he chooses action  $a_i$ , his expected payoff, ignoring the term  $h_i(a_i, s, \theta)$  which does not impact incentives, is equal to

$$E \left[ \pi_i | a_i, \tilde{s}_i, \tilde{\theta}_i \right] = \frac{1}{2} a_i^T E \left[ \Gamma_{ii}(s) | \tilde{s}_i, \tilde{\theta}_i \right] a_i + a_i^T E \left[ \Gamma_{i,-i}(s) a_{-i} | \tilde{s}_i, \tilde{\theta}_i \right] + a_i^T E \left[ g_i(s) | \tilde{s}_i, \tilde{\theta}_i \right].$$

Since  $E \left[ \Gamma_{ii}(s) | \tilde{s}_i, \tilde{\theta}_i \right] = E \left[ \Gamma_{ii}(s) | \tilde{s}_i \right]$  is negative definite for all  $\tilde{s}_i$ , the unique best response of player  $i$  is to set  $a_i(\tilde{s}_i, \tilde{\theta}_i)$  to the value that satisfies the first-order condition

$$E \left[ \Gamma_{ii}(s) | \tilde{s}_i, \tilde{\theta}_i \right] a_i(\tilde{s}_i, \tilde{\theta}_i) + \sum_{j \neq i} E \left[ \Gamma_{ij}(s) a_j(s_j, \theta_j) | \tilde{s}_i, \tilde{\theta}_i \right] + E \left[ g_i(s, \theta) | \tilde{s}_i, \tilde{\theta}_i \right] = 0, \quad (23)$$

and a necessary and sufficient condition for a profile of linear strategies  $a_j(s_j, \theta_j) = \kappa_j(s_j) + \Lambda_j(s_j) \theta_j$  for every player  $j$  to be a linear equilibrium is that the first-order condition (23) is satisfied for every  $i$  and every signal realization  $(\tilde{s}_i, \tilde{\theta}_i)$ .

Bringing the first term under the summation sign in equation (23), we get

$$\sum_{j=1}^n E \left[ \Gamma_{ij}(s) a_j(s_j, \theta_j) | \tilde{s}_i, \tilde{\theta}_i \right] + E \left[ g_i(s, \theta) | \tilde{s}_i, \tilde{\theta}_i \right] = 0. \quad (24)$$

By the law of iterated expectations, conditioning over all  $k_j$  possible realizations of signal  $\tilde{s}_j$ , we can rewrite (24) as

$$\sum_{j=1}^n \sum_{\tilde{s}_j=1}^{k_j} \text{P}[s_j = \tilde{s}_j | s_i = \tilde{s}_i] \text{E} \left[ \Gamma_{ij}(s) a_j(s_j, \theta_j) \middle| \tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i \right] + \text{E} \left[ g_i(s, \theta) \middle| \tilde{s}_i, \tilde{\theta}_i \right] = 0. \quad (25)$$

For simplicity, in the remainder of the proof, we write  $\text{P}[\tilde{s}_j | \tilde{s}_i]$  to denote  $\text{P}[s_j = \tilde{s}_j | s_i = \tilde{s}_i]$ .

By the law of iterated expectations and our assumption that the distribution of  $s$  given  $(\tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i, \tilde{\theta}_j)$  is equal to the distribution of  $s$  given  $(\tilde{s}_i, \tilde{s}_j)$ , we get

$$\text{E} \left[ \Gamma_{ij}(s) \middle| \tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i \right] = \text{E} \left[ \text{E} \left[ \Gamma_{ij}(s) \middle| \tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i, \tilde{\theta}_j \right] \middle| \tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i \right] = \text{E} \left[ \Gamma_{ij}(s) \middle| \tilde{s}_i, \tilde{s}_j \right].$$

Then, recalling our notation  $\text{E} \left[ g_i(s, \theta) \middle| \tilde{s}_i, \tilde{\theta}_i \right] = G_i(\tilde{s}_i) + F_i(\tilde{s}_i) \tilde{\theta}_i$ , equation (25) becomes

$$\sum_{j, \tilde{s}_j} \text{P}[\tilde{s}_j | \tilde{s}_i] \left( \text{E} \left[ \Gamma_{ij}(s) \middle| \tilde{s}_i, \tilde{s}_j \right] \kappa_j(\tilde{s}_j) + \text{E} \left[ \Gamma_{ij}(s) \Lambda_j(\tilde{s}_j) \tilde{\theta}_j \middle| \tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i \right] \right) + G_i(\tilde{s}_i) + F_i(\tilde{s}_i) \tilde{\theta}_i = 0. \quad (26)$$

Let us use again the law of iterated expectations and our assumption that the distribution of  $s$  given  $(\tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i, \tilde{\theta}_j)$  is equal to the distribution of  $s$  given  $(\tilde{s}_i, \tilde{s}_j)$  to obtain the following equalities:

$$\begin{aligned} \text{E} \left[ \Gamma_{ij}(s) \Lambda_j(\tilde{s}_j) \tilde{\theta}_j \middle| \tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i \right] &= \text{E} \left[ \text{E} \left[ \Gamma_{ij}(s) \Lambda_j(\tilde{s}_j) \tilde{\theta}_j \middle| \tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i, \tilde{\theta}_j \right] \middle| \tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i \right] \\ &= \text{E} \left[ \text{E} \left[ \Gamma_{ij}(s) \middle| \tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i, \tilde{\theta}_j \right] \Lambda_j(\tilde{s}_j) \tilde{\theta}_j \middle| \tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i \right] \\ &= \text{E} \left[ \text{E} \left[ \Gamma_{ij}(s) \middle| \tilde{s}_i, \tilde{s}_j \right] \Lambda_j(\tilde{s}_j) \tilde{\theta}_j \middle| \tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i \right] \\ &= \text{E} \left[ \Gamma_{ij}(s) \middle| \tilde{s}_i, \tilde{s}_j \right] \Lambda_j(\tilde{s}_j) \text{E} \left[ \tilde{\theta}_j \middle| \tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i \right]. \end{aligned}$$

Recalling our notation  $\text{E} \left[ \tilde{\theta}_j \middle| \tilde{s}_i, \tilde{s}_j, \tilde{\theta}_i \right] = Q_{ij}^T(\tilde{s}_i, \tilde{s}_j) \tilde{\theta}_i$ , equation (26) can then be written

$$\sum_{j, \tilde{s}_j} \text{P}[\tilde{s}_j | \tilde{s}_i] \text{E} \left[ \Gamma_{ij}(s) \middle| \tilde{s}_i, \tilde{s}_j \right] \left( \kappa_j(\tilde{s}_j) + \Lambda_j(\tilde{s}_j) Q_{ij}^T(\tilde{s}_i, \tilde{s}_j) \tilde{\theta}_i \right) + G_i(\tilde{s}_i) + F_i(\tilde{s}_i) \tilde{\theta}_i = 0. \quad (27)$$

This equation is linear in  $\tilde{\theta}_i$ . By assumption, the variance of  $\theta_i$  has full rank, so for (27) to hold for all possible realizations  $\tilde{\theta}_i$  of  $\theta_i$ , both the constant term and the multiplier of  $\tilde{\theta}_i$  must be equal to zero.<sup>19</sup>

<sup>19</sup>To see this, consider a square matrix  $M \in \mathbb{R}^{\ell \times \ell}$ , a vector  $N \in \mathbb{R}^{\ell}$ , and a random vector  $X \in \mathbb{R}^{\ell}$  with  $\text{E}[X] = 0$  and finite second moments, with full rank variance matrix  $\text{Var}(X)$ . Suppose that the equality  $M\tilde{X} + N = 0$  holds for all possible realizations  $\tilde{X}$  of  $X$ . Then  $\text{E}[M\tilde{X} + N] = 0$ . First, note that as  $\text{E}[X] = 0$ , we must have  $N = 0$ . Second, we now have  $M\tilde{X} = 0$ , which implies  $M\tilde{X}\tilde{X}^T = 0$ , and so  $\text{E}[M\tilde{X}\tilde{X}^T] = 0$ . Since  $\text{E}[M\tilde{X}\tilde{X}^T] = M \text{Var}(X)$  and  $\text{Var}(X)$  has full rank, we must have  $M = 0$ .

The equation corresponding to the constant term is

$$\sum_{j, \tilde{s}_j} \text{P} [\tilde{s}_j | \tilde{s}_i] \text{E} [\Gamma_{ij}(s) | \tilde{s}_i, \tilde{s}_j] \kappa_j(\tilde{s}_j) + G_i(\tilde{s}_i) = 0. \quad (28)$$

and using the notation introduced in Section 2.4, we can rewrite (28) as

$$\sum_{j, \tilde{s}_j} \Phi_{(i, \tilde{s}_i), (j, \tilde{s}_j)} \kappa_j(\tilde{s}_j) + G_i(\tilde{s}_i) = 0. \quad (29)$$

Stacking equations (29) for all  $i = 1, \dots, n$  and all  $\tilde{s}_i = 1, \dots, k_i$ , we get

$$\Phi \kappa + g = 0. \quad (30)$$

Note that (30) is the same as (12), where  $\kappa_j(\tilde{s}_j)$  is used instead of  $a_j(\tilde{s}_j)$  and  $G_i(\tilde{s}_i)$  is used instead of  $\text{E} [g_i(s) | \tilde{s}_i]$ . If  $\Phi$  is invertible, (30) admits the unique solution

$$\kappa = -\Phi^{-1}g.$$

If  $\Phi$  is not invertible, then (30) has either zero or infinitely many solutions.

Now we write the equation corresponding to the multiplier of  $\tilde{\theta}_i$ :

$$\sum_{j, \tilde{s}_j} \text{P} [\tilde{s}_j | \tilde{s}_i] \text{E} [\Gamma_{ij}(s) | \tilde{s}_i, \tilde{s}_j] \Lambda_j(\tilde{s}_j) Q_{ij}^T(\tilde{s}_i, \tilde{s}_j) + F_i(\tilde{s}_i) = 0. \quad (31)$$

We vectorize this equation and apply the identity  $\text{vec}(XYZ) = (Z^T \otimes X) \text{vec}(Y)$ :<sup>20</sup>

$$\sum_{j, \tilde{s}_j} \text{P} [\tilde{s}_j | \tilde{s}_i] (Q_{ij}(\tilde{s}_i, \tilde{s}_j) \otimes \text{E} [\Gamma_{ij}(s) | \tilde{s}_i, \tilde{s}_j]) \text{vec} \Lambda_j(\tilde{s}_j) + \text{vec} F_i(\tilde{s}_i) = 0. \quad (32)$$

Stacking equations (32) for all  $i = 1, \dots, n$  and all  $\tilde{s}_i = 1, \dots, k_i$ , we get

$$\Psi \Lambda + f = 0. \quad (33)$$

If  $\Psi$  is invertible, then (33) admits the unique solution

$$\Lambda = -\Psi^{-1}f.$$

If  $\Psi$  is not invertible, then (33) has either zero or infinitely many solutions.

Therefore, if both  $\Phi$  and  $\Psi$  are invertible, then there exists a unique linear equilibrium given by  $\kappa = -\Phi^{-1}g$  and  $\Lambda = -\Psi^{-1}f$ . Conversely, if  $\Phi$  or  $\Psi$  are not invertible, then either there exists no linear equilibrium, or there are infinitely many linear equilibria.

<sup>20</sup>See [https://en.wikipedia.org/wiki/Vectorization\\_\(mathematics\)#Compatibility\\_with\\_Kronecker\\_products](https://en.wikipedia.org/wiki/Vectorization_(mathematics)#Compatibility_with_Kronecker_products) for details.

## A.8 Proof of Lemma 4

The proof is analogous to the proof of Lemma 2. Similar to that proof, let  $\Phi(\gamma)$  and  $\Psi(\gamma)$  denote the matrices  $\Phi$  and  $\Psi$  corresponding to the quadratic game with parameter  $\gamma$ . As in that proof,  $\det \Phi(\gamma)$  and  $\det \Psi(\gamma)$  are polynomials in  $\gamma$ , and so it is sufficient to prove that neither of these polynomials is identically equal to zero.

Consider the quadratic game with  $\gamma = 0$ , so that payoffs are  $\pi_i(a_i, a_{-i}, s, \theta) = \frac{1}{2}a_i^T \Gamma_{ii}(s)a_i + a_i^T g_i(s, \theta) + h_i(a_{-i}, s, \theta)$ . Since  $E[\Gamma_{ii}(s)|\tilde{s}_i, \tilde{\theta}_i]$  is negative definite for all  $\tilde{s}_i$  and  $\tilde{\theta}_i$ , there is a unique optimal action for each player  $i$ , and this optimal action is independent of the actions of other players. Also, since by assumption  $E[\Gamma_{ii}(s)|\tilde{s}_i, \tilde{\theta}_i]$  does not depend on  $\tilde{\theta}_i$  and  $E[g_i(s, \theta)|\tilde{s}_i, \tilde{\theta}_i]$  is a linear function in  $\tilde{\theta}_i$ , the optimal strategy of player  $i$  is also linear in  $\tilde{\theta}_i$ . Thus, the quadratic game with  $\gamma = 0$  has a unique linear equilibrium, which in turn implies that Conditions 1 and 2 are both satisfied, and so  $\det \Phi(\gamma) \neq 0$  and  $\det \Psi(\gamma) \neq 0$  at  $\gamma = 0$ .

## A.9 Proof of Proposition 4

First, observe that terms  $b_i^*$  do not affect matrices  $\Phi$  and  $\Psi$  in any way: these matrices only depend on the distribution of the profile of finite signals  $s$ , the terms  $\Gamma_{ij}(s)$ , which in the current context depend only on the parameters  $\gamma_{ii}$  and  $\gamma_{ij}$ , and matrices  $Q_{ij}(s_i, s_j)$ , which in the current context depend only on the distribution of the profile of infinite signals  $\theta$ . By Lemma 3, there exists a unique equilibrium if and only if Conditions 1 and 2 are satisfied, i.e., if matrices  $\Phi$  and  $\Psi$  are invertible. Thus, it is sufficient to prove that the game has a unique equilibrium if all bliss points  $b_i^*$  are always equal to zero: this fact would imply that Conditions 1 and 2 are satisfied for that modified game, which would in turn imply that Conditions 1 and 2 are satisfied for the original game, completing the proof.

Consider a beauty contest game with zero bliss points: the payoffs are given by

$$\pi_i(a_i, a_{-i}, s) = -\gamma_{ii}(s)a_i^2 - \sum_{j \neq i} \gamma_{ij}(s)(a_i - a_j)^2. \quad (34)$$

It is immediate that this game has a linear equilibrium in which each player  $i$  always plays  $a_i = 0$  regardless of his information. Let us prove that there are no other linear equilibria in this game.

Consider a linear equilibrium in which each player  $i$  plays  $a_i(s_i, \theta_i) = \kappa_i(s_i) + \Lambda_i(s_i)\theta_i$ . We will first show that for all  $i$  and all realizations  $\tilde{s}_i$ , the constant term  $\kappa_i(\tilde{s}_i)$  is equal to zero, and will then prove an analogous statement about the terms  $\Lambda_i(s_i)$ , thus completing the proof.

Suppose for some  $i$  and some realization  $\tilde{s}_i$ , the term  $\kappa_i(\tilde{s}_i)$  is not equal to zero. Without loss of generality, suppose that for this  $i$  and  $\tilde{s}_i$ , the term  $\kappa_i(\tilde{s}_i)$  has the largest absolute value: for any player  $j$  and signal realization  $\tilde{s}_j$ , we have  $|\kappa_j(\tilde{s}_j)| \leq |\kappa_i(\tilde{s}_i)|$ . Consider the case when player  $i$  observes the realization of his finite signal equal  $\tilde{s}_i$ , and observes the realization of his infinite signal equal to zero. Conditional on this information, the expected value of the infinite signal  $\theta_j$  of any other player  $j$  is also equal to zero, and the expected action of player  $j$  depends only on the  $\kappa_j$

terms of his strategy, and not on  $\Lambda_i$  terms. The optimal action of player  $i$ , which by assumption is given by  $\kappa_i(\tilde{s}_i)$ , must satisfy the following first-order condition:

$$\mathbb{E}[\gamma_{ii}(s)|\tilde{s}_i]\kappa_i(\tilde{s}_i) + \sum_{j \neq i} \mathbb{E}[\gamma_{ij}(s)(\kappa_i(\tilde{s}_i) - \kappa_j(s_j))|\tilde{s}_i] = 0.$$

Since  $\gamma_{ii}(s)$  is always positive,  $\gamma_{ij}(s)$  is always non-negative, and  $\kappa_i(\tilde{s}_i)$  has the largest possible absolute value among all  $\kappa_j(s_j)$ , the above expression has the same sign as  $\kappa_i(\tilde{s}_i)$ , and in particular is not equal to zero, contradicting the assumption of equilibrium. Thus, all terms  $\kappa_i(s_i)$  are equal to zero, and the linear equilibrium behavior of each player  $i$  reduces to  $a_i(s_i, \theta_i) = \Lambda_i(s_i)\theta_i$ .

We will now show that terms  $\Lambda_i(s_i)$  also have to be equal to zero (and thus the actions of all players are constant, identically equal to zero). Suppose that is not the case. Take player  $i$  and finite signal realization  $\tilde{s}_i$  such that the variance  $\text{Var}(a_i|\tilde{s}_i) = \Lambda_i(\tilde{s}_i) \text{Var}(\theta_i)\Lambda_i(\tilde{s}_i)^T$  is maximal, i.e., for every  $j$  and every signal realization  $\tilde{s}_j$ ,  $\text{Var}(a_j|\tilde{s}_j) \leq \text{Var}(a_i|\tilde{s}_i)$  (recall that actions are one-dimensional in this application).

Suppose in addition to the finite signal realization  $\tilde{s}_i$ , player  $i$  also observes an infinite signal realization  $\tilde{\theta}_i$ . His first-order condition then implies that

$$\mathbb{E}[\gamma_{ii}(s)|\tilde{s}_i]\Lambda_i(\tilde{s}_i)\tilde{\theta}_i + \sum_{j \neq i} \mathbb{E}[\gamma_{ij}(s)(\Lambda_i(\tilde{s}_i)\tilde{\theta}_i - \Lambda_j(s_j)\theta_j)|\tilde{s}_i, \tilde{\theta}_i] = 0,$$

which can be rewritten as

$$\mathbb{E}[\gamma_{ii}(s)|\tilde{s}_i]\Lambda_i(\tilde{s}_i)\tilde{\theta}_i + \sum_{j \neq i} \mathbb{E}[\gamma_{ij}(s)(\Lambda_i(\tilde{s}_i)\tilde{\theta}_i - \Lambda_j(s_j) \mathbb{E}[\theta_j|\tilde{\theta}_i])|\tilde{s}_i] = 0.$$

Multiplying this equation by  $\tilde{\theta}_i^T \Lambda_i(\tilde{s}_i)^T$  on the right, and taking the expectation over  $\theta_i$ , we get

$$\mathbb{E}[\gamma_{ii}(s)|\tilde{s}_i] \text{Var}(a_i|\tilde{s}_i) + \sum_{j \neq i} \mathbb{E}[\gamma_{ij}(s)(\text{Var}(a_i|\tilde{s}_i) - \Lambda_j(s_j) \text{Cov}(\theta_j, \theta_i)\Lambda_i(\tilde{s}_i)^T)|\tilde{s}_i] = 0. \quad (35)$$

The final step of the proof is to observe that for every player  $j$  and every realization  $\tilde{s}_j$  of his finite signal, the expression  $\Lambda_j(\tilde{s}_j) \text{Cov}(\theta_j, \theta_i)\Lambda_i(\tilde{s}_i)^T$  is equal to  $\text{Cov}(a_j, a_i|\tilde{s}_j, \tilde{s}_i)$ , for which we have  $\text{Cov}(a_j, a_i|\tilde{s}_j, \tilde{s}_i) \leq \sqrt{\text{Var}(a_j|\tilde{s}_j) \cdot \text{Var}(a_i|\tilde{s}_i)} \leq \text{Var}(a_i|\tilde{s}_i)$  (where the second inequality follows from the choice of player  $i$  and signal realization  $\tilde{s}_i$ ). Thus, the left-hand side of equation (35) is strictly positive, contradicting the assumption that it is equal to zero. Thus, the modified game has a unique linear equilibrium, in which every player always plays zero, and therefore, as explained in the beginning of the proof, the original game also has a unique equilibrium.

## A.10 Proof of Proposition 5

1. The derivative of  $(\beta_L + \beta_H)/2$  with respect to  $\Delta$  is

$$\frac{2\rho^2\Delta(2q-1)(1-\rho)(\alpha(2q-1)\rho-1)}{(\rho^2(2q-1)(\alpha-\Delta)(\alpha+\Delta)-2\alpha q\rho+1)^2}$$

which is negative because  $2q-1 > 0$  (as  $q > 1/2$ ),  $1-\rho > 0$ , and  $\alpha(2q-1)\rho-1 < \alpha\rho-1 < 0$ .

Next, the derivative of  $(\beta_L + \beta_H)/2$  with respect to  $q$  is

$$\frac{2\Delta^2\rho^2(\rho-1)}{(\rho^2(2q-1)(\alpha-\Delta)(\alpha+\Delta)-2\alpha q\rho+1)^2}$$

which is negative, because  $\rho-1 < 0$ .

2. The derivative of  $\beta_H$  with respect to  $q$  is:

$$\frac{2\rho^2\Delta(\rho-1)(\alpha+\Delta)(-\alpha\rho+\Delta\rho+1)}{(\rho^2(2q-1)(\alpha-\Delta)(\alpha+\Delta)-2\alpha q\rho+1)^2}$$

which is negative, because  $\rho-1 < 0$  and because  $-\alpha\rho+\Delta\rho+1 = 1-\rho(\alpha-\Delta) > 0$ .

The derivative of  $\beta_H$  with respect to  $\Delta$  is:

$$\frac{(\rho-1)\rho((2q-1)\rho^2((\alpha+\Delta)^2-4\alpha\Delta q)-2\rho(\alpha q-2\Delta q+\Delta)+1)}{(\rho^2(2q-1)(\alpha-\Delta)(\alpha+\Delta)-2\alpha q\rho+1)^2}. \quad (36)$$

To evaluate the sign of (36) we will show that the following inequality holds:

$$(2q-1)\rho^2((\alpha+\Delta)^2-4\alpha\Delta q)-2\rho(\alpha q-2\Delta q+\Delta)+1 > 0. \quad (37)$$

To see this, recall that  $\alpha = (\alpha_L + \alpha_H)/2$  and  $\Delta = (\alpha_H - \alpha_L)/2$ . Plugging these in the above expression, (37) becomes

$$\alpha_H^2(-2q^2+3q-1)\rho^2 + \alpha_H(q-1)\rho + (\alpha_L q\rho - 1)(\alpha_L(2q-1)\rho - 1) > 0.$$

Note that  $-2q^2+3q-1 = -(q-1)(2q-1) > 0$  and so, since  $\alpha_H > \alpha_L$ ,

$$\begin{aligned} & \alpha_H^2(-2q^2+3q-1)\rho^2 + \alpha_H(q-1)\rho + (\alpha_L q\rho - 1)(\alpha_L(2q-1)\rho - 1) \\ & > \alpha_L^2(-2q^2+3q-1)\rho^2 + \alpha_H(q-1)\rho + (\alpha_L q\rho - 1)(\alpha_L(2q-1)\rho - 1) \\ & = q\rho(\alpha_H + \alpha_L(2\alpha_L\rho - 3)) - \rho(\alpha_H + \alpha_L(\alpha_L\rho - 1)) + 1 \end{aligned}$$

which is linear in  $q$ . If  $q = 1$ , we have

$$q\rho(\alpha_H + \alpha_L(2\alpha_L\rho - 3)) - \rho(\alpha_H + \alpha_L(\alpha_L\rho - 1)) + 1 = (\alpha_L\rho - 1)^2 > 0.$$

If  $q = 1/2$ , we have

$$q\rho(\alpha_H + \alpha_L(2\alpha_L\rho - 3)) - \rho(\alpha_H + \alpha_L(\alpha_L\rho - 1)) + 1 = 1 - \frac{1}{2}\rho(\alpha_H + \alpha_L) > 0.$$

So, for every  $q \in (1/2, 1)$ ,

$$q\rho(\alpha_H + \alpha_L(2\alpha_L\rho - 3)) - \rho(\alpha_H + \alpha_L(\alpha_L\rho - 1)) + 1 > 0$$

which proves (37).

Using (37) and the fact that  $\rho - 1 < 0$ , we obtain that (36) is negative.

3. The derivative of  $\beta_L$  with respect to  $q$  is

$$\frac{2\rho^2\Delta(\rho - 1)(\alpha - \Delta)(\rho(\alpha + \Delta) - 1)}{(\rho^2(2q - 1)(\alpha - \Delta)(\alpha + \Delta) - 2\alpha q\rho + 1)^2}$$

which is positive because  $\rho - 1 < 0$ ,  $\alpha - \Delta > 0$  and  $\rho(\alpha + \Delta) - 1 < \rho - 1 < 0$ .

The derivative of  $\beta_L$  with respect to  $\Delta$  is:

$$-\frac{(\rho - 1)\rho(\rho((2q - 1)\rho(4\alpha\Delta q + (\alpha - \Delta)^2) - 2q(\alpha + 2\Delta) + 2\Delta) + 1)}{((2q - 1)\rho^2(\alpha - \Delta)(\alpha + \Delta) - 2\alpha q\rho + 1)^2}. \quad (38)$$

Fix  $\alpha = \frac{7}{8}$ ,  $q = \frac{15}{16}$ , and  $\rho = \frac{95}{96}$ . Then at  $\Delta = \frac{1}{10}$  expression (38) is negative, while at  $\Delta = \frac{1}{11}$  it is positive. Hence the comparative static is in general ambiguous.

## References

- Angeletos, G.-M. and A. Pavan (2007). Efficient use of information and social value of information. *Econometrica* 75(4), 1103–1142.
- Aumann, R. J. (1976). Agreeing to disagree. *The Annals of Statistics* 4(6), 1236–1239.
- Basar, T. (1975). Equilibrium solutions in two-person quadratic decision problems with static information structures. *IEEE Transactions on Automatic Control* 20(3), 320–328.
- Basar, T. (1978a). Decentralized multicriteria optimization of linear stochastic systems. *IEEE Transactions on Automatic Control* 23(2), 233–243.
- Basar, T. (1978b). Equilibrium solutions in static decision problems with random coefficients in the quadratic cost. *IEEE Transactions on Automatic Control* 23(5), 960–962.
- Bergemann, D., T. Heumann, and S. Morris (2017). Information and interaction. Cowles Foundation Discussion Paper 2088.

- Bergemann, D. and S. Morris (2013). Robust predictions in games with incomplete information. *Econometrica* 81(4), 1251–1308.
- Bergemann, D. and S. Morris (2016). Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics* 11(2), 487–522.
- Golub, B. and S. Morris (2017). Expectations, networks, and conventions. Working paper.
- Hurkens, S. (2014). Bayesian Nash equilibrium in “linear” Cournot models with private information about costs. *International Journal of Economic Theory* 10(2), 203–217.
- Lambert, N., M. Ostrovsky, and M. Panov (2017a). Strategic trading in informationally complex environments. Working paper.
- Lambert, N., M. Ostrovsky, and M. Panov (2017b). Supplementary online appendix to “Strategic trading in informationally complex environments”.
- Leister, C. M. (2016). Information acquisition and welfare in network games. Working paper.
- Li, L., R. D. McKelvey, and T. Page (1987). Optimal research for Cournot oligopolists. *Journal of Economic Theory* 42(1), 140–166.
- Morris, S. and H. S. Shin (2002). Social value of public information. *American Economic Review* 92(5), 1521–1534.
- Myatt, D. and C. Wallace (2017). Information acquisition and use by networked players. Working paper.
- Pachter, M. and K. Pham (2013). Static teams and stochastic games. In A. Sorokin and P. M. Pardalos (Eds.), *Dynamics of Information Systems: Algorithmic Approaches*, pp. 147–175. Springer.
- Palfrey, T. R. (1985). Uncertainty resolution, private information aggregation and the Cournot competitive limit. *Review of Economic Studies* 52(1), 69–83.
- Radner, R. (1962). Team decision problems. *Annals of Mathematical Statistics* 33(3), 857–881.
- Ui, T. (2009). Bayesian potentials and information structures: Team decision problems revisited. *International Journal of Economic Theory* 5(3), 271–291.
- Vives, X. (1988). Aggregation of information in large Cournot markets. *Econometrica* 56(4), 851–876.