Comparing Incomplete Information and Rational Inattention in Games*

Tommaso Denti[†]
Cornell University

Doron Ravid[‡]
University of Chicago

May 1, 2023

PRELIMINARY AND INCOMPLETE
PLEASE DO NOT CIRCULATE

Abstract

We conduct a systematic study of the difference between exogenous information, structured information acquisition, and rational inattention. To do so, we fix a base game, which specifies the set of actions, payoff states, payoff functions, and the common prior over the state, and compare the set of payoffs and outcomes (i.e., action-state distributions) that each informational regime can induce. Under exogenous information, this set is characterized by the Bayes Correlated Equilibrium (Bergemann and Morris, 2016). We show structured information acquisition attains the same outcomes, but allows for lower payoffs from a given outcome. Hence, the welfare implications of an outcome can depend on whether the information generating it is acquired or exogenously given. By contrast, rational inattention can only generate outcomes that satisfy an additional separation constraint. However, this constraint never binds in generic games. Thus, for rational inattention to have bite, one must either commit to a non-generic setting, or impose restrictions on the underlying learning environment. For convex restrictions, we characterize when rational inattention refines the set of attainable outcomes, and show that the refinement is dramatic whenever it occurs.

1. Introduction

In many economic environments, information is abundant but attention is scarce. Rational inattention (Sims, 2003) is an influential model of how economic agents allocate attention. According to this model, agents pay attention "as if" they optimally solve a problem of

*We would like to thank Benjamin Brooks and Luciano Pomatto for helpful discussions.
[†]tjd237@cornell.edu
[‡]dravid@uchicago.edu

costly information acquisition. The cost of acquiring information reflects subjective and psychological factors that limits the agents' ability to pay attention.

Over the last two decades, rational inattention has seen a wide range of applications. It has been used to explain key macroeconomic phenomena such as price rigidity and business cycle dynamics (Maćkowiak and Wiederholt, 2009, 2015), and to justify puzzling features of financial markets; among others, the home bias puzzle and the lack of diversification (Van Nieuwerburgh and Veldkamp, 2009, 2010). Rational inattention has also attracted the interest of behavioral economists who see it as a parsimonious relaxation of the Homo economicus paradigm (Caplin, Dean, and Leahy, 2019; Kőszegi and Matějka, 2020). The sprawling literature on rational inattention reaches fields as diverse as labor economics, trade, development, and political economy.[1]

For all its merits, the rational inattention model also raises important concerns. In the name of tractability, most applications of rational inattention make demanding functional-form assumptions on the cost of acquiring information. Such assumptions are not easily verifiable since attention limitations are affected by factors that are hard to measure directly, such as time, effort, and cognitive resources. A few tests of rational inattention have been carried out in laboratory settings, and the results have been concerning; for example, Dean and Neligh (2022) provide sharp experimental evidence against the ubiquitous entropy-reduction cost (Matějka and McKay, 2015). Overall, this raises the question of what predictions of rational inattention are robust to the specification of the agents' costs of information; this paper aims to address this question.

We study games with information acquisition. There are finitely many players who simultaneously choose from a number of alternative actions. The payoff of an action depends on the other players' actions and an exogenous state of nature. Players begin the game without knowing what the true state is, and have the opportunity to acquire information about it before taking their action. In addition to learning about the state, players can choose to obtain costly information about other "non-fundamental" features of the environment. This class of static games is sufficiently general to cover many settings that are relevant to rational inattention and, more broadly, to information economics.

Our initial finding is a robust characterization of the set of outcomes that can arise in an equilibrium of a game with information acquisition. We take the perspective of an analyst who knows the basic structure of the game (e.g., the analyst knows whether the game is a beauty contest or an auction), but does not know the players' costs of acquiring information (are they based on entropy? homogeneous? etc.). We only make the basic assumption that more information—in the sense of Blackwell (1951, 1953)—is more costly. We show that the implications of this monotonicity condition are captured by a new solution concept for

[1] See Maćkowiak, Matějka, and Wiederholt (2023) for a recent survey.

games with payoff uncertainty that we dub *Blackwell correlated equilibrium* (BKE).

A BKE is a refinement of *Bayes correlated equilibrium* (BCE), a well-known solution concept due to **?**. A BCE is a joint distribution over actions and payoff-relevant states that satisfies an *obedience constraint*: each action must be optimal given the beliefs that the action itself induces about the other players' actions and the state. **?** show the BCE set characterizes the outcomes attainable by some exogenous information structure.

A BKE is a BCE that satisfies an additional constraint: no pair of actions that induce distinct beliefs can be indifferent to a third common action (indifferent according to the respective beliefs). Thus, distinct beliefs have separate best responses; we refer to this condition as the *separation constraint*. Theorem 1 presented below shows that an outcome can be induced by some information acquisition game with Blackwell-monotone costs if and only if it is a BKE. Thus, whereas BCE characterizes the set of outcomes that can arise in an equilibrium of a game with a predetermined information structure, BKE gives the outcomes that can arise when the information structure is endogenous.

In the core of the paper, we study BKE and its relationship with BCE. Theorem 2 shows that BKE is an all-or-nothing refinement of BCE: the set of BKE is either dense or nowhere dense in the set of BCE–there is no middle ground. Theorem 3 shows that for generic games the set of BKE is dense in the set of BCE. To prove these results, we develop new techniques—such as the notion of minimally mixed BCE—and build on existing tools—e.g., the notion of jeopardization (Myerson, 1997) and the semi-algebraic structure of solution concepts (Blume and Zame, 1994).

Finally, we develop an economic application to Bertrand competition where BKE is a significant refinement of BCE. Specifically, we consider Bertrand competition between two firms who are uncertain about production costs. We characterize the consumer surplus that can be achieved in a BKE and in a BCE. We show that the largest consumer surplus that is achievable in a BCE may be much larger than the largest consumer surplus that is achievable in a BKE. The intuition is that competition erodes profits and the incentive to acquire information, substantially reducing the set of information structures that can be sustained in equilibrium.

Related literature. A condition related to the separation constraint that characterizes BKE appears in Denti (2021). Denti (2021) analyzes costly monitoring in signaling games: the receiver has to put costly effort to learn the sender's action. Among other results, he characterizes the equilibrium outcomes that can arise in a signaling game where the receiver's monitoring costs are Blackwell monotone. One of the condition that he finds is a separation constraint: no pair of the receiver's actions that induce distinct beliefs about the sender's action can be indifferent to a third common action. Compared to Denti

(2021), a key contribution of this paper is that our framework involves many players who simultaneously acquire information in a correlated fashion. In Denti (2021), only the receiver actively acquires information; the sender's information is predetermined.

This paper is particularly relevant to the information design literature (Kamenica and Gentzkow, 2011; Bergemann and Morris, 2019). Within this literature, the BCE solution concept has been used to obtain robust predictions that are valid across all *exogenously-given* information structures (e.g., Bergemann and Morris, 2013; Bergemann, Brooks, and Morris, 2017). By spanning the outcomes attainable under *endogenous* information, the BKE solution concept should enable one to obtain sharper conclusions that are valid whenever agents choose what information to collect. Just as BCE has revealed new desirable mechanisms (e.g., Du, 2018; Brooks and Du, 2021), BKE may allow the design of robust mechanisms in scenarios with costly information acquisition. Finally, BKE can be useful to study robust Bayesian persuasion (Dworczak and Pavan, 2022). For example, one may apply BKE in attention management settings where a principal can choose the information available to a rationally inattentive agent (Lipnowski, Mathevet, and Wei, 2020). In such settings the BKE concept could be used to find solutions that are robust to the specific nature of the agents' attention limitations.

This paper contributes to the literature on information acquisition in games. Motivated by questions in macroeconomics and finance, several authors have recently studied information acquisition in coordination games–among others, Hellwig and Veldkamp (2009), Myatt and Wallace (2012), Colombo, Femminis, and Pavan (2014), **?**, Hebert and La'O (forthcoming), **?**, and Morris and Yang (2022). Information acquisition and mechanism design also is a classic research topic (e.g., Persico, 2000; Bergemann and Valimaki, 2002) that has received renewed consideration (e.g., Mensch, 2022; Mensch and Ravid, 2022; Gleyze and Pernoud, forthcoming). Our contribution is to develop tools for robust predictions that do not depend on the specific details of the information technology at the players' disposal.

Our work also expands on the literature on correlated equilibrium and the role of information in incomplete information games (e.g., Aumann, 1974, 1987; Forges, 1993, 2006; Gossner, 2000; Lehrer, Rosenberg, and Shmaya, 2013; Liu, 2015; Doval and Ely, 2020). Whereas this literature focuses on exogenous information structures, we show how to accommodate endogenous information acquisition.

2. Information acquisition games

The object of our study is a class of games where players can preemptively acquire information. Let I be a finite set of players, with typical element i. Each player i has to choose an action a_i from a finite set A_i. For Cartesian products, we adopt the standard notation

$A_{-i} := \prod_{j\neq i} A_j$ and $A := A_i \times A_{-i}$; we use $a_{-i} := (a_j)_{j\neq i}$ and $a := (a_i, a_{-i})$ to denote, respectively, the action profile of all players other than i and the entire action profile.

Players are expected utility maximizers who care about each other's actions as well as an exogenous state θ, which is drawn from a finite set Θ according to a full-support probability measure $\pi \in \Delta(\Theta)$. We denote by $u_i : A \times \Theta \to \mathbb{R}$ player i's utility function. We refer to the tuple $\mathcal{G} := (I, \Theta, \pi, (A_i, u_i)_{i\in I})$ as a **base game**.

Before taking action, each player has the opportunity to acquire information about the payoff-relevant state θ as well as a non-fundamental state z taking values in a finite set Z; a function $\zeta : \Theta \to \Delta(Z)$ details z's conditional distribution given θ.

Following Blackwell (1951), we represent the acquisition of information by an experiment. An **experiment** for player i is a mapping $\xi_i : \Theta \times Z \to \Delta(X_i)$, where X_i is a finite space of signal realizations observable by player i. We denote a profile of signal realizations by $x := (x_i)_{i\in I}$, and take $X := \prod_{i\in I} X_i$ to be the set of all signal profiles.

By construction, the players' signals are conditionally independent given θ and z. However, they may be correlated given θ only. Thus, our framework incorporates a form of correlated information acquisition, as in, among others, Hellwig and Veldkamp (2009), Myatt and Wallace (2012), Hebert and La'O (forthcoming), and **?**.

The acquisition of information faces two kinds of frictions. First, players face a feasibility constraint on the kind of experiments they can use: player i can only choose experiments that lie in a given set $\mathcal{E}_i$. Second, experiments come at a cost, where $C_i : \mathcal{E}_i \to \mathbb{R}_+$ denotes player i's cost function. As a normalization, we assume that there exists an experiment that has cost zero; that is, there exists $\xi_i \in \mathcal{E}_i$ such that $C_i(\xi_i) = 0$. We refer to the tuple $\mathcal{T} := (Z, \zeta, (X_i, \mathcal{E}_i, C_i)_{i\in I})$ as an **information technology**.

Together, a base game $\mathcal{G}$ and an information technology $\mathcal{T}$ define a **game with information acquisition**, our main object of study. The game begins with the realization of the states θ and z. Then, without observing these states, players simultaneously choose experiments, and pay a cost. Finally, players simultaneously observe the outcome of their own experiment and take action (without having knowledge of the experiments chosen by others). We use $\sigma_i : X_i \to \Delta(A_i)$ to denote player i's action plan in this game, and let Σ_i be the set of i's action plans.

The solution concept we adopt is Nash equilibrium. For every player i, a strategy consists of an experiment $\xi_i \in \mathcal{E}_i$ and an action plan $\sigma_i \in \Sigma$. A strategy profile $(\xi_i^*, \sigma_i^*)_{i\in I}$ is an equilibrium if for all players i, (ξ_i^*, σ_i^*) maximizes

$$\left[\sum_{\theta,z,x,a} u_i(a,\theta)\sigma_i(a_i|x_i)\xi_i(x_i|\theta,z)\prod_{j\neq i}\sigma_j^*(a_j|x_j)\xi_j^*(x_j|\theta,z)\zeta(z|\theta)\pi(\theta)\right] - C_i(\xi_i).$$

over all $\xi_i \in \mathcal{E}_i$ and $\sigma_i \in \Sigma_i$. The objective function consists of two terms: the value of information (in square brackets) and the cost of information. As common in applications, value and cost are additively separable.

We summarize the equilibria of information acquisition games using two statistics of the players' behavior: the outcome and the value. The **outcome** is the joint distribution $p \in \Delta(A \times \Theta)$ of the players' actions and the payoff-relevant state. Note that the marginal distribution of p over states must coincide with the prior π; we denote by $\Delta_\pi(A \times \Theta)$ the set of probability measures over $A \times \Theta$ whose marginal on Θ is π. The **value** is the vector $v := (v_i)_{i \in I} \in \mathbb{R}^I$ assigning each player the expected payoff in the game. Observe that this payoff includes players' information acquisition costs and so v_i may differ from the expectation of u_i under p.

Our goal is to develop tools for studying games with information acquisition without committing to the specific details of the information technology. To this end, we analyze the game as one fixes the base game $\mathcal{G}$ and varies the information technology $\mathcal{T}$.

3. Arbitrary information technologies

In this section we present a characterization of the outcome-value pairs that can arise in an equilibrium of a game with information acquisition, keeping fixed the base game and *arbitrarily* varying the information technology. As readily apparent, putting no restrictions on the information technology generates a model that is too permissive and does not reflect the basic idea that information is costly. Nevertheless, the exercise provides a useful benchmark for our main results, which we discuss in the next sections.

In general, the class of information acquisition games includes games in which players' information is predetermined. One can obtain such games by considering information technologies in which each player has only *one* feasible experiment whose cost is zero. We say that such technologies are **degenerate**, and refer to the corresponding information acquisition games as **games with exogenous information**.[2]

Among other results, **?** characterize the equilibrium outcomes that can arise in a game with exogenous information: they call them Bayes correlated equilibria. A **Bayes correlated equilibrium** (BCE) is an outcome $p \in \Delta_\pi(A \times \Theta)$ that satisfies the following constraint: for all $i \in I$ and $a_i, b_i \in A_i$,

$$\sum_{a_{-i},\theta} \left(u_i(a_i, a_{-i}, \theta) - u_i(b_i, a_{-i}, \theta)\right) p(a_i, a_{-i}, \theta) \geq 0. \tag{1}$$

Following Bergemann and Morris, we name (1) the **obedience constraint** since it guarantees that each player is willing to obey the (real or fictitious) mediator who recommends

[2] Of course, these games are just standard Bayesian games à la Harsanyi.

her which action to take.[3]

Because exogenous information comes at zero cost, each player's expected payoff given an outcome p is given by the **gross value** of p,

$$\bar{v}_i(p) := \sum_{a,\theta} u_i(a,\theta)p(a,\theta),$$

which is i's expected payoff under p, ignoring information costs; set $\bar{v}(p) = (\bar{v}_i(p) : i \in I)$. Thus, Bergemann and Morris's analysis delivers the following result.

Proposition 1 (**?**)**.** *Fix a base game $\mathcal{G}$. A degenerate information technology $\mathcal{T}$ exists that induces the outcome-value pair (p, v) in an equilibrium of $(\mathcal{G}, \mathcal{T})$ if and only if*

(i) p is a BCE, and

(ii) for every $i \in I$, $v_i = \bar{v}_i(p)$.

Building on Bergemann and Morris's results, the next proposition characterizes the outcome-value pairs that can arise in an equilibrium of a game with information acquisition, without assumptions on the information technology. To state the proposition, we define the **uninformed** value,

$$\underline{v}_i(p) := \max_{b_i \in A_i} \sum_{a,\theta} u_i(b_i, a_{-i}, \theta)p(a,\theta),$$

which is i's maximal value if she receives no information, and others' actions and θ are distributed according to p; set $\underline{v}(p) = (\underline{v}_i(p))_{i \in I}$.

Proposition 2. *Fix a base game $\mathcal{G}$. An information technology $\mathcal{T}$ exists that induces the outcome-value pair (p, v) in an equilibrium of $(\mathcal{G}, \mathcal{T})$ if and only if*

(i) p is a BCE, and

(ii) for every $i \in I$, $v_i \in [\underline{v}_i(p), \bar{v}_i(p)]$.

For an outcome to be attainable, it must be a BCE: the obedience constraint must be satisfied whether information is exogenous or endogenous. That $\bar{v}_i(p)$ is the maximal payoff player i can attain in an equilibrium that induces the outcome p follows from information acquisition costs being non-negative. To see why $\underline{v}_i(p)$ is a lower bound on player i's payoff, suppose we have an information technology and an equilibrium that induces the outcome p. By assumption, player i can obtain some experiment at zero cost. Therefore, player i cannot get a lower payoff than what she would get if she unilaterally deviated to using this

[3] Our definition of BCE corresponds to the specialization of **?**'s (**?**) definition for the case in which players' original type spaces are degenerate. This specialization was first discussed by Forges (1993), who referred to it as *universal Bayesian solution.*

free experiment. Clearly, the worst payoff the player can get from this deviation is if this free experiment is uninformative. It follows player's i equilibrium payoff must be higher than what she would get if she could get no information at zero cost; that is, player i's payoff must be higher than $\underline{v}_i(p)$.

To show that every outcome-value pair (p, v) satisfying the proposition's conditions (i) and (ii) is attainable in an equilibrium for some information technology, we construct an information technology $\mathcal{T}$ in which each player can use two experiments. The first experiment gives the player no information, and costs zero. The second experiment gives that player's equilibrium information at a cost of $\bar{v}_i(p) - v_i$, which is non-negative because $\bar{v}_i(p) \geq v_i$. This cost implies that, when all players use their equilibrium information, each player i gets a payoff of v_i, whereas unilaterally deviating to no information would give i a payoff of $\underline{v}_i(p)$, which is smaller than v_i by the proposition's conditions. It follows that having each player acquire the informative experiment is an equilibrium with outcome p and value v.

4. Blackwell correlated equilibrium

We now turn to the main focus of the paper: to study environments in which it is harder to acquire more information. To formalize the idea that more information is more costly, we build on a classic notion of informativeness due to Blackwell (1951, 1953).

Given a pair of experiments ξ_i and ξ_i' for a player i, we say that ξ_i **Blackwell dominates** ξ_i' (denoted $\xi_i \succsim \xi_i'$) if there exists a garbling $g : X_i \to \Delta(X_i)$ such that for every $x_i \in X_i$, $\theta \in \Theta$, and $z \in Z$ with $\zeta(z|\theta) > 0$,

$$\xi_i'(x_i|\theta, z) = \sum_{x_i' \in X_i} g(x_i|x_i')\xi(x_i'|\theta, z).$$

As shown by Blackwell, $\xi_i \succsim \xi_i'$ if and only if player i is better off observing the outcome of ξ_i rather than the outcome of ξ_i' (holding fixed other players' behavior). In this sense, ξ_i is more informative than ξ_i'. We write $\xi_i \succ \xi_i'$ whenever $\xi_i \succsim \xi_i'$ and $\xi_i' \not\succsim \xi_i$.

An information technology $\mathcal{T}$ is **Blackwell monotone** if, for every player i, the following two conditions hold. First, if ξ_i is more informative than ξ_i' and ξ_i is feasible, then ξ_i' is feasible as well; that is, if $\xi_i \succsim \xi_i'$ and $\xi_i \in \mathcal{E}_i$, then $\xi_i' \in \mathcal{E}_i$. And second, switching to a less informative experiment lowers the costs of information acquisition: if $\xi_i, \xi_i' \in \mathcal{E}_i$ are such that $\xi_i \succsim \xi_i'$ (resp., $\xi_i \succ \xi_i'$), then $C_i(\xi_i) \geq C_i(\xi_i')$ (resp., $C_i(\xi_i) > C_i(\xi_i')$). Thus, an information technology is monotone if it is always possible to save on costs by reducing the informativeness of one's experiment.

Most applications of rational inattention adopt monotone technologies. For example, one could generate a monotone technology by letting $\mathcal{E}_i$ be the set of all experiments and C_i be the expected reduction in the uncertainty about θ and z as measured by Shannon's entropy (Matějka and McKay, 2015). More broadly, one could substitute Shannon's entropy with any other strictly concave measure of uncertainty (Caplin, Dean, and Leahy, 2022). Since monotonicity is an ordinal concept, one could also consider increasing transformations of these cost functions (Denti, 2022; Zhong, 2022).[4]

Our next theorem characterizes the equilibria that can be attained by monotone information technologies. To present this result, we require a few definitions. Given an outcome $p \in \Delta_\pi(A \times \Theta)$, a player $i \in I$, and an action $a_i \in A_i$, let $p(a_i) := \sum_{a_{-i},\theta} p(a_i, a_{-i}, \theta)$ be the probability of player i taking action a_i under p, and let

$$\mathrm{supp}_i(p) := \{a_i \in A_i : p(a_i) > 0\}$$

be the set of i's actions that have positive probability. For each $a_i \in \mathrm{supp}_i(p)$, let $p_{a_i} \in \Delta(A_{-i} \times \Theta)$ be the conditional distribution of the actions of the players other than i and the payoff-relevant state: for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$,

$$p_{a_i}(a_{-i}, \theta) := \frac{p(a_i, a_{-i}, \theta)}{p(a_i)}.$$

Finally, let

$$BR(p_{a_i}) := \operatorname*{argmax}_{b_i \in A_i} \sum_{a_{-i},\theta} u_i(b_i, a_{-i}, \theta) p_{a_i}(a_{-i}, \theta)$$

be the set of player i's best responses to p_{a_i}.

An outcome $p \in \Delta_\pi(A \times \Theta)$ is a **Blackwell correlated equilibrium** (BKE) if the obedience constraint (1) holds, and, for all $i \in I$ and $a_i, b_i \in \mathrm{supp}_i(p)$,

$$p_{a_i} \neq p_{b_i} \quad \text{implies} \quad BR(p_{a_i}) \cap BR(p_{b_i}) = \varnothing. \tag{2}$$

We call (2) the **separation constraint**: it states that distinct beliefs must have separate best responses.

Theorem 1. *Fix a base game $\mathcal{G}$. A Blackwell-monotone information technology $\mathcal{T}$ exists that induces the outcome-value pair (p, v) in an equilibrium of $(\mathcal{G}, \mathcal{T})$ if and only if*

(i) p is a BKE, and

(ii) for every $i \in I$, $v_i = \underline{v}_i(p) = \bar{v}_i(p)$ or $v_i \in [\underline{v}_i(p), \bar{v}_i(p))$.

[4] In each information acquisition game, prior beliefs are determined by π and ζ, which are exogenous variables. Thus, the issue of experiment-based vs. posterior-based information costs that sometimes arise in applications of rational inattention (see, e.g., Denti, Marinacci, and Rustichini, 2022) is irrelevant here.

Comparing Proposition 2 and Theorem 1, the essential novelty is the separation constraint (2). There is a simple intuition behind it: Consider a player i who takes with positive probability a pair of actions a_i and b_i such that $p_{a_i} \neq p_{b_i}$. As in BCE, we can interpret a_i and b_i as signals. When the information technology is Blackwell monotone, informative signals are costly. To save on information costs, the player could substitute a_i and b_i with a single action recommendation c_i. For this not to be profitable, it must be that either $c_i \notin BR(p_{a_i})$ or $c_i \notin BR(p_{b_i})$. Since the choice of c_i is arbitrary, it must be that $BR(p_{a_i}) \cap BR(p_{b_i}) = \varnothing$.

If there is only one player, that is, I is a singleton, Theorem 1's characterization of equilibrium outcomes follows from Denti (2021), who studies signaling games where the receiver acquires costly information about the sender's action, while the sender's information is exogenous. In these class of games, *only* the information of *one* player is endogenous. Thus, with respect to Denti (2021), the difficulty in proving Theorem 1 is that many players simultaneously acquire information in a correlated fashion. Here we also provide a characterization of the attainable equilibrium payoffs, which has no analogue in Denti (2021).

The notion of Blackwell correlated equilibrium is the cornerstone of our analysis. Next we record that such notion is nonempty, that is, there always exists a BKE.

Proposition 3. *For every base game $\mathcal{G}$, the set of BKE is nonempty.*

A technology where the only feasible experiments are the uninformative experiments (and they all cost zero) is Blackwell monotone. The corresponding game with information acquisition admits an equilibrium by standard arguments (information is de facto exogenous). It follows from Theorem 1 that the outcome of such an equilibrium is a BKE.

5. Bayes vs. Blackwell

Blackwell correlated equilibrium is a refinement of Bayes correlated equilibrium. In this section, we compare the two solution concepts.

We begin with pointing out some basic structural differences. Since BCE is defined by a system of linear inequalities, the BCE set is convex and closed. The BKE set, by contrast, may be neither convex nor closed, as the next two examples highlight.

Example 1 (Not convex)**.** *The set Θ is a singleton and, for every $i \in I$, u_i is a constant function. The players are indifferent over all action profiles: every outcome $p \in \Delta(A)$ is a BCE. The BKE set is the set of product measures: since utilities are constant, the separation constraint is satisfied if and only if $p_{a_i} = p_{b_i}$ for all $i \in I$ and $a_i, b_i \in \mathrm{supp}_i(p)$, that is, if and only if the players' actions are independent of each other. Thus, the BKE set is not convex (as long as at least two distinct players have at least two distinct actions each).*

Example 2 (Not closed)**.** *A single player—i.e., I is a singleton—has two actions: $A = \{c, r\}$. There are two payoff-relevant states, equally likely: $\Theta = \{L, R\}$ with $\pi(L) = \pi(R)$. Action c gives utility 0 regardless of the state, whereas action r gives utility 2 in state R, and -1 otherwise. For every $\epsilon \in [0, 1]$, consider the outcome $p^\epsilon \in \Delta(A \times \Theta)$ given by*

	L	R
c	$\frac{1}{2}$	$\frac{1-\epsilon}{4}$
r	0	$\frac{1+\epsilon}{4}$

For $\epsilon > 0$, p^ϵ is a BKE: each action a is the unique best response to the corresponding belief $p^\epsilon_a \in \Delta(\Theta)$ about the state. However, p^0 is not a BKE: $p^0_c \neq p^0_r$ and $r \in BR(p^0_c) \cap BR(p^0_r)$. Thus, the BKE set is not closed.

In Example 1, there are many more BCEs than BKEs: in fact, the BKE set is *nowhere dense* in the BCE set. In Example 2, not all BCEs are BKEs, but most of them are: one can easily verify that the BKE set is *dense* in the BCE set. Next we show that these two cases are exhaustive, with the latter being generic. As intermediate result, we give a characterization of the games for which the BKE set is dense in the BCE set. The characterization leverages on two concepts: the concept of jeopardization, due to Myerson (1997), and the concept of minimally mixed BCE, which is new to this paper.

For a player i, an action a_i **jeopardizes** an action b_i if, for every BCE p such that $b_i \in \mathrm{supp}_i(p)$, $a_i \in BR(p_{b_i})$.[5] Thus, a_i jeopardizes b_i if, whenever player i takes action b_i in a BCE of the game, they would be equally happy to take action a_i. We denote by $J(b_i)$ the set of actions that jeopardizes b_i.

Every action jeopardizes itself by the obedience constraint; hence, $J(b_i)$ is not empty. A sufficient condition for jeopardization is weak domination: if $u_i(a_i, a_{-i}, \theta) \geq u_i(b_i, a_{-i}, \theta)$ for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$, then a_i jeopardizes b_i. But the concept of jeopardization is broader than weak domination. For example, in Matching pennies, Heads and Tails jeopardize each other, even if neither action is weakly dominant.

A BCE p has **maximal support** if the support of every other BCE is contained by the support of p. A BCE p is **minimally mixed** if it has maximal support and

$$q_{a_i} \neq q_{b_i} \quad \text{implies} \quad p_{a_i} \neq p_{b_i}$$

for every BCE q, $i \in I$, and $a_i, b_i \in \mathrm{supp}_i(q)$.

For an interpretation of minimal mixing, take the perspective of a mediator who wants to implement a BCE p. When $p_{a_i} = p_{b_i}$, the mediator can replace the distinct recommendations of playing a_i and b_i with a single recommendation of mixing between the two actions with

[5] Myerson defines jeopardization for games without payoff uncertainty; here we give the obvious extension to games where Θ is not a singleton.

probabilities $p(a_i)/(p(a_i) + p(b_i))$ and $p(b_i)/(p(a_i) + p(b_i))$. Thus, a BCE p is minimally mixed if a mediator has the least amount of opportunities to implement p recommending mixed actions. This notion may seem esoteric at first. On the contrary, the set of minimally mixed BCEs is open and dense in the BCE set (see Lemma 8 in the appendix).

Our next result uses the concepts of jeopardization and minimally mixed BCEs to characterize when the BCE and the BKE sets coincide.

Proposition 4. *The following statements are equivalent:*

(i) The BKE set is dense in the BCE set.

(ii) A minimally mixed BKE exists.

(iii) For every BCE p, $i \in I$, $a_i, b_i \in \text{supp}_i(p)$,

$$p_{a_i} \neq p_{b_i} \quad \textit{implies} \quad J(a_i) \cap J(b_i) = \varnothing.$$

The result shows how jeopardization and minimal mixing can be used in applications to study BKE. To verify that the BKE set is dense in the BCE set, it is enough to produce a minimally mixed BKE. To verify that the BKE set is not dense in the BCE set, it is enough to produce a BCE in which two actions induce distinct beliefs and share a common jeopardizing action. As shown by Myerson (1997), the jeopardizing actions can be easily computed from the dual of the system of linear inequalities that defines BCE.

To check the conditions of Proposition 4, knowing which actions induce different beliefs for some BCE is useful. The following result shows how to find these actions by examining the extreme points of the BCE set, which we denote by $\text{ext}(BCE)$. To succinctly state the proposition, we borrow some terminology from Nau and McCardle (1990) and call an action $a_i \in A_i$ for player i **coherent** if that player takes that action in some BCE; that is, if some BCE p exists with $p(a_i) > 0$.

Proposition 5. *Fix any player i and two coherent actions $a_i, b_i \in A_i$. Then every BCE p with $a_i, b_i \in \text{supp}_i(p)$ has $p_{a_i} = p_{b_i}$ if and only if one of the following two conditions hold:*

(i) A $\mu \in \Delta(A_{-i} \times \Theta)$ exists such that for all $p \in \text{ext}(BCE)$ and $c_i \in \{a_i, b_i\} \cap \text{supp}_i(p)$, one has $p_{c_i} = \mu$.

(ii) A constant $\lambda > 0$ exists such that for all $p \in \text{ext}(BCE)$, $a_{-i} \in A_{-i}$, and $\theta \in \Theta$, one has $p(a_i, a_{-i}, \theta) = \lambda p(b_i, a_{-i}, \theta)$.

Thus, to know whether or not a pair of actions leads to the same beliefs in all BCEs, it is enough to check the extreme points of the BCE set for one of two properties. The first property states these actions always induce the same belief in all extreme BCEs. The second property requires the likelihood ratios of these two actions to be constant across all extreme BCEs.

Next, we build on Proposition 4 and obtain BKE is an all-or-nothing refinement of BCE:

Theorem 2. *The BKE set is either dense or nowhere dense in the BCE set.*

For an intuition, consider first the case in which a minimally mixed BKE exists. Then, by Proposition 4, the BKE set is dense in the BCE set. Consider now the case in which a minimally BKE does not exist. By Proposition 4, the BKE set is not dense in the BCE set. To reach the stronger conclusion that the BKE set is nowhere dense in the BCE set, we use the fact that the set of minimally mixed BCE is open and dense in the BCE set (see the formal proof in the appendix for more details).

The following result shows that the BKE set generically is dense in the BCE set. We adopt the following notion of genericity: We fix a finite set of players I, a finite set of payoff states Θ, a full-support prior $\pi \in \Delta(\Theta)$, and a finite set of actions A_i for each player $i \in I$. To specify a base game, it remains to specify a profile of utility functions $u = (u_i)_{i \in I}$. We identify u with an element of the Euclidean space $\mathbb{R}^{I \times A \times \Theta}$ and say that a statement is true for generic u if the closure of the subset in $\mathbb{R}^{I \times A \times \Theta}$ for which it is false has Lebesgue measure zero.

Theorem 3. *For generic u, the BKE set is dense in the BCE set.*

Thus, the environments in which costly information acquisition have different predictions than exogenous incomplete information are "special." An important caveat to this result is the notion of genericity we use: it is the most common for the static base games we study in this paper, but also the most permissive. For example, according to this notion of genericity, many important economic applications—such as auctions and oligopolistic competition, which we consider in Section 6—are non-generic. The notion of genericity also is not appropriate if the base game is not actually static but represents the strategic form of a primitive dynamic game. Extending BKE to dynamic games is an interesting direction for future research.

If there is only one player, then Theorem 3 easily follows from the fact that jeopardization coincides with weak dominance in single-agent settings:

Example 3. *Let I be a singleton. An action a jeopardizes an action b, with $a \neq b$, if and only if a mixed action $\alpha \in \Delta(A \setminus \{b\})$ exists that weakly dominates b, that is,*

$$\sum_a u(a, \theta)\alpha(a) \geq u(b, \theta)$$

for every $\theta \in \Theta$.[6] *For generic u, a mixed action $\alpha \in \Delta(A \setminus \{b\})$ exists that weakly dominates b if and only if a mixed action $\beta \in \Delta(A)$ exists that strictly dominates b, that is,*

$$\sum_a u(a,\theta)\beta(a) > u(b,\theta)$$

for every $\theta \in \Theta$.[7] *Thus, for generic u, either $J(b) = \{b\}$ or b is never played in a BCE; it follows from Proposition 8 that the BKE set is dense in the BCE set.*

The situation substantially changes when there are at least two players. If I is not a singleton, then jeopardization does not coincide with weak dominance. As noted before, in Matching Pennies, Heads and Tails jeopardize each other but neither action is weakly dominant. Moreover, this is is not a knife-edge property of Matching Pennies, since any nearby game has a unique correlated equilibrium that is a completely mixed Nash equilibrium.[8]

The next example describes a game in which no action is weakly dominated and yet the BKE set is nowhere dense in the BCE set.

Example 4. *Consider the following two-player game without payoff uncertainty:*

	a_2	b_2	c_2
a_1	$8,8$	$3,7$	$2,6$
b_1	$7,3$	$5,1$	$0,5$
c_1	$6,2$	$1,4$	$4,0$

No action is weakly dominated; in fact, every action is the unique best response to some belief about the opponent.

One can verify that the game has two Nash equilibria, (a_1, a_2) and $(\frac{1}{2}b_1 + \frac{1}{2}c_1, \frac{1}{2}b_2 + \frac{1}{2}c_2)$, and that the BCEs are their convex combinations: for $t \in [0,1]$,

$$p^t = t(a_1, a_2) + (1-t)\left(\frac{1}{2}b_1 + \frac{1}{2}c_1, \frac{1}{2}b_2 + \frac{1}{2}c_2\right).$$

The game has only two BKE, namely, the two Nash equilibria. Indeed, for $t \in (0,1)$ and $i \in I$, the action recommendations a_i and b_i (or c_i) induce distinct beliefs about the action of the opponent; yet, a_i is best response to the belief induced by b_i. Thus, for $t \in (0,1)$, p^t is not a BKE. This shows that the BKE set is nowhere dense in the BCE set.

[6] We expect the result to be known in the literature, but we did not manage to find a good reference; for the reader's convenience, we give a proof in the appendix (see Proposition 10 in Appendix A.5).

[7] We expect the result to be known in the literature, but we could not find a good reference; for the reader's convenience, we give a proof in the appendix (see Proposition 11 in Appendix A.5).

[8] For the case of no payoff uncertainty, Viossat (2008) shows that the set of games with a unique correlated equilibrium is open.

In sum, to prove Theorem 3, we had to rely on different ideas than the simple ones behind Example 3. As we explain in the appendix, Theorem 3 follows from combining two independent lemmas. The first lemma shows that one can perturb any game to make any BCE a BKE (different BCEs may require different perturbations). The second lemma shows that the closures of the BKE sets of generic nearby game also are nearby. To prove this result of generic continuity, we use the observation that the BKE set is semi-algebraic. This allows us to exploit tools from real algebraic geometry; in particular, we build on Blume and Zame (1994), who study the algebraic geometry of solution concepts.

6. Bertrand competition

In this section we discuss an applications to Bertrand competition where BKE is a significant refinement of BCE. In the application, players have continuous actions; we omit the definitions of BCE and BKE for continuous games: they are straightforward extensions of the definitions we gave for finite games.

Two profit-maximizing firms produce homogeneous goods and compete a la Bertrand. Each firm $i \in I := \{1, 2\}$ chooses a price $a_i \in A_i := \mathbb{R}_+$. The demand curve is $D : \mathbb{R}_+ \to \mathbb{R}_+$. To ease the exposition, we assume a specific functional form: $D(t) = \max\{1 - t, 0\}$ where $t \in \mathbb{R}_+$ is the market price. Our analysis extends to most common demand curves.

The marginal costs of production are uncertain and heterogeneous. To keep things simple, we assume that the marginal cost of firm 1 is either $\theta_1 = \underline{\theta}_1$ or $\theta_1 = \overline{\theta}_1$ with equal probability, and that the marginal cost of firm 2 is commonly know, denoted by θ_2. The firms' production costs are related as follows:

$$0 < \underline{\theta}_1 < \bar{\theta}_1 < \theta_2 < 1.$$

In particular, firm 1 has always a lower marginal cost. Overall, $\Theta = \{(\underline{\theta}_1, \theta_2), (\overline{\theta}_1, \theta_2)\}$.

All consumers buy from the firm that sets the lowest price. In the case of a tie, we assume that all consumers buy from the low-cost firm (firm 1). Our results extend to other standard tie-breaking rules (e.g., equal splitting of the market).

Next we compare the predictions of BCE and BKE. We focus on consumer surplus. Given an outcome $p \in \Delta_\pi(A \times \Theta)$, the expected **consumer surplus** is

$$CS(p) = \int_{A \times \Theta} \left[\int_{\min\{a_1, a_2\}}^{\infty} D(t)\, \mathrm{d}t \right] \mathrm{d}p(a, \theta).$$

The following result shows the largest consumer surplus that is attainable in a BCE is significantly larger than the largest consumer surplus that is attainable in a BKE. In addition, as long the cost of acquiring information is positive (but possibly small), the firms acquire no information: in every BKE, prices and marginal costs of production are independent of each other.

Proposition 6. *In the Bertrand competition game,*

$$\max_{p \in BCE} CS(p) = \frac{1}{2}\int_{\underline{\theta}_1}^{\infty} D(t)\,\mathrm{d}t + \frac{1}{2}\int_{\bar{\theta}_1}^{\infty} D(t)\,\mathrm{d}t > \int_{\frac{1}{2}\underline{\theta}_1 + \frac{1}{2}\bar{\theta}_1}^{\infty} D(t)\,\mathrm{d}t = \max_{p \in BKE} CS(p).$$

In addition, in every BKE the firm's prices are stochastically independent of each other and of the marginal cost of firm 1.

To get some intuition for the result, consider the game with exogenous information where firm 1's marginal cost of production is common knowledge. This game has an equilibrium where $a_1 = a_2 = \theta_1$ and all consumers buy from firm 1. In this equilibrium, the consumer surplus is

$$\frac{1}{2}\int_{\underline{\theta}_1}^{\infty} D(t)\,\mathrm{d}t + \frac{1}{2}\int_{\bar{\theta}_1}^{\infty} D(t)\,\mathrm{d}t,$$

which the proposition shows is the highest consumer surplus attainable in a BCE. Notice, however, that both firms have no incentive to to acquire the information prescribed by this equilibrium, since their profits are zero. Yet, both firms' actions are perfectly correlated with the state. Clearly, such is a situation is not sustainable if acquiring information is costly.

Proposition 6 generalizes this insight: in every BKE the firm's prices are independent of each other and of θ_1. As a result, the largest consumer surplus across BKE is

$$\int_{\frac{1}{2}\underline{\theta}_1 + \frac{1}{2}\bar{\theta}_1}^{\infty} D(t)\,\mathrm{d}t,$$

corresponding to a game with exogenous information where the firms' experiments are uninformative and the prices are equal to the expected marginal cost of firm 1.

The formal proof of Proposition 6 builds on the analysis of correlated equilibria in Bertrand competition by Jann and Schottmüller (2015). For completeness, the next proposition characterizes the smallest consumer surplus, which is the same in BCE and BKE.

Proposition 7. *In the Bertrand competition game,*

$$\min_{p \in BCE} CS(p) = \int_{\theta_2}^{\infty} D(t)\,\mathrm{d}t = \min_{p \in BKE} CS(p).$$

Since the set of BCE is convex, any level of consumer surplus between smallest and largest can be achieved:

$$\{CS(p) : p \in BCE\} = \left[\int_{\theta_2}^{\infty} D(t)\,\mathrm{d}t, \frac{1}{2}\int_{\underline{\theta}_1}^{\infty} D(t)\,\mathrm{d}t + \frac{1}{2}\int_{\bar{\theta}_1}^{\infty} D(t)\,\mathrm{d}t\right].$$

In general, the set of BKE is not convex (see Example 1 above). However, in this application, it is true that any level of consumer surplus between smallest and largest can be achieved; that is,

$$\{CS(p) : p \in BKE\} = \left[\int_{\theta_2}^{\infty} D(t)\,\mathrm{d}t, \int_{\frac{1}{2}\underline{\theta}_1+\frac{1}{2}\bar{\theta}_1}^{\infty} D(t)\,\mathrm{d}t\right].$$

Indeed, each point in this interval corresponds to a different equilibrium of a game with exogenous information where the firms' experiments are uninformative.

A. Proofs

A.1. Proof of Proposition 2

Proof of the "if" statement. Let (p, v) be an outcome-value pair such that p is a BCE and, for every $i \in I$, $v_i \in [\underline{v}_i(p), \bar{v}_i(p)]$. Since p is a BCE, by Proposition 1 there exist a set Z of non-fundamental states, a stochastic kernel $\zeta : \Theta \to \Delta(Z)$, a signal space X, a profile of experiments ξ, and a profile of action plans σ such that p is the outcome of (ξ, σ), and for every player i, σ_i maximizes

$$\sum_{\theta,z,x,a} u_i(a,\theta)\sigma_i'(a_i|x_i)\xi_i(x_i|\theta,z)\prod_{j\neq i}\sigma_j(a_j|x_j)\xi_j(x_j|\theta,z)\zeta(z|\theta)\pi(\theta) \tag{3}$$

over all $\sigma_i' \in \Sigma_i$.

For every player i, we define $\mathcal{E}_i = \{\xi_i, \xi_i'\}$ where ξ_i' is an uninformative experiment. We also define $C_i(\xi_i) = \bar{v}_i(p) - v_i$ and $C_i(\xi_i') = 0$. Note that $C_i(\xi_i) \geq 0$ because $v_i \in [\underline{v}_i(p), \bar{v}_i(p)]$ by Proposition 2-(ii). We obtain that

$$\begin{aligned}
&\sum_{\theta,z,x,a} u_i(a,\theta)\prod_{j}\sigma_j(a_j|x_j)\xi_j(x_j|\theta,z)\zeta(z|\theta)\pi(\theta) - C(\xi_i)\\
&= \max_{\sigma_i'}\sum_{\theta,z,x,a} u_i(a,\theta)\sigma_i'(a_i|x_i)\xi_i(x_i|\theta,z)\prod_{j\neq i}\sigma_j(a_j|x_j)\xi_j(x_j|\theta,z)\zeta(z|\theta)\pi(\theta) - C(\xi_i)\\
&= \bar{v}_i(p) - (\bar{v}_i(p) - v_i) = v_i\\
&\geq \underline{v}_i(p) = \max_{\sigma_i'}\sum_{\theta,z,x,a} u_i(a,\theta)\sigma_i'(a_i|x_i)\xi_i'(x_i|\theta,z)\prod_{j\neq i}\sigma_j(a_j|x_j)\xi_j(x_j|\theta,z)\zeta(z|\theta)\pi(\theta) - C(\xi_i')
\end{aligned}$$

where the first equality follows from (3) and the last inequality from $v_i \in [\underline{v}_i(p), \bar{v}_i(p)]$.

Thus, (ξ, σ) is an equilibrium of $(\mathcal{G}, \mathcal{T})$ with $\mathcal{T} := (Z, \zeta, (X_i, \mathcal{E}_i, C_i)_{i \in I})$; in addition, (p, v) is the outcome-value pair corresponding to (ξ, σ). ■

Proof of the 'only if" statement. Let (p, v) be the outcome-value pair of an equilibrium (ξ, σ) of an information acquisition game $(\mathcal{G}, \mathcal{T})$. For every player i, let $\mathcal{E}_i' = \{\xi_i\}$ and $C_i'(\xi_i) = 0$. The information technology $\mathcal{T}' := (Z, \zeta, (X_i, \mathcal{E}_i', C_i')_{i \in I})$ is degenerate. Since (ξ, σ) is an equilibrium of $(\mathcal{G}, \mathcal{T})$, it is also an equilibrium of $(\mathcal{G}, \mathcal{T}')$. It follows from Proposition 1 that p is a BCE.

For every player i, $C_i(\xi_i) \geq 0$, which implies that $v_i \leq \bar{v}_i(p)$. In addition, by hypothesis there exists an experiment ξ_i' such that $C_i(\xi_i') = 0$. Thus, since (ξ_i, σ_i) is a best response to (ξ_{-i}, σ_{-i}), we have that

$$v_i \geq \max_{\sigma_i'} \sum_{\theta, z, x, a} u_i(a, \theta) \sigma_i'(a_i | x_i) \xi_i(x_i | \theta, z) \prod_{j \neq i} \sigma_j(a_j | x_j) \xi_j(x_j | \theta, z) \zeta(z | \theta) \pi(\theta) \geq \underline{v}_i(p).$$

We conclude that $v_i \in [\bar{v}_i(p), \underline{v}_i(p)]$. ■

A.2. Proof of Theorem 1

The proof of the theorem proceeds as follows. First, we recall a single-agent lemma due Denti (2021). Next, with the help of the single-agent lemma, we prove the "if" and "only if" statements of the theorem.

A single-agent lemma. We take the perspective of a decision maker i who has to choose an action $a_i \in A_i$ whose utility $w_i(a_i, \omega)$ depends on an uncertain state of nature $\omega \in \Omega$. Both A_i and Ω are finite. Let $\rho \in \Delta(\Omega)$ be the prior distribution of the state; ρ may not have full support.

Before choosing an action, the decision maker can run an experiment $\xi_i : \Omega \to \Delta(X_i)$ at a cost $C_i(\xi_i) \in \mathbb{R}_+$. Let $\mathcal{E}_i$ be the set of feasible experiments. The signal space X_i is finite. Overall, the decision maker faces the following information acquisition problem:

$$\max_{\xi_i \in \mathcal{E}_i, \sigma_i \in \Sigma_i} \left[\sum_{\omega, x_i, a_i} w_i(a_i, \omega) \sigma_i(a_i | x_i) \xi_i(x_i | \omega) \rho(\omega) \right] - C_i(\xi_i) \tag{4}$$

where Σ_i is the set of all action plans $\sigma_i : X_i \to \Delta(A_i)$.

In accordance with the terminology used in the main text, $\xi_i \succsim \xi_i'$ if there is a garbling $g : X_i \to \Delta(X_i)$ such that for every $x_i \in X_i$ and $\omega \in \Omega$ with $\rho(\omega) > 0$,

$$\xi_i'(x_i | \omega) = \sum_{x_i' \in X_i} g(x_i | x_i') \xi(x_i' | \omega).$$

We also say that a pair $(\mathcal{E}_i, C_i)$ is **Blackwell monotone** if the following two conditions hold. First, if $\xi_i \succsim \xi_i'$ and $\xi_i \in \mathcal{E}_i$, then $\xi_i' \in \mathcal{E}_i$. And second, if $\xi_i, \xi_i' \in \mathcal{E}_i$ are such that $\xi_i \succsim \xi_i'$ (resp., $\xi_i \succ \xi_i'$), then $C_i(\xi_i) \geq C_i(\xi_i')$ (resp., $C_i(\xi_i) > C_i(\xi_i')$).

Next we characterize the pairs (ξ_i, σ_i) that are optimal solutions of (4) for some Blackwell monotone $(\mathcal{E}_i, C_i)$. To state the result, we need some auxiliary notation. Let

$$\operatorname{supp}(\xi_i) = \left\{ x_i : \sum_{\omega} \xi_i(x_i|\omega)\rho(\omega) > 0 \right\} \tag{5}$$

be the set of signals that have positive probability. For every $x_i \in \operatorname{supp}(\xi_i)$, we denote by $\mu_{x_i} \in \Delta(\Omega)$ the posterior distribution of the state:

$$\mu_{x_i}(\omega) = \frac{\xi_i(x_i|\omega)\rho(\omega)}{\sum_{\omega'} \xi_i(x_i|\omega')\rho(\omega')}.$$

Finally, let

$$BR(\mu_{x_i}) := \arg\max_{a_i \in A_i} \sum_{\omega} w_i(a_i, \omega)\mu_{x_i}(\omega)$$

be the set of best responses to μ_{x_i}.

Lemma 1 (Denti, 2021)**.** *A Blackwell monotone $(\mathcal{E}_i, C_i)$ exists such that the pair (ξ_i, σ_i) is an optimal solution of (4) if and only if the following conditions hold:*

(i) For all $x_i \in \operatorname{supp}(\xi_i)$,

$$\sigma_i(BR(\mu_{x_i})|x_i) = 1.$$

(ii) For all $x_i, x_i' \in \operatorname{supp}(\xi_i)$,

$$\mu_{x_i} \neq \mu_{x_i'} \quad \textit{implies} \quad BR(\mu_{x_i}) \cap BR(\mu_{x_i'}) = \varnothing.$$

In addition, C_i can be chosen such that

$$C_i(\xi_i) = \sum_{\omega, x_i, a_i} w_i(a_i, \omega)\sigma_i(a_i|x_i)\xi_i(x_i|\omega)\rho(\omega) - \max_{a_i \in A_i} \sum_{\omega} w_i(a_i, \omega)\rho(\omega).$$

Comparing this paper and Denti (2021), the reader should keep in mind that Denti (2021) uses a stronger version of Blackwell's informativeness: $\xi_i \succsim^* \xi_i'$ if there is a garbling $g : X_i \to \Delta(X_i)$ such that for every $x_i \in X_i$ and $\omega \in \Omega$,

$$\xi_i'(x_i|\omega) = \sum_{x_i' \in X_i} g(x_i|x_i')\xi(x_i'|\omega). \tag{6}$$

The difference between $\succsim$ and $\succsim^*$ is that (5) holds on the support of ρ while (6) on Ω.

Proof of the "if" statement of Theorem 1. Let (p, v) be an outcome-value pair that satisfies the conditions (i) and (ii) of Theorem 1. First we consider the case in which $v = \underline{v}(p)$.

Lemma 2. *For every BKE p, there is a Blackwell monotone $\mathcal{T}$ and an equilibrium (ξ, σ) of $(\mathcal{G}, \mathcal{T})$ whose outcome-value pair is $(p, \underline{v}(p))$.*

Proof of the lemma. For every player i and pair of actions $a_i, b_i \in \operatorname{supp}_i(p)$, let $a_i \sim_i b_i$ be $p_{a_i} = p_{b_i}$. Note that $\sim_i$ is an equivalence relation on $\operatorname{supp}_i(p)$. Let Z_i be the corresponding set of equivalence classes. Take $Z = \prod_{i \in I} Z_i$ and define $\zeta : \Theta \to Z$ by

$$\zeta(z|\theta) = \frac{p(\{(a, \theta) : a_i \in z_i \text{ for all } i \in I\})}{\pi(\theta)}.$$

We also take $X_i = Z_i$ and define $\xi_i : \Theta \times Z \to \Delta(X_i)$ by

$$\xi_i(x_i|\theta, z) = \begin{cases} 1 & \text{if } x_i = z_i, \\ 0 & \text{otherwise.} \end{cases}$$

Finally, we define $\sigma_i : X_i \to \Delta(A_i)$ by

$$\sigma_i(a_i|x_i) = \begin{cases} \frac{p(a_i)}{\sum_{b_i \in x_i} p(b_i)} & \text{if } a_i \in x_i, \\ 0 & \text{otherwise.} \end{cases}$$

We observe that p is the outcome induced (ξ, σ), that is, for all $a \in A$ and $\theta \in \Theta$,

$$p(a, \theta) = \sum_z \prod_{i \in I} \sigma_i(a_i|z_i)\zeta(z|\theta)\pi(\theta). \tag{7}$$

To obtain (7), first observe that, for all $i \in I$, $a \in A$, and $\theta \in \Theta$, Bayes rule implies that

$$p(a, \theta) = p(a_i)p_{a_i}(a_{-i}, \theta).$$

Now take $z_i \in Z_i$ such that $a_i \in z_i$. For all $b_i \in z_i$, $p_{b_i} = p_{a_i}$. Thus,

$$p(a, \theta) = p(a_i) \sum_{b_i \in z_i} p_{b_i}(a_{-i}, \theta)\sigma_i(b_i|z_i).$$

It follows that

$$\begin{aligned}
p(a,\theta) &= \frac{p(a_i)}{\sum_{c_i \in z_i} p(c_i)} \sum_{b_i \in z_i} p_{b_i}(a_{-i},\theta) \left(\sigma_i(b_i|z_i) \sum_{c_i \in z_i} p(c_i) \right) \\
&= \sigma_i(a_i|z_i) \sum_{b_i \in z_i} p_{b_i}(a_{-i},\theta) p(b_i) \\
&= \sigma_i(a_i|z_i) \sum_{b_i \in z_i} p(b_i, a_{-i}, \theta). \qquad (8)
\end{aligned}$$

Applying (8) iteratively, we obtain that

$$p(a,\theta) = \prod_{i \in I} \sigma_i(a_i|z_i) \zeta(z|\theta) \pi(\theta)$$

for all $a \in A$ and $\theta \in \Theta$ such that, for all $i \in I$, $a_i \in z_i$. Since $\sigma_i(a_i|z_i) = 0$ whenever $a_i \notin z_i$, we conclude that (7) holds.

Next we use Lemma 1 to construct $\mathcal{E}_i$ and C_i. Let $\Omega = \Theta \times Z$; we define $\rho \in \Delta(\Omega)$ and $w_i : A_i \times \Omega \to \mathbb{R}$ by

$$\begin{aligned}
\rho(\theta, z) &= \pi(\theta)\zeta(z|\theta), \\
w_i(a_i, \theta, z) &= \sum_{a_{-i}} u_i(a,\theta) \prod_{j \neq i} \sigma_j(a_j|z_j).
\end{aligned}$$

We note that for all $z_i \in Z_i$ and $a_i \in z_i$, $BR(p_{a_i}) = BR(\mu_{z_i})$. Indeed, it follows from (7) that for all $b_i \in A_i$,

$$\sum_{\omega} w_i(b_i, \omega) \mu_{z_i}(\omega) = \sum_{a_{-i},\theta} u_i(b_i, a_{-i}, \theta) p_{a_i}(a_{-i}, \theta).$$

To verify Lemma 1-(i), take $z_i \in Z_i$ and $a_i \in A_i$ such that $\sigma_i(a_i|z) > 0$. By the construction of the action plan, $a_i \in z_i$. By the obedience constraint for p, $a_i \in BR(p_{a_i})$. As noted above, $BR(p_{a_i}) = BR(\mu_{z_i})$. Therefore, $a_i \in BR(\mu_{z_i})$. We deduce that Lemma 1-(i) holds.

To verify Lemma 1-(ii), take $z_i, z'_i \in Z_i$ such that $\mu_{z_i} \neq \mu_{z'_i}$. Let $a_i \in z_i$ and $b_i \in z'_i$. Since $\mu_{z_i} \neq \mu_{z'_i}$, we must have $z_i \neq z'_i$ and, therefore, $p_{a_i} \neq p_{b_i}$. By the separation constraint for p, $BR(p_{a_i}) \cap BR(p_{b_i}) = \varnothing$. As noted above, $BR(p_{a_i}) = BR(\mu_{z_i})$ and $BR(p_{b_i}) = BR(\mu_{z'_i})$. We deduce $BR(\mu_{z_i}) = BR(\mu_{z'_i}) = \varnothing$: Lemma 1-(ii) holds.

In sum, for every player i, we can invoke Lemma 1 to find a Blackwell monotone $(\mathcal{E}_i, C_i)$ such that (ξ_i, σ_i) is a best reply to (ξ_{-i}, σ_{-i}). This shows that $\mathcal{T} = (Z, \zeta, (X_i, \mathcal{E}_i, C_i)_{i \in I})$ is a Blackwell-monotone information technology such that (ξ, σ) is an equilibrium $(\mathcal{G}, \mathcal{T})$ with outcome p. In addition, as in the last part of Lemma 1, we can choose costs so that, for every player i, $C_i(\xi_i) = \bar{v}_i(p) - \underline{v}_i(p)$, that is, $\bar{v}_i(p) - C_i(\xi_i) = \underline{v}_i(p)$. $\square$

Now we consider the general case in which v may differ from $\underline{v}(p)$. Let $\mathcal{T}$ and (ξ, σ) as in Lemma 2. Without loss of generality, we assume that for every player i,

$$\mathcal{E}_i = \{\xi_i' : \xi_i \succsim \xi_i'\}.$$

For every $\lambda_i \in (0, 1]$, we define $C_{\lambda_i} : \mathcal{E}_i \to \mathbb{R}_+$ by

$$C_{\lambda_i}(\xi_i') = \lambda_i C_i(\xi_i').$$

For a vector $\lambda = (\lambda_i)_{i \in I}$, we define the information technology $\mathcal{T}_\lambda = (Z, \zeta, (X_i, \mathcal{E}_i, C_{\lambda_i})_{i\in I})$. Since $\mathcal{T}$ is Blackwell monotone, $\mathcal{T}_\lambda$ is Blackwell monotone. Since (ξ, σ) is an equilibrium $(\mathcal{G}, \mathcal{T})$ whose outcome is p, (ξ, σ) is an equilibrium $(\mathcal{G}, \mathcal{T}_\lambda)$ whose outcome is p. In addition,

$$\bar{v}_i(p) - C_{\lambda_i}(\xi_i) = (1 - \lambda_i)\underline{v}_i(p) + \lambda_i \bar{v}_i(p).$$

Thus, by appropriately choosing the vector λ, we can be sure that v is the value of (ξ, σ).$\blacksquare$

Proof of the "only if" statement of Theorem 1. Let (p, v) be the outcome-value pair induced by an equilibrium (ξ, σ) of an information acquisition game $(\mathcal{G}, \mathcal{T})$ where $\mathcal{T}$ is Blackwell monotone. We organize the proof of the "only if" statement of Theorem 1 in two lemmas:

Lemma 3. *The equilibrium outcome p is a BKE.*

Proof of the lemma. By Proposition 2, p is a BCE. To verify the separation constraint, we begin by fixing a player i. By the definition of equilibrium, player i's strategy (ξ_i, σ_i) maximizes

$$\left[\sum_{\theta,z,x,a} u_i(a, \theta)\sigma_i'(a_i|x_i)\xi_i'(x_i|\theta, z)\prod_{j\neq i}\sigma_j(a_j|x_j)\xi_j(x_j|\theta, z)\zeta(z|\theta)\pi(\theta)\right] - C_i(\xi_i)$$

over all $\xi_i' \in \mathcal{E}_i$ and $\sigma_i' \in \Sigma_i$. Since $\mathcal{T}$ is Blackwell monotone, we can apply Lemma 1 with $\Omega = \Theta \times Z$, and $\rho \in \Delta(\Omega)$ and $w_i : A_i \times \Omega \to \mathbb{R}$ given by

$$\begin{aligned} \rho(\theta, z) &= \pi(\theta)\zeta(z|\theta), \\ w_i(a_i, \theta, z) &= \sum_{x_{-i},a_{-i}} u_i(a, \theta)\prod_{j\neq i}\sigma_j(a_j|x_j)\xi_j(x_j|\theta, z). \end{aligned} \tag{9}$$

Thus, for all $x_i \in \text{supp}(\xi_i)$, $\sigma(BR(\mu_{x_i})|\mu_{x_i}) = 1$ (Lemma 1-(i)). Moreover, for all $x_i, x_i' \in \text{supp}(\xi_i)$, if $\mu_{x_i} \neq \mu_{x_i'}$ then $BR(\mu_{x_i}) \neq BR(\mu_{x_i'})$ (Lemma 1-(ii)).

Let $a_i, b_i \in \text{supp}_i(p)$ such that $p_{a_i} \neq p_{b_i}$. Simple algebra shows that, for every $a_{-i} \in A_{-i}$ and $\theta \in \Theta$,

$$p_{a_i}(a_{-i}, \theta) = \sum_{z, x_{-i}} \left(\prod_{j \neq i} \sigma_j(a_j | x_j) \xi_j(x_j | \theta, z) \right) \sum_{x_i} \mu_{x_i}(\theta, z) \xi_i(x_i | a_i) \tag{10}$$

where, with a slight abuse of notation, we denote by $\xi_i(x_i|a_i)$ the probability that player i's signal is x_i given that her action is a_i; i.e.,

$$\xi_i(x_i | a_i) = \frac{\sigma_i(a_i|x_i) \sum_\omega \xi_i(x_i|\omega)\rho(\omega)}{\sum_{x_i'} \sigma_i(a_i|x_i') \sum_\omega \xi_i(x_i'|\omega)\rho(\omega)}.$$

Similar expressions hold for $p_{b_i}(a_{-i}, \theta)$ and $\xi_i(x_i|b_i)$. Thus, since $p_{a_i} \neq p_{b_i}$, there must be a pair of signals $x_{a_i}, x_{b_i} \in \text{supp}(\xi_i)$ such that $\sigma_i(a_i|x_{a_i}) > 0$, $\sigma_i(b_i|x_{b_i}) > 0$, and $\mu_{x_{a_i}} \neq \mu_{x_{b_i}}$. By Lemma 1-(ii),

$$BR(\mu_{x_{a_i}}) \cap BR(\mu_{x_{b_i}}) = \varnothing. \tag{11}$$

We claim that $BR(p_{a_i}) \subseteq BR(\mu_{x_{a_i}})$. Take $c_i \in BR(p_{a_i})$. By Lemma 1-(i), for all x_i such that $\sigma(a_i|x_i) > 0$,

$$\sum_\omega w_i(a_i, \omega)\mu_{x_i}(\omega) \geq \sum_\omega w_i(c_i, \omega)\mu_{x_i}(\omega). \tag{12}$$

Using (9) and (10), we note that

$$\sum_{x_i} \xi_i(x_i|a_i) \left(\sum_\omega w_i(a_i, \omega)\mu_{x_i}(\omega) \right) = \sum_{a_{-i}, \theta} u_i(a_i, a_{-i}, \theta) p_{a_i}(a_{-i}, \theta),$$
$$\sum_{x_i} \xi_i(x_i|a_i) \left(\sum_\omega w_i(c_i, \omega)\mu_{x_i}(\omega) \right) = \sum_{a_{-i}, \theta} u_i(c_i, a_{-i}, \theta) p_{a_i}(a_{-i}, \theta).$$

As a consequence, since $c_i \in BR(p_{a_i})$, we have that

$$\sum_{x_i} \xi_i(x_i|a_i) \left(\sum_\omega w_i(a_i, \omega)\mu_{x_i}(\omega) \right) \leq \sum_{x_i} \xi_i(x_i|a_i) \left(\sum_\omega w_i(c_i, \omega)\mu_{x_i}(\omega) \right).$$

It follows that (12) must hold with equality. Therefore, in particular,

$$\sum_\omega w_i(a_i, \omega)\mu_{x_{a_i}}(\omega) = \sum_\omega w_i(c_i, \omega)\mu_{x_{a_i}}(\omega).$$

Since $a_i \in BR(\mu_{x_{a_i}})$, this shows that $c_i \in BR(\mu_{x_{a_i}})$. Overall, we conclude that $BR(p_{a_i}) \subseteq BR(\mu_{x_{a_i}})$.

By a similar argument, $BR(p_{b_i}) \subseteq BR(\mu_{x_{b_i}})$. Hence,

$$BR(p_{a_i}) \cap BR(p_{b_i}) \subseteq BR(\mu_{x_{a_i}}) \cap BR(\mu_{x_{b_i}}).$$

By (11), $BR(\mu_{x_{a_i}}) \cap BR(\mu_{x_{b_i}}) = \varnothing$, which implies that $BR(p_{a_i}) \cap BR(p_{b_i}) = \varnothing$. This shows that the separation constraint holds: we deduce that p is a BKE. □

Lemma 4. *For every player i,*

$$v_i = \underline{v}_i(p) = \bar{v}_i(p) \quad \textit{or} \quad v_i \in [\underline{v}_i(p), \bar{v}_i(p)).$$

Proof of the lemma. By Proposition 2, $v_i \in [\underline{v}_i(p), \bar{v}_i(p)]$. If $\underline{v}_i(p) = \bar{v}_i(p)$, then

$$v_i = \underline{v}_i(p) = \bar{v}_i(p). \tag{13}$$

Suppose instead that $\underline{v}_i(p) < \bar{v}_i(p)$. If

$$\sum_{\theta,z,x,a} u_i(a,\theta)\sigma_i(a_i|x_i)\xi_i(x_i|\theta,z) \prod_{j\neq i} \sigma_j(a_j|x_j)\xi_j(x_j|\theta,z)\zeta(z|\theta)\pi(\theta) < \bar{v}(p),$$

then obviously $v_i < \bar{v}(p)$ because $C_i(\xi_i) \geq 0$. If, on the other hand,

$$\sum_{\theta,z,x,a} u_i(a,\theta)\sigma_i(a_i|x_i)\xi_i(x_i|\theta,z) \prod_{j\neq i} \sigma_j(a_j|x_j)\xi_j(x_j|\theta,z)\zeta(z|\theta)\pi(\theta) = \bar{v}(p),$$

then ξ_i cannot be uninformative because $\underline{v}_i(p) < \bar{v}_i(p)$. By Blackwell monotonicity, $C_i(\xi_i) > 0$, which implies that $v_i < \bar{v}(p)$. In sum, if $\underline{v}_i(p) < \bar{v}_i(p)$, then $v_i \in [\underline{v}_i(p), \bar{v}_i(p))$. □

By Lemmas 3 and 4, the conditions (i) and (ii) of Theorem 1 hold. ■.

A.3. Proofs of Proposition 4 and Theorem 2

In this section we prove generalizations of Proposition 4 and Theorem 2 that apply locally to a closed convex set of BCEs.

With a slight abuse of notation, we denote by BCE the set of all BCEs and by BKE the set of all BKEs. Let $B \subseteq BCE$ be a non-empty closed convex set. For a player i, an action a_i B-**jeopardizes** an action b_i if, for every $p \in B$ with $b \in \text{supp}_i(p)$, $a_i \in BR(p_{b_i})$. We denote by $J_B(b_i)$ the set of actions that B-jeopardizes b_i. Just like the standard jeopardization concept, one has $b_i \in J_B(b_i)$ for all b_i.

An outcome $p \in B$ has B-**maximal support** if the support of every other $q \in B$ is contained by the support of p. An outcome $p \in B$ is B-**minimally mixed** if it has

B-maximal support and

$$q_{a_i} \neq q_{b_i} \quad \text{implies} \quad p_{a_i} \neq p_{b_i}.$$

if for every $q \in B$, $i \in I$, and $a_i, b_i \in \text{supp}_i(p)$,

We are now ready to state the local versions of Proposition 4 and Theorem 2 that we prove in this section.

Proposition 8. *The following statements are equivalent:*

(i) The set $BKE \cap B$ is dense in B.

(ii) A B-minimally mixed BKE exists.

(iii) For all $p \in B$, $i \in I$, $a_i, b_i \in \text{supp}_i(p)$,

$$p_{a_i} \neq p_{b_i} \quad \textit{implies} \quad J_B(a_i) \cap J_B(b_i) = \varnothing.$$

Theorem 4. *The set $BKE \cap B$ is either dense or nowhere dense in B.*

As a first step, we prove a basic lemma about best responses. In what follows, for $p \in \Delta(A \times \Theta)$, $a_i \in \text{supp}_i(p)$, and $b_i \in A_i$, take $u_i(b_i, p_{a_i}) \in \mathbb{R}$ to be

$$u_i(b_i, p_{a_i}) = \sum_{a_{-i}, \theta} u_i(b_i, a_{-i}, \theta) p_{a_i}(a_{-i}, \theta).$$

Lemma 5. *For every $t \in (0,1)$, $p, q \in BCE$, $i \in I$, and $a_i \in \text{supp}_i(p)$,*

$$BR\left((tp + (1-t)q)_{a_i}\right) \subseteq BR\left(p_{a_i}\right).$$

Proof. Take $b_i \in BR\left((tp + (1-t)q)_{a_i}\right)$. If $a_i \notin \text{supp}_i(q)$, then

$$(tp + (1-t)q)_{a_i} = p_{a_i},$$

which immediately implies the desired result.

Suppose now that $a_i \in \text{supp}_i(q)$. Since $p, q \in BCE$, we have

$$u_i(a_i, p_{a_i}) \geq u_i(b_i, p_{a_i}) \quad \text{and} \quad u_i(a_i, q_{a_i}) \geq u_i(b_i, q_{a_i}).$$

Simple algebra shows that there exists $s \in (0,1)$ such that

$$(tp + (1-t)q)_{a_i} = sp_{a_i} + (1-s)q_{a_i}.$$

Since $b_i \in BR\left((tp + (1-t)q)_{a_i}\right)$, we obtain that

$$\begin{aligned} su_i(b_i, p_{a_i}) + (1-s)u_i(b_i, q_{a_i}) &= u_i(b_i, sp_{a_i} + (1-s)q_{a_i}) \\ &\geq u_i(a_i, sp_{a_i} + (1-s)q_{a_i}) \\ &= su_i(a_i, p_{a_i}) + (1-s)u_i(a_i, q_{a_i}). \end{aligned}$$

We conclude that $u_i(a_i, p_{a_i}) = u_i(b_i, p_{a_i})$ and $u_i(a_i, q_{a_i}) = u_i(b_i, q_{a_i})$. It follows from $p \in BCE$ that $b_i \in BR\left(p_{a_i}\right)$. □

Next, we show that taking convex combinations of BCEs usually preserve the set of action recommendations that lead to different beliefs.

Lemma 6. *For every $p, q \in \Delta(A \times \Theta)$, $i \in I$, and $a_i, b_i \in \mathrm{supp}_i(p)$ with $p_{a_i} \neq p_{b_i}$, there are at most two $t \in (0,1)$ such that*

$$(tp + (1-t)q)_{a_i} = (tp + (1-t)q)_{b_i}. \tag{14}$$

Proof. Note that $t \in (0,1)$ is a solution of (14) if and only if for every $a_{-i} \in A_{-i}$ and $\theta \in \Theta$,

$$\begin{aligned} &(tp(a_i, a_{-i}, \theta) + (1-t)q(a_i, a_{-i}, \theta))\,(tp(b_i) + (1-t)q(b_i)) \\ =&(tp(b_i, a_{-i}, \theta) + (1-t)q(b_i, a_{-i}, \theta))\,(tp(a_i) + (1-t)q(a_i))\,. \end{aligned} \tag{15}$$

Each equation (15) is polynomial in t, with degree at most two. Since $p_{a_i} \neq p_{b_i}$, at least one such polynomial equation does not have degree zero and, therefore, has at most two solutions. We deduce that (14) has at most two solutions for $t \in (0,1)$. □

Our next goal is to show that B-minimally mixed BCEs are the norm rather than the exception. As an intermediate step, we first show the set of B-minimally mixed BCEs is non-empty.

Lemma 7. *A B-minimally mixed BCE exists.*

Proof. For every $p \in B$, define the set

$$X(p) = \bigcup_i \left\{(a_i, b_i) : a_i, b_i \in \mathrm{supp}_i(p) \text{ and } p_{a_i} \neq p_{b_i}\right\}.$$

Note that $p \in B$ is B-minimally mixed if and only if it has B-maximal support and for every $q \in B$, $X(q) \subseteq X(p)$.

Since the set $A \times \Theta$ is finite and the set B is convex, we can find a B-maximal support $p \in B$ such that for every B-maximal-support $q \in B$, the cardinality of $X(p)$ is larger than the cardinality of $X(q)$.

We now show that p is B-minimally mixed. Fix an arbitrary $q \in B$. For every $t \in (0,1)$, define $p^t = tp + (1-t)q$, which belongs to B because B is convex. Since p has B-maximal support, the same is true for p^t. Thus, the cardinality of $X(p)$ is larger than the cardinality of $X(p^t)$. By Lemma 6, we can find $t \in (0,1)$ such that $X(p) \subseteq X(p^t)$ and $X(q) \subseteq X(p^t)$. This shows that $X(q) \subseteq X(p)$; otherwise, the cardinality of $X(p^t)$ would be strictly larger than the cardinality of $X(p)$. We conclude that p is B-minimally mixed. $\square$

We now show that the B-minimally mixed BCEs includes most of the BCEs in B in a precise sense.

Lemma 8. *The set of B-minimally mixed BCEs is open and dense in B.*

Proof. Let B_M denote the set of B-minimally mixed BCEs. We first argue that B_M is open in B. Towards this goal, note the following sets are open in B for every $i \in I$ and $a_i, b_i \in A_i$:

$$\{p \in B : p(a_i) > 0\}, \quad \text{and} \quad \{p \in B : p(a_i)p(b_i) > 0 \text{ and } p_{a_i} \neq p_{b_i}\}.$$

Since A is finite, we obtain that B_M equals the intersection of a finite number of open subsets of B_M. It follows B_M is open in B.

To see B_M is dense in B, fix some $q \in B$. Take p to be a B-minimally mixed BCE, which exists by Lemma 7. For every $t \in (0,1)$, define $p^t = tp + (1-t)q$. Because p has B-maximal support, the same is true for p^t for all $t \in (0,1)$. Moreover, by Lemma 6, a finite set $T \subseteq (0,1)$ exists such that for all $t \in (0,1) \setminus T$, $i \in I$, and $a_i, b_i \in \text{supp}_i(p)$,

$$p_{a_i} \neq p_{b_i} \quad \text{implies} \quad p^t_{a_i} \neq p^t_{b_i}.$$

Thus, p^t is a B-minimally mixed BCE for all $t \in (0,1) \setminus T$. Thus, q is a limit point of $\{p^t : t \in (0,1) \setminus T\}$, which implies it is a limit point of B_M. $\square$

We are now ready to prove Proposition 8 and Theorem 4.

Proof of Proposition 8. That (i) implies (ii) follows from Lemma 8.

We now show (ii) implies (iii). Let q be a B-minimally mixed BKE. Fix any $p \in B$, $i \in I$ and $a_i, b_i \in \text{supp}_i(p)$ such that $p_{a_i} \neq p_{b_i}$. Since q is B-minimally mixed, $a_i, b_i \in \text{supp}_i(q)$ (because q has B-maximal support) and $q_{a_i} \neq q_{b_i}$. Thus,

$$\varnothing = BR(q_{a_i}) \cap BR(q_{b_i}) \supseteq J_B(a_i) \cap J_B(b_i),$$

where first we use the separation constraint, and then the fact that $J_B(c_i) = \cap_{\tilde{p} \in B} BR(\tilde{p}_{c_i})$ for all $c_i \in A_i$. We conclude (ii) implies (iii).

Finally, we argue (iii) implies (i). Fix any $p \in B$. Because A is finite and B is convex, it follows from Lemma 5 that we can find $q \in B$ such that q has B-maximal support and

$$BR(q_{a_i}) = J_B(q_{a_i}) \tag{16}$$

for all $i \in I$ and $a_i \in \text{supp}_i(q)$.

For $t \in (0,1)$, let $p^t = tp + (1-t)q$. We claim that $p^t \in BKE \cap B$. That $p^t \in B$ follows from B being convex. To see p^t is a BKE, take any $i \in I$ and $a_i, b_i \in \text{supp}_i(p^t)$ such that $p^t_{a_i} \neq p^t_{b_i}$. Since q has maximal support, $a_i, b_i \in \text{supp}_i(q)$. Then,

$$BR(p^t_{a_i}) \cap BR(p^t_{b_i}) \subseteq BR(q_{a_i}) \cap BR(q_{b_i}) = J_B(a_i) \cap J_B(b_i) = \varnothing,$$

where first we use Lemma 5, then (16), and finally Proposition 8-(iii). We conclude $p^t \in BKE \cap B$ for all $t \in (0,1)$. Proposition 8-(i) then follows from $p = \lim_{t \to 1} p^t$. □

Proof of Theorem 4. It is enough to prove that if $BKE \cap B$ is not nowhere dense in B, then it is dense in B. Suppose $BKE \cap B$ is dense in some non-empty set $\tilde{B} \subseteq B$ that is open in B. Let B_M the set of $p \in B$ that are B-minimally mixed. Note $\tilde{B} \cap B_M$ is open (in B) and non-empty by Lemma 8. Because $BKE \cap B$ is dense in $\tilde{B}$, we obtain that $(BKE \cap B) \cap (\tilde{B} \cap B_{MM})$ is non-empty. Thus, we have found a B-minimally mixed BKE. That $BKE \cap B$ is dense in B then follows from Proposition 8. □

A.4. Proof of Proposition 5

We prove here a slightly more general version of Proposition 5 that applies locally to any closed convex set of outcomes.[9] Thus, fix a closed convex set $B \subseteq \Delta(A \times \Theta)$. For a player i, say an action a_i is B-**coherent** if a $p \in B$ exists with $p(a_i) > 0$. Let $\mathbf{0}$ be the all-zeros vector in $\mathbb{R}^{A_{-i} \times \Theta}$; in what follows, we use the convention that $p_{a_i} = \mathbf{0}$ for every $p \in \Delta(A \times \Theta)$ and $a_i \in A_i$ such that $p(a_i) = 0$. As in the previous section, we say that an outcome $p \in B$ has B-**maximal support** if the support of every other $q \in B$ is contained by the support of p.

Proposition 9. *Fix a player i and two B-coherent actions $a_i, b_i \in A_i$. Then every $p \in B$ with $a_i, b_i \in \text{supp}_i(p)$ has $p_{a_i} = p_{b_i}$ if and only if one of the following two conditions hold:*

(i) A $\mu \in \Delta(A_{-i} \times \Theta)$ exists such that for all $p \in \text{ext}(B)$, $\{p_{a_i}, p_{b_i}\} \subseteq \{\mu, \mathbf{0}\}$.

(ii) A constant $\lambda > 0$ exists such that for all $p \in \text{ext}(B)$, $p(a_i)p_{a_i} = \lambda p(b_i)p_{b_i}$.

To prove the proposition, we need the following lemma.

[9] Neither the obedience nor the separation constraint play any role in this section.

Lemma 9. *Fix a player i and two actions $a_i, b_i \in A_i$. Let $p, q \in \Delta(A \times \Theta)$ such that $\{a_i, b_i\} \subseteq \mathrm{supp}_i(p) \cup \mathrm{supp}_i(q)$. Suppose $r_{a_i} = r_{b_i}$ for all $r \in \{p, q\}$ with $\{a_i, b_i\} \subseteq \mathrm{supp}_i(r)$. If $(tp + (1-t)q)_{a_i} = (tp + (1-t)q)_{b_i}$ for some $t \in (0,1)$, then one of the following two conditions hold:*

(i) A $\mu \in \Delta(A_{-i} \times \Theta)$ exists such that for all $r \in \{p, q\}$, $\{r_{a_i}, r_{b_i}\} \subseteq \{\mu, \boldsymbol{0}\}$.

(ii) A constant $\lambda > 0$ exists such that for all $r \in \{p, q\}$, $r(a_i) = \lambda r(b_i)$.

Proof. Let $p^t := tp + (1-t)q$. We proceed by contradiction: we assume that Lemma 9-(i) and Lemma 9-(ii) both fail and show that $p^t_{a_i} \neq p^t_{b_i}$.

We begin by noting that one can rewrite the condition that $r_{a_i} = r_{b_i}$ for all $r \in \{p, q\}$ with $\{a_i, b_i\} \subseteq \mathrm{supp}_i(r)$ as

$$r(a_i)r(b_i)r_{a_i} = r(a_i)r(b_i)r_{b_i} \text{ for all } r \in \{p, q\}. \tag{17}$$

Because $\mathrm{supp}_i(p^t) = \mathrm{supp}_i(p) \cup \mathrm{supp}_i(q)$ and $\{a_i, b_i\} \subseteq \mathrm{supp}_i(p) \cup \mathrm{supp}_i(q)$, we have $\{a_i, b_i\} \subseteq \mathrm{supp}_i(p^t)$. Thus, applying Bayes rule, we obtain that $p^t_{a_i} = p^t_{b_i}$ if and only if for every $a_{-i} \in A_{-i}$ and $\theta \in \Theta$, one has

$$p^t(a_i)p^t(b_i, a_{-i}, \theta) - p^t(b_i)p^t(a_i, a_{-i}, \theta) = 0.$$

Expanding the right hand side of the above equation by substituting in the definition of p^t, rearranging terms as a polynomial in t, and using (17), delivers that the above display equation is equivalent to

$$(t - t^2)\Big[p(a_i)q(b_i, a_{-i}, \theta) + q(a_i)p(b_i, a_{-i}, \theta) - q(b_i)p(a_i, a_{-i}, \theta) - p(b_i)q(a_i, a_{-i}, \theta)\Big] = 0.$$

Since $t \in (0,1)$, we get that $p^t_{a_i} = p^t_{b_i}$ if and only if for every $a_{-i} \in A_{-i}$ and $\theta \in \Theta$, one has

$$p(a_i)q(b_i, a_{-i}, \theta) + q(a_i)p(b_i, a_{-i}, \theta) - q(b_i)p(a_i, a_{-i}, \theta) - p(b_i)q(a_i, a_{-i}, \theta) = 0.$$

Writing the above in vector notation delivers that $p^t_{a_i} = p^t_{b_i}$ is equivalent to

$$p(a_i)q(b_i)q_{b_i} + q(a_i)p(b_i)p_{b_i} - q(b_i)p(a_i)p_{a_i} - p(b_i)q(a_i)q_{a_i} = \mathbf{0}. \tag{18}$$

We now divide the proof into cases. Consider first the case in which $\{a_i, b_i\} \subseteq \mathrm{supp}_i(p) \cap \mathrm{supp}_i(q)$. In this case, (17) implies $p_{a_i} = p_{b_i}$ and $q_{a_i} = q_{b_i}$, and so we get that

$$\begin{aligned} &p(a_i)q(b_i)q_{b_i} + q(a_i)p(b_i)p_{b_i} - q(b_i)p(a_i)p_{a_i} - p(b_i)q(a_i)q_{a_i} = \\ &\qquad\qquad = (p(a_i)q(b_i) - p(b_i)q(a_i))(q_{a_i} - p_{a_i}) \neq \mathbf{0}, \end{aligned}$$

where the inequality follows from failure of Lemma 9-(i) and Lemma 9-(ii). We conclude (18) fails.

Consider now the case in which $\{a_i, b_i\} \nsubseteq \text{supp}_i(p) \cap \text{supp}_i(q)$. Because Lemma 9-(ii) fails, we can assume $p(a_i) = 0 < p(b_i)$ without loss of generality. Since the lemma assume $a_i \in \text{supp}_i(p) \cup \text{supp}_i(q)$, it follows $q(a_i) > 0$. Therefore, we can use failure of Lemma 9-(ii) to deduce that $p_{b_i} \neq q_{a_i}$. Using these facts, we obtain that

$$p(a_i)q(b_i)q_{b_i} + q(a_i)p(b_i)p_{b_i} - q(b_i)p(a_i)p_{a_i} - p(b_i)q(a_i)q_{a_i} = q(a_i)p(b_i)(p_{b_i} - q_{a_i}) \neq \mathbf{0}.$$

It follows that (18) fails. □

We are now ready to prove Proposition 9. The "if" portion is straightforward; the "only if" portion uses Lemma 9.

Proof of Proposition 9. We first prove the "if" portion. Let $p \in B$ and $a_i, b_i \in \text{supp}_i(p)$. Let $t^1, \ldots, t^n > 0$ and $p^1, \ldots, p^n \in \text{ext}(B)$ such that

$$p = \sum_{m=1}^{n} t^m p^m.$$

Simple algebra shows that for all $c_i \in \text{supp}_i(p)$

$$p_{c_i} = \sum_{m=1}^{n} \frac{t^m p^m(c_i)}{\sum_{l=1}^{n} t^l p^l(c_i)} p^m_{c_i}.$$

If Proposition 9-(i) holds, then

$$\begin{aligned} p_{a_i} = \sum_{m=1}^{n} \frac{t^m p^m(a_i)}{\sum_{l=1}^{n} t^l p^l(a_i)} p^m_{a_i} &= \sum_{m=1}^{n} \frac{t^m p^m(a_i)}{\sum_{l=1}^{n} t^l p^l(a_i)} \mu \\ &= \mu \\ &= \sum_{m=1}^{n} \frac{t^m p^m(b_i)}{\sum_{l=1}^{n} t^l p^l(b_i)} \mu = \sum_{m=1}^{n} \frac{t^m p^m(b_i)}{\sum_{l=1}^{n} t^l p^l(b_i)} p^m_{b_i} = p_{b_i}. \end{aligned}$$

Suppose now that Proposition 9-(ii) holds. For every m, $p^m(a_i)p^m_{a_i} = \lambda p^m(b_i)p^m_{b_i}$ implies $p^m(a_i) = \lambda p^m(b_i)$ and $p^m_{a_i} = p^m_{b_i}$. Thus,

$$p_{a_i} = \sum_{m=1}^{n} \frac{t^m p^m(a_i)}{\sum_{l=1}^{n} t^l p^l(a_i)} p^m_{a_i} = \sum_{m=1}^{n} \frac{t^m \lambda p^m(b_i)}{\sum_{l=1}^{n} t^l \lambda p^l(b_i)} p^m_{b_i} = \sum_{m=1}^{n} \frac{t^m p^m(b_i)}{\sum_{l=1}^{n} t^l p^l(b_i)} p^m_{b_i} = p_{b_i}.$$

This concludes the proof of the proposition's "if" portion.

We now show the proposition's "only if" portion. We proceed by contradiction: we assume that Proposition 9-(i) and Proposition 9-(ii) both fail and show that there exists $p \in B$ such that $a_i, b_i \in \mathrm{supp}_i(p)$ and $p_{a_i} \neq p_{b_i}$. As we are done if $p_{a_i} \neq p_{b_i}$ for some $p \in \mathrm{ext}(B)$ with $a_i, b_i \in \mathrm{supp}_i(p)$, assume $p_{a_i} = p_{b_i}$ holds for all such p.

Since Proposition 9-(i) fails, and a_i and b_i are B-coherent, there exist $p, q \in \mathrm{ext}(B)$ such that $p(a_i) > 0$, $q(b_i) > 0$, and $p_{a_i} \neq q_{b_i}$. As we are done if $(0.5p + 0.5q)_{a_i} \neq (0.5p + 0.5q)_{b_i}$, assume $(0.5p+0.5q)_{a_i} = (0.5p+0.5q)_{b_i}$. Since $p_{a_i} \neq q_{b_i}$, Lemma 9-(i) fails. Thus, Lemma 9-(ii) must hold: there exist $\lambda > 0$ such that $p(a_i) = \lambda p(b_i)$ and $q(a_i) = \lambda q(b_i)$; in particular, $p(b_i) > 0$ and $q(a_i) > 0$.

Since Proposition 9-(ii) fails, there must exist $r \in \mathrm{ext}(B)$ such that $r(a_i) \neq \lambda r(b_i)$; in particular, $r(a_i) > 0$ or $r(b_i) > 0$. Let $c_i \in \{a_i, b_i\}$ such that $r(c_i) > 0$. Since $p_{a_i} \neq q_{b_i}$, either $r_{c_i} \neq p_{a_i}$, or $r_{c_i} \neq p_{b_i}$, or both. Thus, by Lemma 9, either $(0.5p + 0.5r)_{a_i} \neq (0.5p + 0.5r)_{c_i}$, or $(0.5q + 0.5r)_{b_i} \neq (0.5q + 0.5r)_{c_i}$, or both. In any case, we have found $p \in B$ such that $a_i, b_i \in \mathrm{supp}_i(p)$ and $p_{a_i} \neq p_{b_i}$. $\square$

A.5. Proofs for Example 3

Proposition 10. *Let I be a singleton. The following statements are equivalent:*

(i) The set $J(a)$ contains at least two elements.

(ii) There is no belief $\mu \in \Delta(\Theta)$ for which a is the unique best response.

(iii) There is a mixed action $\alpha \in \Delta(A \setminus \{a\})$ that weakly dominates a.

Proof. "(i) implies (ii)." We prove the contrapositive. Assume there is a belief $\mu \in \Delta(\Theta)$ for which a is the unique best response. Since π has full support, we can find $\nu \in \Delta(\Theta)$ and $t \in (0, 1)$ such that

$$\pi = t\mu + (1 - t)\nu.$$

Let b be a best response to the belief ν. Define the outcome $p \in \Delta_\pi(A \times \Theta)$

$$p(c, \theta) = \begin{cases} t\mu(\theta) & \text{if } c = a, \\ (1-t)\nu(\theta) & \text{if } c = b, \\ 0 & \text{otherwise.} \end{cases}$$

The outcome p is a BCE. If $a = b$, then $p_a = \pi$; if $a \neq b$, then $p_a = \mu$. In any case, a is the unique best response to p_a. Hence, $J(a) = \{a\}$.

"(ii) implies (iii)." Assume that there is no belief μ for which a is the unique best response, that is,

$$\max_{\mu \in \Delta(\Theta)} \min_{b \in A \setminus \{a\}} \sum_\theta (u(a, \theta) - u(b, \theta))\mu(\theta) \leq 0.$$

Equivalently,

$$\max_{\mu\in\Delta(\Theta)}\min_{\alpha\in\Delta(A\setminus\{a\})}\sum_{b,\theta}(u(a,\theta)-u(b,\theta))\mu(\theta)\alpha(b)\leq 0.$$

By the minimax theorem (e.g., Rockafellar, 1970, Corollary 37.3.2),

$$\min_{\alpha\in\Delta(A\setminus\{a\})}\max_{\mu\in\Delta(\Theta)}\sum_{b,\theta}(u(a,\theta)-u(b,\theta))\mu(\theta)\alpha(b)\leq 0.$$

Equivalently,

$$\min_{\alpha\in\Delta(A\setminus\{a\})}\max_{\theta\in\Theta}\left(u(a,\theta)-\sum_{b}u(b,\theta)\alpha(b)\right)\leq 0.$$

Thus, there exists a mixed action $\alpha\in\Delta(A\setminus\{a\})$ that weakly dominates a.

"(iii) implies (ii)." Let $\alpha\in\Delta(A\setminus\{a\})$ be a mixed action that weakly dominates a. Then

$$\{b:\alpha(b)>0\}\subseteq J(a).$$

Thus, since $a\in J(a)$ and $\alpha(a)=0$, the set $J(a)$ contains at two elements. $\square$

Proposition 11. *Let I be a singleton. For generic u, if an action a is weakly dominated by some mixed action $\alpha\in\Delta(A\setminus\{a\})$, then it is strictly dominated by some mixed action (possibly different from α).*

Proof. Let a be an action that is weakly dominated by a mixed action $\alpha\in\Delta(A\setminus\{a\})$. Let A' be the support of α, and let Θ' be set of states θ for which

$$u(a,\theta)=\sum_{b}u(b,\theta)\alpha(b). \tag{19}$$

Let m be the cardinality of A', and let n be the cardinality of Θ'. We consider the $m\times n$ matrix $M\in\mathbb{R}^{A'\times\Theta'}$ given by

$$M(b,\theta)=u(a,\theta)-u(b,\theta).$$

For generic u, the matrix M has full rank. By (19), the rows of M are linearly dependent. Thus, the rank of M must be n, the number of columns. We obtain that the row space of M has dimension n. Hence, we can find $\beta\in\mathbb{R}^{A'}$ such that for every $\theta\in\Theta'$

$$\sum_{b}(u(a,\theta)-u(b,\theta))\beta(b)<0.$$

For every $t > 0$, we define $\alpha_t \in \mathbb{R}^{A'}$ by

$$\alpha_t(b) = \frac{\alpha(b) + t\beta(b)}{\sum_c \alpha(c) + t\beta(c)}.$$

For t sufficiently small, α_t is a mixed action that strictly dominates a. $\square$

A.6. Proof of Theorem 3

For a utility vector $u \in \mathbb{R}^{I \times A \times \Theta}$, we denote by $BCE(u)$ the set of BCEs, by $BKE(u)$ the set of BKEs, by $\mathrm{cl}(BKE(u))$ the closure of $BKE(u)$, and by $\|u\|$ the Euclidean norm.

Lemma 10. *For every $u \in \mathbb{R}^{I \times A \times \Theta}$, $p \in BCE(u)$, and $\epsilon > 0$, there exists $u' \in \mathbb{R}^{I \times A \times \Theta}$ such that $\|u - u'\| \leq \epsilon$ and $p \in BKE(u')$.*

Proof. For each player $i \in I$, we consider a set $P_i \subseteq \Delta(A_{-i} \times \Theta)$ given by

$$P_i = \{p_{a_i} : a_i \in \mathrm{supp}_i(p)\}.$$

Let n_i be the cardinality of P_i (of course, it could be that n_i is smaller than the cardinality of $\mathrm{supp}_i(p)$). Reasoning inductively, we can find an enumeration $p_1, \ldots, p_{n_i}$ of the elements of P_i such that, for every $m_i \in \{1, \ldots, n_i\}$, p_{m_i} is an extreme point of the convex hull of $\{p_1, \ldots, p_{m_i}\}$.

By an hyperplane separation theorem (e.g., Rockafellar, 1970, Corollary 11.4.2) for every $m_i \in \{2, \ldots, n_i\}$ we can find a function $f_{m_i} : A_{-i} \times \Theta \to \mathbb{R}$ such that

$$\sum_{a_{-i},\theta} f_{m_i}(a_{-i}, \theta) p_{m_i}(a_{-i}, \theta) > 0 \geq \max_{l_i \in \{1,\ldots,m_i - 1\}} \sum_{a_{-i},\theta} f_{m_i}(a_{-i}, \theta) p_{l_i}(a_{-i}, \theta). \tag{20}$$

For $m_i = 1$, we define $f_1(a_{-i}, \theta) = 1$ for all $a_{-i} \in A_{-i}$ and $\theta \in \Theta$.

For every $l_i \in \{1, \ldots, n_i - 1\}$, we choose $t_{l_i} \in (0, 1]$ such that for every $m_i \in \{l_i + 1, \ldots, n_i\}$,

$$\sum_{a_{-i},\theta} f_{m_i}(a_{-i}, \theta) p_{m_i}(a_{-i}, \theta) > t_{l_i} \sum_{a_{-i}} f_{l_i}(a_{-i}) p_{m_i}(a_{-i}). \tag{21}$$

We can choose such a t_{l_i} because the left-hand side of (21) is positive—see (20). For $l_i = n_i$, we simply define $t_{n_i} = 1$.

For every $l_i \in \{1, \ldots, n_i\}$, we define

$$s_{l_i} = \prod_{m_i = l_i}^{n_i} t_{m_i}.$$

Using (21), simple algebra shows that for every $l_i \in \{1, \ldots, n_i - 1\}$ and $m_i \in \{l_i + 1, \ldots, n_i\}$,

$$s_{m_i} \sum_{a_{-i},\theta} f_{m_i}(a_{-i}, \theta) p_{m_i}(a_{-i}, \theta) > s_{l_i} \sum_{a_{-i},\theta} f_{l_i}(a_{-i}, \theta) p_{m_i}(a_{-i}, \theta). \tag{22}$$

For $a_i \in \mathrm{supp}_i(p)$, we define the function $g_{a_i} : A_{-i} \times \Theta \to \mathbb{R}$ by

$$g_{a_i}(a_{-i}, \theta) = s_{m_i} \cdot f_{m_i}(a_{-i}, \theta)$$

where m_i is such that $p_{a_i} = p_{m_i}$. For $a_i \notin \mathrm{supp}_i(p)$, we define $g_{a_i} = 0$.

We claim that for all $a_i \in \mathrm{supp}_i(p)$ and $b_i \notin \{c_i \in \mathrm{supp}_i(p) : p_{c_i} = p_{a_i}\}$,

$$\sum_{a_{-i},\theta} g_{a_i}(a_{-i}, \theta) p_{a_i}(a_{-i}, \theta) > \sum_{a_{-i},\theta} g_{b_i}(a_{-i}, \theta) p_{a_i}(a_{-i}, \theta). \tag{23}$$

To verify the claim, pick $m_i \in \{1, \ldots, n_i\}$ such that $p_{a_i} = p_{m_i}$. From the left-hand side of (20) and the fact that $s_{m_i} > 0$, we obtain that

$$\sum_{a_{-i},\theta} g_{a_i}(a_{-i}, \theta) p_{a_i}(a_{-i}, \theta) = s_{m_i} \sum_{a_{-i},\theta} f_{m_i}(a_{-i}, \theta) p_{m_i}(a_{-i}, \theta) > 0. \tag{24}$$

Hence, for $b_i \notin \mathrm{supp}_i(p)$, we have

$$\sum_{a_{-i},\theta} g_{a_i}(a_{-i}, \theta) p_{a_i}(a_{-i}, \theta) > 0 = \sum_{a_{-i},\theta} g_{b_i}(a_{-i}, \theta) p_{a_i}(a_{-i}, \theta)$$

where the equality follows from $g_{b_i} = 0$.

Assume now that $b_i \in \mathrm{supp}_i(p)$. Choose l_i such that $p_{b_i} = p_{l_i}$. Since $p_{a_i} \neq p_{b_i}$, $m_i \neq l_i$. Suppose that $l_i > m_i$. It follows from the right-hand side of (20)—in (20) the roles of l_i and m_i are inverted—that

$$0 \geq \sum_{a_{-i},\theta} f_{l_i}(a_{-i}, \theta) p_{m_i}(a_{-i}, \theta).$$

Thus, given that $s_{l_i} > 0$, we deduce that

$$0 \geq \sum_{a_{-i},\theta} g_{b_i}(a_{-i}, \theta) p_{a_i}(a_{-i}, \theta) = s_{l_i} \sum_{a_{-i},\theta} f_{l_i}(a_{-i}, \theta) p_{m_i}(a_{-i}, \theta).$$

We obtain that

$$\sum_{a_{-i},\theta} g_{a_i}(a_{-i}, \theta) p_{a_i}(a_{-i}, \theta) > 0 \geq \sum_{a_{-i},\theta} g_{b_i}(a_{-i}, \theta) p_{a_i}(a_{-i}, \theta)$$

where we use again (24). For the case $l_i < m_i$, the condition

$$\sum_{a_{-i},\theta} g_{a_i}(a_{-i},\theta) p_{a_i}(a_{-i},\theta) > \sum_{a_{-i},\theta} g_{b_i}(a_{-i},\theta) p_{a_i}(a_{-i},\theta)$$

is equivalent to (22). We conclude that (23) holds.

To complete the proof of the lemma, for every $\delta > 0$ we define $u' = (u'_i)_{i \in I}$ by

$$u'_i(a,\theta) = u_i(a,\theta) + \delta g_{a_i}(a_{-i},\theta).$$

By choosing δ sufficiently small, we can be make sure that $\|u - u'\| \leq \epsilon$. Since $p \in BCE(u)$, it follows from (23) that for all $i \in I$ and $a_i \in \operatorname{supp}_i(p)$,

$$a_i \in BR'(p_{a_i}) \subseteq \{b_i \in \operatorname{supp}_i(p) : p_{a_i} = p_{b_i}\}$$

where $BR'(p_{a_i})$ is the set of i's best response to a belief p_{a_i} given utility function u'_i. Thus, $p \in BKE(u')$. □

A subset of a Euclidean space is **semi-algebraic** if it is defined by finite systems of polynomial inequalities. A correspondence between Euclidean spaces is semi-algebraic if its graph is semi-algebraic. The background knowledge on semi-algebraic sets that we use in this proof can be gathered from Blume and Zame (1994, Section 2).

Lemma 11. *The correspondences $u \mapsto BCE(u)$, $u \mapsto BKE(u)$, and $u \mapsto \operatorname{cl}(BKE(u))$ are semi-algebraic.*

Proof. The BCE correspondence is semi-algebraic: for all $u \in \mathbb{R}^{I \times A \times \Theta}$ and $p \in \mathbb{R}^{A \times \Theta}$, $p \in BCE(u)$ if and only if the pair (u,p) is a solution to the following finite system of polynomial inequalities:

$$\begin{aligned} p(a,\theta) &\geq 0 && \text{for all } a \in A \text{ and } \theta \in \Theta, \\ \sum_a p(a,\theta) &= \pi(\theta) && \text{for all } \theta \in \Theta, \\ \sum_{a_{-i},\theta} (u(a,\theta) - u(b_i,a_{-i},\theta)) p(a,\theta) &\geq 0 && \text{for all } i \in I \text{ and } a_i, b_i \in A_i. \end{aligned}$$

The BKE correspondence is also semi-algebraic. To prove it, for every $u \in \mathbb{R}^{I \times A \times \Theta}$, $p \in \mathbb{R}^{A \times \Theta}$, $i \in I$, and $a_i, b_i, c_i \in A_i$, we denote by $F(u,p,a_i,b_i,c_i)$ the quantity

$$\sum_{a_{-i},\theta} (u(a_i,a_{-i},\theta) - u(c_i,a_{-i},\theta)) p(a_i,a_{-i},\theta) + (u(b_i,a_{-i},\theta) - u(c_i,a_{-i},\theta)) p(b_i,a_{-i},\theta).$$

We observe that $p \in BKE(u)$ if and only if $p \in BCE(u)$ and for every $i \in I$ there is $T_i \subseteq A_i \times A_i$ such that the pair (u, p) is a solution of the following finite system of polynomial inequalities:

$$\sum_{a_{-i},\theta} (p(a_i, a_{-i}, \theta)p(b_i) - p(b_i, a_{-i}, \theta)p(a_i))^2 = 0 \qquad \text{for all } (a_i, b_i) \in T_i,$$

$$F(u, p, a_i, b_i, c_i) > 0 \qquad \text{for all } (a_i, b_i) \notin T_i \text{ and } c_i \in A_i.$$

Thus, $p \in BKE(u)$ if and only if it the solution of one of finitely many systems of polynomial inequalities; we conclude that the BKE correspondence is semi-algebraic.

The correspondence $u \mapsto \text{cl}(BKE(u))$ is also semi-algebraic. Indeed, $p \in \text{cl}(BKE(u))$ if and only if for every $\epsilon > 0$ there exists $q \in \mathbb{R}^{A\times\Theta}$ such that $\|p - q\| \leq \epsilon$ and $q \in BKE(u)$. Thus, since the BKE correspondence is semi-algebraic, the graph of the correspondence $u \mapsto \text{cl}(BKE(u))$ is defined by a first-order formula and therefore semi-algebraic by the Tarski-Seidenberg theorem (Blume and Zame, 1994, page 787). □

We are ready to complete the proof of the theorem. By Lemma 11, the correspondence $u \mapsto \text{cl}(BKE(u))$ is semi-algebraic. Hence, there is an open subsets U of $\mathbb{R}^{I\times A\times\Theta}$ such that the complement of U has Lebesgue measure zero, and $u \mapsto \text{cl}(BKE(u))$ is continuous on U (Blume and Zame, 1994, page 786).

We claim that for all $u \in U$, $BCE(u) = \text{cl}(BKE(u))$. To prove the claim, take any $u \in U$. Since $BCE(u)$ is closed and $BKE(u) \subseteq BCE(u)$, $\text{cl}(BKE(u)) \subseteq BCE(u)$. To verify the other inclusion, we use Lemma 10 to find a sequence of games $(u^n)_{n=1}^{\infty}$ such that $u^n \to u$ and, for every n, $p \in BKE(u^n) \subseteq \text{cl}(BKE(u^n))$. Since $\text{cl}(BKE(u))$ is continuous at u, we have $p \in \text{cl}(BKE(u))$; see Aliprantis and Border (2006, Theorem 17.16). Hence, $BCE(u) \subseteq \text{cl}(BKE(u))$. We deduce that $BCE(u) = \text{cl}(BKE(u))$, as desired.

We conclude that for generic u, the BKE set is dense in the BCE set. ■

A.7. Proofs of Propositions 6 and 7

In what follows, define $t^c := 1$ (the "choke" price). To prove the proposition, we rely on a lemma that is modeled on Jann and Schottmüller (2015):

Lemma 12. *For every $p \in BCE$,*

$$p(\{(\theta, a) : \min\{a_1, a_2\} \leq \theta_2\}) = 1.$$

Proof. Let $p \in \Delta_\pi(\Theta \times A)$ be a Bayes correlated equilibrium. We denote by $t^{\sup}$ the essential supremum of the market price:

$$t^{\sup} = \inf\{t \in [0, \infty) : p(\{(\theta, a) : \min\{a_1, a_2\} \leq t\}) = 1\}.$$

The goal is to show that $t^{\sup} \leq \theta_2$, which is equivalent to

$$p(\{(\theta, a) : \min\{a_1, a_2\} \leq \theta_2\}) = 1.$$

First, we verify that $t^{\sup} < 1$.

Claim 1. *$t^{\sup} < t^c$.*

Proof of the claim. Let $E := \{(\theta, a) : \min\{a_1, a_2\} > a^m_{\theta_2}\}$. Firm 2 can deviate to the following alternative strategy:

$$s_2(a_2) = \begin{cases} a_2 & \text{if } a_2 \leq a^m_{\theta_2}, \\ a^m_{\theta_2} & \text{otherwise.} \end{cases}$$

The loss from deviating, which must be non-positive, is

$$\begin{aligned} L &= \int_{\Theta \times A} u_2(\theta, a) - u_2(\theta, a_1, s_2(a_2))\, \mathrm{d}p(\theta, a) \\ &= \int_E u_2(\theta, a) - D\left(a^m_{\theta_2}\right)\left(a^m_{\theta_2} - \theta_2\right)\, \mathrm{d}p(\theta, a). \end{aligned}$$

For all $a_2 \neq a^m_{\theta_2}$, $u_2(\theta, a) < D\left(a^m_{\theta_2}\right)\left(a^m_{\theta_2} - \theta_2\right)$. Since $L \geq 0$, we conclude that $p(E) = 0$. This shows that $t^{\sup} \leq a^m_{\theta_2} < t^c$. □

For every $t < t^{\sup}$, we define the events

$$\begin{aligned} E_t &= \{(\theta, a) : a_1 \in (t, t^{\sup}] \text{ and } a_1 \leq a_2\}, \\ F_t &= \{(\theta, a) : a_2 \in (t, t^{\sup}] \text{ and } a_2 < a_1\}. \end{aligned}$$

Claim 2. *If $t^{\sup} > \theta_2$, then for every $t \in (\theta_2, t^{\sup})$, $p(E_t) > 0$.*

Proof of the claim. By contradiction, suppose that $p(E_t) = 0$. Consider the following deviation for firm 1:

$$s_1(a_1) = \begin{cases} a_1 & \text{if } a_1 \notin (t, t^{\sup}], \\ t & \text{otherwise.} \end{cases}$$

The loss from deviating, which must be non-negative, is

$$\begin{aligned}L &= \int_{\Theta \times A} u_1(\theta, a) - u_1(\theta, s_1(a_1), a_2)\, \mathrm{d}p(\theta, a) \\ &= \int_{E_t} D(a_1)(a_1 - \theta_1)\mathrm{d}p(\theta, a) - \int_{\Theta \times (t, t^{\mathrm{sup}}] \times [t, \infty)} D(t)(t - \theta_1)\, \mathrm{d}p(\theta, a).\end{aligned}$$

Using the hypothesis that $p(E_t) = 0$, we obtain that

$$L = -\int_{\Theta \times (t, t^{\mathrm{sup}}] \times [t, \infty)} D(t)(t - \theta_1)\, \mathrm{d}p(\theta, a).$$

Since $t \geq \theta_2 > \max \Theta_1$ and $t^{\mathrm{sup}} < t^c$ (see Claim 1), the integrand is positive. Since $L \geq 0$, we must have

$$p(\Theta \times (t, t^{\mathrm{sup}}] \times [t, \infty)) = 0. \tag{25}$$

Next we show that $p(\{(\theta, a) : \min\{a_1, a_2\} > t) = 0$, reaching a contradiction with $t < t^{\mathrm{sup}}$.

Consider the following deviation for firm 1:

$$s_1'(a_1) = \begin{cases} a_1 & \text{if } a_1 \leq t^{\mathrm{sup}}, \\ t & \text{otherwise.} \end{cases}$$

The loss from deviating, which must be non-negative, is

$$\begin{aligned}L' &= \int_{\Theta \times A} u_1(\theta, a) - u_1(\theta, s_1'(a_1), a_2)\, \mathrm{d}p(\theta, a) \\ &= -\int_{\Theta \times (t^{\mathrm{sup}}, \infty) \times [t, \infty)} D(t)(t - \theta_1)\, \mathrm{d}p(\theta, a).\end{aligned}$$

Since $t \geq \theta_2 > \max \Theta_1$ and $t^{\mathrm{sup}} < t^c$ (see Claim 1), the integrand is positive. Since $L' \geq 0$, we must have

$$p(\Theta \times (t^{\mathrm{sup}}, \infty) \times [t, \infty)) = 0. \tag{26}$$

Putting together (25) and (26), we conclude that

$$p(\{(\theta, a) : \min\{a_1, a_2\} > t) = 0,$$

which contradicts $t^{\mathrm{sup}} > t$. We conclude that $p(E_t) > 0$. $\square$

Claim 3. *Set $t = \frac{1}{4}\theta_2 + \frac{3}{4}t^{\mathrm{sup}}$. If $t^{\mathrm{sup}} > \theta_2$, then $p(E_t) > p(F_t)$.*

Proof of the claim. Consider the following deviation for firm 1:

$$s_1(a_1) = \begin{cases} a_1 & \text{if } a_1 \notin (t, t^{\text{sup}}], \\ t & \text{otherwise.} \end{cases}$$

The loss from deviating, which must be non-negative, is

$$\begin{aligned} L &= \int_{\Theta \times A} u_1(\theta, a) - u_1(\theta, s_1(a_1), a_2)\, \mathrm{d}p(\theta, a) \\ &= \int_{E_t} D(a_1)(a_1 - \theta_1) \mathrm{d}p(\theta, a) - \int_{\Theta \times (t, t^{\text{sup}}] \times [t, \infty)} D(t)(t - \theta_1)\, \mathrm{d}p(\theta, a) \\ &\leq \int_{E_t} D(t)(a_1 - \theta_1) \mathrm{d}p(\theta, a) - \int_{\Theta \times (t, t^{\text{sup}}] \times (t, t^{\text{sup}}]} D(t)(t - \theta_1)\, \mathrm{d}p(\theta, a) \\ &= \int_{E_t} D(t)(a_1 - t) \mathrm{d}p(\theta, a) - \int_{(\Theta \times (t, t^{\text{sup}}] \times (t, t^{\text{sup}}]) \setminus E_t} D(t)(t - \theta_1)\, \mathrm{d}p(\theta, a), \end{aligned}$$

where the inequality holds because $a_1 > t \geq \theta_2 \geq \max \Theta_1$ and $D(t) \geq D(a_1)$. We observe that

$$(\Theta \times (t, t^{\text{sup}}] \times (t, t^{\text{sup}}]) \setminus E_t = F_t \setminus (\Theta \times (t^{\text{sup}}, \infty) \times (t, t^{\text{sup}}]).$$

Note that $p(\Theta \times (t^{\text{sup}}, \infty) \times (t, t^{\text{sup}}]) = 0$—otherwise, firm 1 would benefit by playing t whenever the recommendation is $a_1 > t^{\text{sup}}$. We deduce that

$$\begin{aligned} L &\leq \int_{E_t} D(t)(a_1 - t) \mathrm{d}p(\theta, a) - \int_{F_t} D(t)(t - \theta_1)\, \mathrm{d}p(\theta, a) \\ &\leq \int_{E_t} D(t)(a_1 - t) \mathrm{d}p(\theta, a) - \int_{F_t} D(t)(t - \bar{\theta}_1)\, \mathrm{d}p(\theta, a) \\ &= D(t)(t - \bar{\theta}_1) \left(\int_{E_t} \frac{a_1 - t}{t - \bar{\theta}_1} \mathrm{d}p(\theta, a) - p(F_t) \right). \end{aligned}$$

Note that

$$0 < \frac{a_1 - t}{t - \bar{\theta}_1} \leq \frac{t^{\text{sup}} - t}{t - \bar{\theta}_1} \leq \frac{t^{\text{sup}} - t}{t - \theta_2} < 1.$$

Thus,

$$0 \leq L \leq \int_{E_t} \frac{a_1 - t}{t - \bar{\theta}_1} \mathrm{d}p(\theta, a) - p(F_t) < p(E_t) - p(F_t).$$

where the strict inequality uses the fact that $p(E_t) > 0$ (see Claim 2). We conclude that $p(E_t) > p(F_t)$. □

Claim 4. *Set $t = \frac{1}{4}\theta_2 + \frac{3}{4}t^{\text{sup}}$. If $t^{\text{sup}} > \theta_2$, then $p(F_t) \geq p(E_t)$.*

Proof of the claim. Consider the following deviation for firm 2:

$$s_2(a_2) = \begin{cases} a_2 & \text{if } a_2 \notin (t, t^{\text{sup}}], \\ t & \text{otherwise.} \end{cases}$$

The loss from deviating, which must be non-negative, is

$$\begin{aligned} L &= \int_{\Theta \times A} u_2(\theta, a) - u_2(\theta, a_1, s_2(a_2)) \, \mathrm{d}p(\theta, a) \\ &= \int_{F_t} D(a_2)(a_2 - \theta_2) \, \mathrm{d}p(\theta, a) - \int_{\Theta \times [t, \infty) \times (t, t^{\text{sup}}]} D(t)(t - \theta_2) \, \mathrm{d}p(\theta, a) \\ &\leq \int_{F_t} D(t)(a_2 - \theta_2) \, \mathrm{d}p(\theta, a) - \int_{\Theta \times (t, t^{\text{sup}}] \times (t, t^{\text{sup}}]} D(t)(t - \theta_2) \, \mathrm{d}p(\theta, a) \\ &= \int_{F_t} D(t)(a_2 - t) \, \mathrm{d}p(\theta, a) - \int_{(\Theta \times (t, t^{\text{sup}}] \times (t, t^{\text{sup}}]) \setminus F_t} D(t)(t - \theta_2) \, \mathrm{d}p(\theta, a) \end{aligned}$$

where the inequality holds because $a_2 > t \geq \theta_2$ and $D(t) \geq D(a_2)$. We observe that

$$(\Theta \times (t, t^{\text{sup}}] \times (t, t^{\text{sup}}]) \setminus F_t = E_t \setminus (\Theta \times (t, t^{\text{sup}}] \times (t^{\text{sup}}, \infty)).$$

Note that $p(\Theta \times (t, t^{\text{sup}}] \times (t^{\text{sup}}, \infty)) = 0$—otherwise, firm 2 would benefit by playing t whenever the recommendation is $a_2 > t^{\text{sup}}$. We deduce that

$$\begin{aligned} L &\leq \int_{F_t} D(t)(a_2 - t) \, \mathrm{d}p(\theta, a) - \int_{E_t} D(t)(t - \theta_2) \, \mathrm{d}p(\theta, a) \\ &= D(t)(t - \theta_2) \int_{F_t} \frac{a_2 - t}{t - \theta_2} \, \mathrm{d}p(\theta, a) - p(E_t). \end{aligned}$$

Note that

$$0 < \frac{a_2 - t}{t - \theta_2} \leq \frac{t^{\text{sup}} - t}{t - \theta_2} < 1.$$

Thus,

$$0 \leq L \leq \int_{E_t} \frac{a_1 - t}{t - \bar{\theta}_1} \mathrm{d}p(\theta, a) - p(F_t) \leq p(F_t) - p(E_t).$$

We conclude that $p(F_t) \geq p(E_t)$. □

Note that Claims 3 and 4 contradict each other. Thus, we must have $t^{\text{sup}} \leq \theta_2$, as stated in the lemma. □

Under complete information, the game has a Bayes Nash equilibrium where the market price is equal to the marginal cost of firm 1. Thus

$$\max_{p \in BCE} CS(p) \geq \frac{1}{2} \int_{\underline{\theta}_1}^{\infty} D(t) \, \mathrm{d}t + \frac{1}{2} \int_{\bar{\theta}_1}^{\infty} D(t) \, \mathrm{d}t. \tag{27}$$

Next we show that the inequality cannot be strict:

Lemma 13. *We have*

$$\max_{p \in BCE} CS(p) = \frac{1}{2}\int_{\underline{\theta}_1}^{\infty} D(t)\,\mathrm{d}t + \frac{1}{2}\int_{\overline{\theta}_1}^{\infty} D(t)\,\mathrm{d}t.$$

Proof. Let $p \in BCE$. We claim that

$$p(\{(\theta, a) : \theta_2 \leq a_2 < a_1\}) = 0.$$

To prove it, define the event

$$E = \{(\theta, a) : \min\{a_1, a_2\} \leq \theta_2\}.$$

By Lemma 12, $p(E) = 1$. Thus,

$$\begin{aligned} p(\{(\theta, a) : \theta_2 \leq a_2 < a_1\}) &= p(\{(\theta, a) : \theta_2 \leq a_2 < a_1\} \cap E) \\ &= p(\{(\theta, a) : \theta_2 = a_2 < a_1\}). \end{aligned}$$

Set $F = \{(\theta, a) : \theta_2 = a_2 < a_1\}$. Consider the following deviation for firm 1:

$$s_1(a_1) = \begin{cases} a_1 & \text{if } a_1 \leq \theta_2 \\ \theta_2 & \text{otherwise.} \end{cases}$$

The loss from deviating, which must be non-negative, is

$$\begin{aligned} L &= \int_{\Theta \times A} u_1(\theta, a) - u_1(\theta, s_1(a_1), a_2)\,\mathrm{d}p(\theta, a) \\ &= \int_{\Theta \times (\theta_2, \infty) \times A_2} u_1(\theta, a) - u_1(\theta, \theta_2, a_2)\,\mathrm{d}p(\theta, a) \\ &= \int_{(\Theta \times (\theta_2, \infty) \times A_2) \cap E} u_1(\theta, a) - u_1(\theta, \theta_2, a_2)\,\mathrm{d}p(\theta, a) \\ &= -\int_F D(\theta_2)(\theta_2 - \theta_1)\,\mathrm{d}p(\theta, a). \end{aligned}$$

Since $\theta_2 < t^c$ and $\theta_2 > \max \Theta_1$, $D(\theta_2)(\theta_2 - \theta_1) > 0$. Since $L \geq 0$, we conclude that $p(F) = 0$. As explained above, this implies that

$$p(\{(\theta, a) : \theta_2 \leq a_2 < a_1\}) = 0.$$

It follows that

$$p(\{(a,\theta) : a_2 < a_1\}) = p(\{(a,\theta) : a_2 < a_1 \text{ and } a_2 < \theta_2\}).$$

This shows that $p(\{(a,\theta) : a_2 < a_1\}) = 0$—otherwise, firm 2 would make negative profits.

Overall, we conclude that firm 1 gets the whole market with probability one. This implies that

$$p(\{(a,\theta) : \min\{a_1, a_2\} = a_1 \geq E_p[\theta_1|a_1]\} = 1$$

where $E_p[\theta_1|a_1]$ is a version of the conditional expectation of θ_1 given a_1. Thus,

$$CS(p) \leq \int_{A_1} CS\left(E_p[\theta_1|a_1]\right) \mathrm{d}p_1(a_1).$$

By standard arguments, the distribution of $E_p[\theta_1|a_1]$ is a mean-preserving contraction of π_1. Since $CS(t)$ is convex in t,

$$\int_{A_1} CS\left(E_p[\theta_1|a_1]\right) \mathrm{d}p_1(a_1) \leq \frac{1}{2}\int_{\underline{\theta}_1}^{\infty} D(t)\,\mathrm{d}t + \frac{1}{2}\int_{\bar{\theta}_1}^{\infty} D(t)\,\mathrm{d}t.$$

Since the choice of $p \in BCE$ is arbitrary, we conclude that

$$\max_{p\in BCE} CS(p) \leq \frac{1}{2}\int_{\underline{\theta}_1}^{\infty} D(t)\,\mathrm{d}t + \frac{1}{2}\int_{\bar{\theta}_1}^{\infty} D(t)\,\mathrm{d}t.$$

The equality follows from (27). □

Under mo information, the game has a Bayes Nash equilibrium where the market price is equal to the expected marginal cost of firm 1. Thus

$$\max_{p\in BKE} CS(p) \geq \int_{\frac{1}{2}\underline{\theta}_1 + \frac{1}{2}\bar{\theta}_1}^{\infty} D(t)\,\mathrm{d}t. \tag{28}$$

Next we show that the inequality cannot be strict:

Lemma 14. *The set BKE corresponds to the set of Bayes Nash equilibria of the no-information game. In particular,*

$$\max_{p\in BKE} CS(p) = \int_{\frac{1}{2}\underline{\theta}_1 + \frac{1}{2}\bar{\theta}_1}^{\infty} D(t)\,\mathrm{d}t.$$

Proof. Let $p \in BKE$. Define

$$E = \{(\theta, a) : \min\{a_1, a_2\} \leq \theta_2\}.$$

By Lemma 12, $p(E) = 1$.

The fact that $P(E) = 1$ implies that, almost surely, firm 2 makes non-positive profits. Firm 2 can achieve zero profits simply by pricing at marginal cost. Thus, almost surely, firm 2 makes zero profits. In addition, since $p \in BKE$, the price of firm 2 must be independent of the state and of the price of firm 1.

Next we show that the price of firm 1 is independent of the state. Let $a_2^{\inf}$ be the essential infimum of the price of firm 2:

$$a_2^{\inf} = \sup\{t : p(\{(\theta, a) : t \leq a_2\}) = 1\}.$$

We distinguish between two cases: $a_2^{\inf} \geq \theta_2$ and $a_2^{\inf} < \theta_2$.

First, assume that $a_2^{\inf} \geq \theta_2$. Regardless of the recommendation, Firm 1 can deviate to $a'_1 = \theta_2$ incurring an ex-ante loss of

$$\begin{aligned} L_1 &= \int_{\Theta \times A} (u_1(\theta, a) - u_1(\theta, a'_1, a_2)) \, \mathrm{d}p(\theta, a) \\ &= \int_E (u_1(\theta, a) - D(\theta_2)(\theta_2 - \theta_1)) \, \mathrm{d}p(\theta, a), \end{aligned}$$

where we use the fact that $p(E) = 1$. We observe that $L_1 \leq 0$. Indeed, if $\min\{a_1, a_2\} \leq \theta_2$ and $a_1 > a_2$, then $u_1(\theta, a) = 0$, in which case

$$u_1(\theta, a) - D(\theta_2)(\theta_2 - \theta_1) < 0$$

since $\theta_2 > \bar{\theta}_1 > \underline{\theta}_1$ and $D(\theta_2) \geq D(a^m_{\theta_2}) > 0$. If instead $\min\{a_1, a_2\} \leq \theta_2$ and $a_1 \leq a_2$, then $u_1(\theta, a) = D(a_1)(a_1 - \theta_1)$, in which case

$$u_1(\theta, a) - D(\theta_2)(\theta_2 - \theta_1) \leq 0$$

since $a^m_{\theta_1} \geq a^m_{\theta_2} \geq \theta_2$ and $D(a_1)(a_1 - \theta_1)$ is increasing in $a_1 \in [0, a^m_{\theta_1}]$. Overall, we conclude that $L_1 \leq 0$. Since $p \in BKE$, this shows that, when $a_2^{\inf} \geq \theta_2$, the price of firm 1 must be independent of the state.

Now, assume that $a_2^{inf} < \theta_2$. We claim that

$$p\left(\left\{(\theta, a) : a_1 > a_2^{inf}\right\}\right) = 0.$$

To prove the claim, take any $t \in (a_2^{inf}, \theta_2)$. By the definition of the essential infimum,

$$p\left(\{(\theta, a) : a_2 < t\}\right) > 0.$$

Since the price of firm 2 is independent of the price of firm 1,

$$p(\{(\theta,a):a_2<t\leq a_1\}) = p(\{(\theta,a):a_2<t\})p(\{(\theta,a):t\leq a_1\}).$$

Since, almost surely, firm 2 makes zero profits,

$$p(\{(\theta,a):a_2<t\leq a_1\}) = 0.$$

Thus, $p(\{a_1\geq t\})=0$. by taking the limit as t decreases to $a_2^{\inf}$, we conclude that

$$p\left(\left\{(\theta,a):a_1>a_2^{\inf}\right\}\right)=0,$$

as desired.

Define the event

$$F=\left\{(\theta,a):a_1\leq a_2^{\inf}\right\}.$$

Regardless of the recommendation, firm 1 can deviate to $a_1'=a_2^{\inf}$, incurring a loss of

$$\begin{aligned} L_1 &= \int_{\Theta\times A}(u_1(\theta,a)-u_1(\theta,a_1',a_2))\,\mathrm{d}p(\theta,a) \\ &= \int_F\left(D(a_2)(a_2-\theta_1)-D(a_2^{\inf})(a_2^{\inf}-\theta_1)\right)\mathrm{d}p(\theta,a), \end{aligned}$$

where we use the fact that $p(F)=1$. Since $a_{\theta_1}^m\geq a_{\theta_2}^m$, $a_2^{\inf}<\theta_2<a_{\theta_2}^m$, and $D(a_2)(a_2-\theta_2)$ in increasing in $a_1\in\left[0,a_{\theta_1}^m\right]$, we deduce that the integrand is non-negative. Thus, $L_1\leq 0$. Since $p\in BKE$, this shows that, when $a_2^{\inf}<\theta_2$, the price of firm 1 must be independent of the state.

We conclude that in every Blackwell correlated equilibrium, firms' actions are independent of each other and of the state. Thus, the set BKE corresponds to the set of Bayes Nash equilibria of the no-information game. □

Propositions 6 and 7 follow from Lemmas 13 and 14, together with the observation that $CS(t)$ is strictly convex in $t\in[0,t^c]$.

References

C. D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis*. Springer, third edition, 2006.

R. J. Aumann. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics*, 1:67–96, 1974.

R. J. Aumann. Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, 55:1–18, 1987.

D. Bergemann and S. Morris. Robust predictions in games with incomplete information. *Econometrica*, 81:1251–1308, 2013.

D. Bergemann and S. Morris. Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, 11:487–522, 2016.

D. Bergemann and S. Morris. Information design: A unified perspective. *Journal of Economic Literature*, 57:44–95, 2019.

D. Bergemann and J. Valimaki. Information acquisition and efficient mechanism design. *Econometrica*, 70:1007–1033, 2002.

D. Bergemann, B. Brooks, and S. Morris. First-price auctions with general information structures: Implications for bidding and revenue. *Econometrica*, 85:107–143, 2017.

D. Blackwell. Comparison of experiments. In J. Neyman, editor, *Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability*, pages 93–102. University of California Press, 1951.

D. Blackwell. Equivalent comparisons of experiments. *Annals of Mathematical Statistics*, 24:265–272, 1953.

L. E. Blume and W. R. Zame. The algebraic geometry of perfect and sequential equilibrium. *Econometrica*, 62:783–794, 1994.

B. Brooks and S. Du. Optimal auction design with common values: An informationally robust approach. *Econometrica*, 89:1313–1360, 2021.

A. Caplin, M. Dean, and J. Leahy. Rational inattention, optimal consideration sets, and stochastic choice. *Review of Economic Studies*, 86:1061–1094, 2019.

A. Caplin, M. Dean, and J. Leahy. Rationally inattentive behavior: Characterizing and generalizing shannon entropy. *Journal of Political Economy*, 130:1676–1715, 2022.

L. Colombo, G. Femminis, and A. Pavan. Information acquisition and welfare. *Review of Economic Studies*, 81:1438–1483, 2014.

M. Dean and N. L. Neligh. Experimental tests of rational inattention. 2022. URL `http://www.columbia.edu/~md3405/`.

T. Denti. Costly monitoring in signaling games. 2021. URL `https://sites.google.com/site/tommasojdenti/home`.

T. Denti. Posterior separable cost of information. *American Economic Review*, 112:3215–59, 2022.

T. Denti, M. Marinacci, and A. Rustichini. Experimental cost of information. *American Economic Review*, 112:3106–23, 2022.

L. Doval and J. C. Ely. Sequential information design. *Econometrica*, 88:2575–2608, 2020.

S. Du. Robust mechanisms under common valuation. *Econometrica*, 86:1569–1588, 2018.

P. Dworczak and A. Pavan. Preparing for the worst but hoping for the best: Robust (Bayesian) persuasion. *Econometrica*, 90:2017–2051, 2022.

F. Forges. Five legitimate definitions of correlated equilibrium in games with incomplete information. *Theory and Decision*, 35:277–310, 1993.

F. Forges. Correlated equilibrium in games with incomplete information revisited. *Theory and Decision*, 61:329–344, 2006.

S. Gleyze and A. Pernoud. Informationally simple incentives. *Journal of Political Economy*, forthcoming.

O. Gossner. Comparison of information structures. *Games and Economic Behavior*, 30: 44–63, 2000.

B. Hebert and J. La'O. Information acquisition, efficiency, and non-fundamental volatility. *Journal of Political Economy*, forthcoming.

C. Hellwig and L. Veldkamp. Knowing what others know: Coordination motives in information acquisition. *Review of Economic Studies*, 76:223–251, 2009.

O. Jann and C. Schottmüller. Correlated equilibria in homogeneous good Bertrand competition. *Journal of Mathematical Economics*, 57:31–37, 2015.

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

B. Kőszegi and F. Matějka. Choice simplification: A theory of mental budgeting and naive diversification. *Quarterly Journal of Economics*, 135:1153–1207, 2020.

E. Lehrer, D. Rosenberg, and E. Shmaya. Garbling of signals and outcome equivalence. *Games and Economic Behavior*, 81:179–191, 2013.

E. Lipnowski, L. Mathevet, and D. Wei. Attention management. *American Economic Review: Insights*, 2:17–32, 2020.

Q. Liu. Correlation and common priors in games with incomplete information. *Journal of Economic Theory*, 157:49–75, 2015.

B. Maćkowiak and M. Wiederholt. Optimal sticky prices under rational inattention. *American Economic Review*, 99:769–803, 2009.

B. Maćkowiak and M. Wiederholt. Business cycle dynamics under rational inattention. *Review of Economic Studies*, 82:1502–1532, 2015.

B. Maćkowiak, F. Matějka, and M. Wiederholt. Rational inattention: A review. *Journal of Economic Literature*, 61:226–273, 2023.

F. Matějka and A. McKay. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105:272–298, 2015.

J. Mensch. Screening inattentive buyers. *American Economic Review*, 112:1949–84, 2022.

J. Mensch and D. Ravid. Monopoly, product quality, and flexible learning. *arXiv preprint arXiv:2202.09985*, 2022.

S. Morris and M. Yang. Coordination and continuous stochastic choice. *Review of Economic Studies*, 89:2687–2722, 2022.

D. P. Myatt and C. Wallace. Endogenous information acquisition in coordination games. *Review of Economic Studies*, 79:340–374, 2012.

R. B. Myerson. Dual reduction and elementary games. *Games and Economic Behavior*, 21:183–202, 1997.

R. F. Nau and K. F. McCardle. Coherent behavior in noncooperative games. *Journal of Economic Theory*, 50:424–444, 1990.

N. Persico. Information acquisition in auctions. *Econometrica*, 68:135–148, 2000.

R. T. Rockafellar. *Convex analysis*. Princeton university press, 1970.

C. A. Sims. Implications of rational inattention. *Journal of Monetary Economics*, 50:665–690, 2003.

S. Van Nieuwerburgh and L. Veldkamp. Information immobility and the home bias puzzle. *Journal of Finance*, 64:1187–1215, 2009.

S. Van Nieuwerburgh and L. Veldkamp. Information acquisition and under-diversification. *Review of Economic Studies*, 77:779–805, 2010.

Y. Viossat. Is having a unique equilibrium robust? *Journal of Mathematical Economics*, 44:1152–1160, 2008.

W. Zhong. Optimal dynamic information acquisition. *Econometrica*, 90:1537–1582, 2022.