THE COMPARISON OF INFORMATION STRUCTURES IN GAMES: BAYES CORRELATED EQUILIBRIUM AND INDIVIDUAL SUFFICIENCY

By

Dirk Bergemann and Stephen Morris

September 2013

COWLES FOUNDATION DISCUSSION PAPER NO. 1909

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

The Comparison of Information Structures in Games: Bayes Correlated Equilibrium and Individual Sufficiency*

Dirk Bergemann† Stephen Morris‡

September 16, 2013

Abstract

We define and characterize a notion of correlated equilibrium for games with incomplete information, which we call *Bayes correlated equilibrium*. The set of outcomes that can arise in Bayes Nash equilibria of an incomplete information game where players may have access to additional signals beyond the information structure is characterized and shown to be equivalent to the set of *Bayes correlated equilibria*.

A game of incomplete information can be decomposed into a basic game, given by actions sets and payoff functions, and an information structure. We introduce a partial order on many player information structures - which we call *individual sufficiency* - under which more information shrinks the set of Bayes correlated equilibria. We discuss the relation of the solution concept to alternative definitions of correlated equilibrium in incomplete information games and of the partial order on information structures to others, including Blackwell's for the single player case.

Keywords: Correlated equilibrium, incomplete information, robust predictions, information structure, sufficiency, Blackwell ordering.
JEL Classification: C72, D82, D83.

*We acknowledge financial support through NSF Grants SES 0851200 and ICES 1215808. Earlier versions (Bergemann and Morris (2011)) were circulated under the title "Correlated Equilibrium in Games of Incomplete Information." We are grateful for discussions with Ben Brooks, with whom we are working on related projects, and to Hanming Fang and Ilya Segal, with whom we discussed related ideas for possible joint work in the past. We have benefited from comments from seminar participants and also from Mark Dean, Francoise Forges, Matthew Gentzkow, Olivier Gossner, Atsushi Kajii, Emir Kamenica, Qingmin Liu, Eran Shmaya, Bruno Strulovici, Jonathan Weinstein and research assistance from Constantinos Kalfarentzos, Alex Smolin and Áron Tóbiás.

†Department of Economics, Yale University, New Haven, U.S.A., dirk.bergemann@yale.edu.
‡Department of Economics, Princeton University, Princeton, U.S.A., smorris@princeton.edu.

1 Introduction

1.1 Motivation and Results

We investigate behavior in a given game of incomplete information, where the latter is described by a "basic game" and by an "information structure". The basic game refers to the set of actions, the set of payoff states, the utility functions of the players, and the common prior over the payoff states. The information structure refers to the type space of the game, which is generated by a mapping from the payoff states to a probability distribution over types, or signals. We ask what might happen in equilibrium if players may have access to additional signals beyond the given "information structure"? We show that behavior corresponds to a Bayes Nash equilibrium for some extra information that the players might observe if and only if it is an incomplete information version of correlated equilibrium that we dub *Bayes correlated equilibrium*. Aumann (1974), (1987) introduced the notion of correlated equilibrium in games with complete information and a number of definitions of correlated equilibrium in games with incomplete information have been suggested, notably in Forges (1993). Our definition is driven by a different motivation from the earlier literature and is weaker than the weakest definition of incomplete information correlated equilibrium (the Bayesian solution in Forges (1993)), because it allows play to be correlated with states that are not known by any player.

We define the Bayes correlated equilibrium as a mapping from the payoff states and type profile into a distribution over actions. The mapping, referred to as *decision rule*, is only required to satisfy an obedience condition for every pair of action and type of every player. There are a number of distinct reasons why the notion of Bayes correlated equilibrium - even though it is a straightforward variation of earlier definitions - as well as its subsequent characterization results are of particular interest. First, it allows the analyst to identify properties of equilibrium outcomes that are going to hold independent of features of the information structure that the analyst does not know; in this sense, properties that hold in all Bayes correlated equilibria of a given incomplete information game constitute robust predictions. Second, it provides a way to partially identify parameters of the underlying economic environment independently of knowledge of the information structure. Third, it provides an indirect method of identifying socially or privately optimal information structures without explicitly working with a space of all information structures. In Bergemann and Morris (2013), we illustrate these uses of the characterization result in a particular class of continuum player, linear best response games, focussing on normal distributions of types and actions and symmetric information structures and outcomes. While special, these games and equilibria can be used to model many economic phenomena of interest. In this paper, we work with general (finite player, finite action and finite state) games, and illustrate these uses with examples.

The separation between the basic game and the information structure enables us to ask how changes in the information structure affect the equilibrium set for a fixed basic game. A second contribution of the paper and the main formal contribution is that (i) we introduce a natural, statistical, partial order on information structures - called *individual sufficiency* - that captures intuitively when one information structure contains more information than another; and (ii) we show that the set of Bayes correlated equilibria shrinks in all games if and only if the informativeness of the information structure increases. Thus, if the information structure of the players contains more information, then the prediction of the analyst improves as a smaller set of outcomes is incentive compatible.

To describe our order on information structures, it is useful to note that a one player version of an information structure is an "experiment" in the sense studied by Blackwell (1951), (1953). An experiment consists of a set of signals and a mapping from states to probability distributions over signals. Suppose that we are interested in comparing a pair of experiments. A *combination* of the two experiments is a new experiment where a pair of signals - one from each experiment - is observed and the marginal probability over signals from each of the original experiments corresponds to the original distribution. One way of characterizing the classic sufficiency condition of Blackwell (1951) is the following: one experiment is *sufficient* for another if it is possible to construct a combined experiment such that the signal in the former experiment is a sufficient statistic for the signal in the latter experiment.

Our partial order on (many player) information structures is a natural generalization of sufficiency. One information structure is *individually sufficient* for another if there exists a combined information structure where each player's signal from the former information structure is a sufficient statistic for his beliefs over both states and others' signals in the combined information structure. This partial order has a couple of key properties - each generalizing well known properties in the one player case - that suggest that it is the "right" ordering on (many player) information structures. First, two information structures are individually sufficient for each other if and only if they are "higher-order belief equivalent" in the sense that they correspond to the same probability distribution over beliefs and higher-order beliefs about states (for any given prior on states). Second, one information structure is individually sufficient for another if and only if it is possible to start with the latter information structure and then have each player observe an extra signal, so that the expanded information structure is higher-order belief equivalent to the former information structure.

We introduce (in the many player case) an "incentive ordering" on information structures: an information structure is more *incentive constrained* than another if it gives rise to a smaller set of Bayes correlated equilibria. Our main result, stated in this language, is that one information structure is more incentive constrained than another if and only if the former is individually sufficient for the latter. Thus we show the equivalence between a statistical ordering and an incentive ordering. Blackwell's theorem showed that if one experiment was sufficient for another, then making decisions based on the former experiment allows a decision maker to attain a richer set of outcomes. Thus we will argue that Blackwell's theorem showed the equivalence of a "statistical ordering" on experiments (sufficiency) and a "feasibility ordering" (more valuable than). Our main result, restricted to the one person case, has a natural interpretation and shows an equivalence between a statistical ordering and an incentive ordering, and thus can be seen as an extension of Blackwell's theorem. To further understand the connection to Blackwell's theorem, we also describe a feasibility ordering on many player information structures which is equivalent to individual sufficiency and more incentive constrained than.

Taken together, our main result and discussion of the relation to Blackwell's theorem, highlight the dual role of information. By making more outcomes feasible, more information allows more outcomes to occur. By adding incentive constraints, more information restricts the set of outcomes that can occur. We show that the same partial order - individual sufficiency, reducing to sufficiency in the one player case - captures both roles of information simultaneously.

1.2 Related Literature

In a seminal paper, Hirshleifer (1971) showed how information might be damaging in a many player context because it removed options to insure ex ante. Our result on the incentive constrained ordering can be seen as a formalization of the idea behind the observation of Hirshleifer (1971): we give a general statement of how information creates more incentive constraints and thus reduces the set of incentive compatible outcomes. Importantly, in the one player context - but not in the many player context - more incentive constraints associated with additional information restrict the set of incentive compatible outcome, but never removes the most valuable action.

Our characterization result also has a one player analogue. Consider a decision maker who has access to an experiment, but may have access to more information. What can we say about the joint distribution of actions and states that might result in a given decision problem? We show that they are one person Bayes correlated equilibria. Such one person Bayes correlated equilibria have already arisen in a variety of contexts. Kamenica and Gentzkow (2011) consider the problem of cheap talk with commitment. In order to understand the behavior that a sender/speaker can induce a receiver/decision maker to choose, we might first want to characterize all outcomes that can arise for some committed cheap talk (independent of the objectives of the speaker). This, in our language, is the set of one person Bayes correlated equilibria. The earlier work of Aumann and Maschler (1995) on Nash equilibria of infinitely repeated zero sum games with one sided uncertainty and without discounting, a central result established that the outcome of the repeated game is as if the informed player can commit to reveal only certain information about the state in the corresponding static game. In turn, they showed that a concavification of the complete information payoff function yields the complete characterization of the set of feasible payoffs in the one player game of private information. In this sense, our analysis can be interpreted as bringing Aumann (1987) to environments with incomplete information by extending the analysis of Aumann and Maschler (1995) to many players and general, many player information structures.

Caplin and Martin (2013) study experiments with imperfect perception of a set of physical signals. Since they do not know how the decision maker perceives the state of the world, they interpret the subject as if she has observed more or less information unknown to the experimenter, and thus outcomes are, in our language, one person Bayes correlated equilibria.

There is a literature studying and comparing alternative definitions of correlated equilibrium under incomplete information, with the papers of Forges (1993), (2006) being particularly important. A standard assumption in that literature - which we dub "join feasibility" - is that play can only depend on the joint information of all the players. This restriction makes sense under the maintained assumption that correlated equilibrium is intended to capture the role of correlation of the players' actions but not unexplained correlation with the state of nature. Our different motivation leads us to allow such unexplained correlation. Liu (2011) also relaxes the join feasibility assumption, but imposes a belief invariance assumption (introduced and studied in combination with join feasibility in Forges (2006)), requiring that, from each player's point of view, the action recommendation that he receives from the mediator does not change his beliefs about others' types and the state. Intuitively, the belief invariant Bayes correlated equilibria of Liu (2011) capture the implications of common knowledge of rationality and a fixed information structure, while our Bayes correlated equilibria capture the implications of common knowledge of rationality and the fact that the player have observed at least the signals in the information structure.

Two papers - Lehrer, Rosenberg, and Shmaya (2010), (2013) - have examined comparative statics of how changing the information structure affects the set of predictions that can be made about players' actions, under Bayes Nash equilibrium and definitions of incomplete information correlated equilibrium stronger than Bayes correlated equilibrium. In the language of our paper, they construct statistical orderings on information structures and show how these orderings are relevant for - in Lehrer, Rosenberg, and Shmaya (2010) - feasibility orderings and - in Lehrer, Rosenberg, and Shmaya (2013) - incentive orderings. Our ordering - individual sufficiency - is a less incomplete variation on the orderings they construct. Importantly, all of their solution concepts maintain join feasibility. This has two important implications and leads to crucial differences in the results. First, as their solution concepts maintain join feasibility, their orderings are always refinements of sufficiency, i.e., they require players' joint (or pooled) information in one information structure to be sufficient for their joint (or pooled) information in the other structure, and then impose additional restrictions. By construction, individual sufficiency is a many player analogue of sufficiency but neither implies nor is implied by sufficiency of joint information. Second, because they work with solution concepts that include feasibility restrictions, the results relating information structure orders to incentive constraints in Lehrer, Rosenberg, and Shmaya (2013) are weaker than ours: they characterize when two information structures support the same set of equilibria in all games, thus they obtain equivalence classes, but not when one information structure supports a larger or smaller set, thus we obtain rankings.

The structure of the remaining paper is as follows. In Section 2, we define the notion of Bayes correlated equilibrium for a general finite game and establish the first main result, Theorem 1, namely the epistemic relationship between Bayes correlated equilibrium and Bayes Nash equilibrium. We illustrate the concepts with the analysis of a class of binary games. In Section 3, we offer a many player generalization of the sufficiency ordering of information structures, dubbed *individual sufficiency.* We also relate individual sufficiency to beliefs and higher-order beliefs, and illustrate the different notions with binary information structures. In Section 4 we present the second main result, Theorem 2, which establishes an equivalence between the incentive based ordering and the statistical ordering. We thus report results on comparing information structures in many player environments. In Section 5, we explain how the solution concept we dub "Bayes correlated equilibrium" relates to the literature, in particular Forges (1993), (2006) and conclude with a discussion of the relation to the large literature on the value of information in games.

2 Bayes Correlated Equilibrium

2.1 Definition

There are I players, $1, 2, ..., I$, and we write i for a typical player. There is a finite set of states, Θ, and we write θ for a typical state. A *basic game* G consists of (1) for each player i, a finite set of actions A_i,where we write $A = A_1 \times \cdots \times A_I$, and a utility function $u_i : A \times \Theta \to \mathbb{R}$; and (2) a full support prior $\psi \in \Delta_{++}(\Theta)$, Thus $G = \left((A_i, u_i)_{i=1}^I , \psi \right)$. An *information structure* S consists of (1) for each player i, a finite set of signals (or types) T_i,where we write $T = T_1 \times \cdots \times T_I$; and (2) a signal distribution $\pi : \Theta \to \Delta(T)$. Thus $S = \left((T_i)_{i=1}^I , \pi \right)$.

Together, the basic game G and the information structure S define a standard incomplete information game. While we use different notation, this division of an incomplete information game into the "basic game" and the "information structure" has been used in the literature (see, for example, Lehrer, Rosenberg, and Shmaya (2010)).

A *decision rule* in the incomplete information game (G, S) is a mapping σ :

$$\sigma : T \times \Theta \to \Delta(A). \tag{1}$$

One way to mechanically understand the notion of the decision rule is to view σ as the strategy of a mediator who first observes the realization of $\theta \in \Theta$ chosen according to ψ and the realization of $t \in T$ chosen according to $\pi(\cdot|\theta)$; and then picks an action profile $a \in A$, and privately announces to each player i the draw of a_i. For players to have an incentive to follow the mediator's recommendation in this scenario, it would have to be the case that the recommended action a_i was always preferred to any other action a_i' conditional on the signal t_i that player i had received and his knowledge of the recommended action a_i. This is reflected in the following "obedience" condition.

Definition 1 (Obedience)
Decision rule σ is obedient for (G, S) if, for each $i = 1, ..., I$, $t_i \in T_i$ and $a_i \in A_i$, we have

$$\begin{aligned} & \sum_{a_{-i}, t_{-i}, \theta} \psi(\theta)\, \pi((t_i, t_{-i})\,|\theta)\, \sigma((a_i, a_{-i})\,|\,(t_i, t_{-i}), \theta)\, u_i((a_i, a_{-i}), \theta) \\ \geq & \sum_{a_{-i}, t_{-i}, \theta} \psi(\theta)\, \pi((t_i, t_{-i})\,|\theta)\, \sigma((a_i, a_{-i})\,|\,(t_i, t_{-i}), \theta)\, u_i\left(\left(a_i', a_{-i}\right), \theta\right); \end{aligned} \tag{2}$$

for all $a_i' \in A_i$.

Our definition of Bayes correlated equilibrium requires obedience and nothing else. In particular, we shall not refer to or explicitly use a mediator in what follows.

Definition 2 (Bayes Correlated Equilibrium)
A decision rule σ is a Bayes correlated equilibrium (BCE) of (G, S) if it is obedient for (G, S).

If there is complete information, i.e., if Θ is a singleton, and S is the degenerate information structure where each player's signal set is a singleton, then this definition reduces to the Aumann (1987) definition of correlated equilibrium for a complete information game. We postpone until Section 5 a discussion of how this relates to (and why it is weaker than) other definitions in the literature on incomplete information correlated equilibrium. We provide our motivation for studying this particular definition next.

Consider an analyst who knew that

1. The basic game G describes actions, payoff functions depending on states, and a prior distribution on states.

2. The players observe at least information structure S, but may observe more.

3. The players' actions constitute a Bayes Nash equilibrium given the actual information structure.

What can she deduce about the joint distribution of actions, signals (in the minimal information structure, S) and states? We will formalize this question and show that all she can deduce is that the distribution will be a Bayes correlated equilibrium of (G, S).

To formalize this, we state the standard definition of Bayes Nash equilibrium in this setting. A (behavioral) strategy for player i in the incomplete information game (G, S) is $\beta_i : T_i \to \Delta(A_i)$.

Definition 3 (Bayes Nash Equilibrium)
A strategy profile β is a Bayes Nash equilibrium (BNE) of (G, S) if for each $i = 1, ..., I$, $t_i \in T_i$ and $a_i \in A_i$ with $\beta_i(a_i|t_i) > 0$, we have

$$\begin{aligned} & \sum_{a_{-i}, t_{-i}, \theta} \psi(\theta) \pi((t_i, t_{-i})|\theta) \left(\prod_{j \neq i} \beta_j(a_j|t_j) \right) u_i((a_i, a_{-i}), \theta) \\ \geq & \sum_{a_{-i}, t_{-i}, \theta} \psi(\theta) \pi((t_i, t_{-i})|\theta) \left(\prod_{j \neq i} \beta_j(a_j|t_j) \right) u_i\left((a_i', a_{-i}), \theta\right), \end{aligned} \tag{3}$$

for each $a_i' \in A_i$.

2.2 Foundations

We want to discuss situations where players observe more information than that contained in a given information structure. To formalize this, we must introduce combinations of information structures. If we have two information structures $S^1 = (T^1, \pi^1)$ and $S^2 = (T^2, \pi^2)$, we will say that information structure $S^* = (T^*, \pi^*)$ is a *combination* of information structures S^1 and S^2 if the combination, or *combined* information structure $S^* = (T^*, \pi^*)$ is obtained by forming a product space of the signals, $T_i^* = T_i^1 \times T_i^2$ for each i, and a likelihood function $\pi^* : \Theta \to \Delta(T_1 \times T_2)$ that preserves the marginal distribution of its constituent information structures.

Definition 4 (Combination)
The information structure $S^* = (T^*, \pi^*)$ is a combination *of information structures $S^1 = (T^1, \pi^1)$ and $S^2 = (T^2, \pi^2)$ if*

$$T_i^* = T_i^1 \times T_i^2 \text{ for each } i; \tag{4}$$

and

$$\begin{aligned} \sum_{t^2 \in T^2} \pi^*\left(t^1, t^2|\theta\right) &= \pi^1\left(t^1|\theta\right) \text{ for each } t^1 \in T^1 \text{ and } \theta \in \Theta; \\ \sum_{t^1 \in T^1} \pi^*\left(t^1, t^2|\theta\right) &= \pi^2\left(t^2|\theta\right) \text{ for each } t^2 \in T^2 \text{ and } \theta \in \Theta. \end{aligned} \tag{5}$$

Note that the above definition places no restrictions on whether signals $t^1 \in T^1$ and $t^2 \in T^2$ are independent or correlated, conditional on θ, under π^*. Thus any pair of information structures S^1 and S^2 will have many combined information structures.

Definition 5 (Expansion)
An information structure S^* is an expansion of S^1 if S^* is a combination of S^1 and some other information structure S^2.

Suppose strategy profile β was played in (G, S^*), where S^* is a combination of two information structures S^1 and S^2. Now, if the analyst did not observe the signals of the combined information structure S^*, but only the signals of S^1, then the behavior under the strategy profile β would induce a decision rule for $\left(G, S^1\right)$. Formally, the strategy profile β for (G, S^*) *induces* the decision rule σ for (G, S) given by:

$$\sigma\left(a|t^1, \theta\right) \triangleq \frac{\sum\limits_{t^2 \in T^2} \pi^*\left(t^1, t^2|\theta\right) \prod\limits_{j=1}^{I} \beta_j\left(a_j|t_j^1, t_j^2\right)}{\pi^1\left(t^1|\theta\right)},$$

for each $a \in A$ whenever $\pi^1\left(t^1|\theta\right) > 0$.

Theorem 1 (Epistemic Relationship)
A decision rule σ is a Bayes correlated equilibrium of (G, S) if and only if, for some expansion S^* of S, there is a Bayes Nash equilibrium of (G, S^*) which induces σ.

Thus this is an incomplete information analogue of the Aumann (1987) characterization of correlated equilibrium for complete information games. An alternative interpretation of this result - following Aumann (1987) - would be to say that BCE captures the implications of common certainty of rationality (and the common prior assumption) in the game G when players have at least information S, since requiring BNE in some game with expanded information is equivalent to describing a belief closed subset where the game G is being played, players have access to (at least) information S and there is common certainty of rationality. The proof follows the logic of the classic result of Aumann (1987) for complete information and that of Forges (1993) for the Bayesian solution for incomplete information games (discussed in Section 5).

Proof. Suppose that σ is a Bayes correlated equilibrium of (G, S). Thus

$$\begin{aligned} & \sum_{a_{-i}, t_{-i}, \theta} \psi(\theta) \pi\left(\left(t_i, t_{-i}\right)|\theta\right) \sigma\left(\left(a_i, a_{-i}\right) \mid \left(t_i, t_{-i}\right), \theta\right) u_i\left(\left(a_i, a_{-i}\right), \theta\right) \\ \geq & \sum_{a_{-i}, t_{-i}, \theta} \psi(\theta) \pi\left(\left(t_i, t_{-i}\right)|\theta\right) \sigma\left(\left(a_i, a_{-i}\right) \mid \left(t_i, t_{-i}\right), \theta\right) u_i\left(\left(a_i', a_{-i}\right), \theta\right) \end{aligned}$$

for each i, $t_i \in T_i$, $a_i \in A_i$ and $a_i' \in A_i$. Let $S^* = \left(\left(T_i^*\right)_{i=1}^{I}, \pi^*\right)$ be an expansion of S, and, in particular, a combination of $S = \left((T_i)_{i=1}^{I}, \pi \right)$ and $S' = \left((T'_i)_{i=1}^{I}, \pi' \right)$, where $T'_i = A_i$ for each i and π^* satisfies

$$\pi^* \left((t_i, a_i)_{i=1}^{I} \middle| \theta \right) = \pi \left(t \middle| \theta \right) \sigma \left(a \middle| t, \theta \right), \tag{6}$$

for each $t \in T$, $a \in A$ and $\theta \in \Theta$. Now, in the game (G, S^*), consider the "truthful" strategy β_j^* for player j, with

$$\beta_j^* \left(a'_j \middle| (t_j, a_j) \right) = \begin{cases} 1, & \text{if } a'_j = a_j, \\ 0, & \text{if } a'_j \neq a_j, \end{cases} \tag{7}$$

for all $t_j \in T_j$ and $a_j \in A_j$. Now the interim payoff to player i observing signal (t_i, a_i) and choosing action a'_i in (G, S^*) if he anticipates that each opponent will follow strategy β_j^* is

$$\begin{aligned}
& \sum_{a'_{-i}, a_{-i}, t_{-i}, \theta} \psi(\theta) \pi^* \left((t_i, t_{-i}), (a_i, a_{-i}) \middle| \theta \right) \left(\prod_{j \neq i} \beta_j^* \left(a'_j | t_j, a_j \right) \right) u_i \left(\left(a'_i, a'_{-i} \right), \theta \right), \text{ by (7)} \\
= & \sum_{a_{-i}, t_{-i}, \theta} \psi(\theta) \pi^* \left((t_i, t_{-i}), (a_i, a_{-i}) \middle| \theta \right) u_i \left(\left(a'_i, a_{-i} \right), \theta \right), \text{ by (6)} \\
= & \sum_{a_{-i}, t_{-i}, \theta} \psi(\theta) \pi \left((t_i, t_{-i}) | \theta \right) \sigma \left((a_i, a_{-i}) | (t_i, t_{-i}), \theta \right) u_i \left(\left(a'_i, a_{-i} \right), \theta \right),
\end{aligned}$$

by (6) and (7), and thus Bayes Nash equilibrium optimality conditions for the truth telling strategy profile β^* are implied by the obedience conditions on σ. By construction, β induces σ.

Conversely, suppose that β is a Bayes Nash equilibrium of (G, S^*), where S^* is a combined experiment for S and S'. Write $\sigma : T \times \Theta \to \Delta(A)$ for the decision rule for (G, S) induced by β, so that

$$\pi (t|\theta) \sigma (a|t, \theta) = \sum_{t' \in T'} \pi^* \left((t_i, t'_i)_{i=1}^{I} \middle| \theta \right) \prod_{j=1}^{I} \beta_j \left(a_j | t_j, t'_j \right)$$

for each $t \in T$, $a \in A$ and $\theta \in \Theta$. Now $\beta_i \left(a_i | (t_i, t'_i) \right) > 0$ implies

$$\begin{aligned}
& \sum_{a_{-i}, t_{-i}, t'_{-i}, \theta} \psi(\theta) \pi^* \left((t_i, t'_i)_{i=1}^{I} \middle| \theta \right) \left(\prod_{j \neq i} \beta_j \left(a_j | t_j, t'_j \right) \right) u_i \left((a_i, a_{-i}), \theta \right) \\
\geq & \sum_{a_{-i}, t_{-i}, t'_{-i}, \theta} \psi(\theta) \pi^* \left((t_i, t'_i)_{i=1}^{I} \middle| \theta \right) \left(\prod_{j \neq i} \beta_j \left(a_j | t_j, t'_j \right) \right) u_i \left(\left(a'_i, a_{-i} \right), \theta \right),
\end{aligned}$$

for each i, $t_i \in T_i$, $t'_i \in T'_i$ and $a'_i \in A_i$. Thus

$$\begin{aligned}
& \sum_{t'_i} \beta_i \left(a_i | (t_i, t'_i) \right) \sum_{a_{-i}, t_{-i}, t'_{-i}, \theta} \psi(\theta) \pi^* \left((t_i, t'_i)_{i=1}^{I} \middle| \theta \right) \left(\prod_{j \neq i} \beta_j \left(a_j | t_j, t'_j \right) \right) u_i \left((a_i, a_{-i}), \theta \right) \\
\geq & \sum_{t'_i} \beta_i \left(a_i | (t_i, t'_i) \right) \sum_{a_{-i}, t_{-i}, t'_{-i}, \theta} \psi(\theta) \pi^* \left((t_i, t'_i)_{i=1}^{I} \middle| \theta \right) \left(\prod_{j \neq i} \beta_j \left(a_j | t_j, t'_j \right) \right) u_i \left(\left(a'_i, a_{-i} \right), \theta \right),
\end{aligned}$$

for each i, $t_i \in T_i$ and $a_i' \in A_i$. But

$$\sum_{t_i'} \beta_i \left(a_i | \left(t_i, t_i'\right)\right) \sum_{a_{-i}, t_{-i}, t_{-i}', \theta} \psi\left(\theta\right) \pi^* \left(\left(t_i, t_i'\right)_{i=1}^{I} \Big| \theta \right) \left(\prod_{j \neq i} \beta_j \left(a_j | t_j, t_j'\right) \right) u_i \left(\left(a_i', a_{-i}\right), \theta \right)$$
$$= \sum_{a_{-i}, t_{-i}, \theta} \psi\left(\theta\right) \pi\left(\left(t_i, t_{-i}\right) | \theta\right) \sigma\left(\left(a_i, a_{-i}\right) | \left(t_i, t_{-i}\right), \theta\right) u_i \left(\left(a_i', a_{-i}\right), \theta \right),$$

and thus BNE equilibrium conditions (G, S^*) imply obedience conditions of σ for (G, S). ■

2.3 A Class of Binary Games

We now present a simple class of binary games, with an intuitive economic interpretation, to illustrate the structure of Bayes correlated equilibria and Theorem 1. In particular, we identify, in the class of games, the expanded information structures that support or "decentralize" welfare maximizing Bayes correlated equilibria as Bayes Nash equilibria. We use a "$2 \times 2 \times 2$” basic game, where there are two players, two actions for each player and two states. We also use a one-dimensional family of binary information structures with public signals. And we restrict attention to a two dimensional class of (symmetric) decision rules. We find it useful to restrict attention to these classes because we can visualize the results using pictures. Calvó-Armengol (2006) showed that - even in complete information games - characterizing and visualizing all correlated equilibria of all two player two actions games is not easy. We emphasize that we are using this class of examples to illustrate results that apply to general, asymmetric, information structures and general, asymmetric, decision rules.

Each player can either invest, $a = I$ or not invest, $a = N$ and the payoffs are given in the bad state θ_B and the good state θ_G by the following matrices:

θ_B	I	N
I	$z - 1 + y_B, z - 1 + y_B$	$-1, z$
N	$z, -1$	$0, 0$

θ_G	I	N
I	$z + 1 + y_G, z + 1 + y_G$	$1, z$
N	$z, 1$	$0, 0$

. (8)

The payoffs are symmetric across players and have three components: (i) there is a payoff 1 to invest in the good state θ_G and a payoff -1 to invest in the bad state θ_B; (ii) there is always an externality $z > 0$ if the other player invests, and (iii) there is an additional, possibly state dependent payoff y_j, $j = B, G$, to invest if the other player invests as well. The payoff y_j can be positive or negative, but of uniform sign across states, leading to a game with strategic complements or substitutes, respectively. We will focus on the case where $z \gg 1$ and $y_j \approx 0$.[1] Thus, if the players were to know the state, i.e. under complete information, then each player would have a strict dominant strategy to invest in θ_G and not to invest in

[1] Formally, we require that $z > 1$ and that $z > 1 - 2y_j$ for $j = B, G$.

θ_B. Importantly, given that the externality z is assumed to be large, i.e. $z \gg 1$, the sum of the payoffs is maximized if both players invest in both states, θ_B and θ_G. Notice that z is a pure externality that influences players' utilities but not the best responses and a fortiori not the set of BCE. Finally, we assume that state θ_G occurs with probability ψ, while state θ_B occurs with probability $1-\psi$.

We consider a binary information structure S where, if the state is bad, each player observes a signal t_b, saying that the state is bad, with probability q. If a player doesn't receive the signal t_b, then he receives the signal t_g, and thus the signal t_g is always observed in the good state. The signal distribution $\pi : \Theta \to T$ is given by:

$\pi(\cdot \mid \theta_B)$	t_b	t_g
t_b	q	0
t_g	0	$1-q$

$\pi(\cdot \mid \theta_G)$	t_b	t_g
t_b	0	0
t_g	0	1

. (9)

Each player observes his signal realization privately but the signal realizations are perfectly correlated. The information structure is thus symmetric across players, but not across states. In particular, the conditional probability q is a measure of the accuracy of the information structure. An increase in q leads, after a realization of t_g to a strict increase in the posterior probability that the state is θ_G, and after a realization of t_b the posterior probability that the state is θ_B is always 1 (and thus is *weakly* increasing in q).

We restrict attention to decision rules σ, as defined earlier in (1), that are symmetric across players. Accordingly, we must specify the action profile for each state-signal profile (θ, t). After observing the negative signal t_b, each player knows that the state is θ_B and has a strictly dominant strategy to choose N, so we will take this behavior as given. We can parameterize the symmetric (across players) decision rule σ conditional on the positive signals t_g and the state θ_j, for $j = B, G$, by:

$\sigma(\theta_j, t_g)$	I	N
I	γ_j	$\alpha_j - \gamma_j$
N	$\alpha_j - \gamma_j$	$\gamma_j + 1 - 2\alpha_j$

. (10)

We thus have four parameters, $\alpha_B, \alpha_G, \gamma_B, \gamma_G$, where α_j is the probability that any one player invests in state θ_j and γ_j is the probability that both players invest under the nonnegativity restrictions:

$$\alpha_j \geq 0, \gamma_j \geq 0, \text{ and } 2\alpha_j - 1 \leq \gamma_j \leq \alpha_j\,, \quad \text{for } j = B, G. \tag{11}$$

The set of parameters $(\alpha_B, \alpha_G, \gamma_B, \gamma_G)$ which form a Bayes correlated equilibrium are completely characterized by the obedience conditions for $a = I, N$. Thus, explicitly, if a player is observing signal t_g and is advised to invest, then he will invest if:

$$\psi\left(\alpha_G + \gamma_G y_G\right) + (1-\psi)(1-q)\left(-\alpha_B + \gamma_B y_B\right) \geq 0; \tag{12}$$

and a player advised to not invest will not invest if:

$$\psi\left(1-\alpha_G+\left(\alpha_G-\gamma_G\right)y_G\right)+\left(1-\psi\right)\left(1-q\right)\left(-\left(1-\alpha_B\right)+\left(\alpha_B-\gamma_B\right)y_B\right)\leq 0. \tag{13}$$

We will focus on the characterization of the "second-best BCE" which maximizes the sum of players' utility subject to being a BCE and then describe the expanded information structures that can achieve the Bayes correlated equilibrium as a Bayes Nash equilibrium. To this end, it suffices to identify the parameters $(\alpha_B, \alpha_G, \gamma_B, \gamma_G)$ that maximize the expected utility of a (representative) player:

$$\psi\left(\alpha_G\left(z+1\right)+\gamma_G y_G\right)+\left(1-\psi\right)\left(1-q\right)\left(\alpha_B\left(z-1\right)+\gamma_B y_B\right) \tag{14}$$

subject to the obedience conditions (12) and (13) and the nonnegativity restrictions (11). In the analysis it will prove useful to distinguish between the strategic complements, $y_j \geq 0$, and strategic substitutes, $y_j \leq 0$.

Strategic Complements We begin with strategic complements. As a player never invests after observing the negative signal t_b, after correctly inferring that the state is θ_B, we immediately ask under what conditions investment can occur after the realization of positive signal t_g. If investment could always be achieved, independent of the true state, then the resulting decision rule σ would have $\alpha_G = \gamma_G = \alpha_B = \gamma_B = 1$, and inserting these values in the obedience constraint for investing, see (12), yields:

$$\psi\left(1+y_G\right)+\left(1-\psi\right)\left(1-q\right)\left(y_B-1\right)\geq 0 \quad \Leftrightarrow \quad q\geq 1-\frac{\psi}{1-\psi}\frac{1+y_G}{1-y_B}. \tag{15}$$

Thus, if the information structure S, as represented by q, is sufficiently accurate, then investment following the realization of the signal t_g can be achieved with probability one. In fact, the above condition (15) is a necessary and sufficient condition for a Bayes Nash equilibrium with investment after the signal t_g to exist. Hence, we know that this decision rule can be informationally decentralized without any additional information if q is sufficiently large.

By contrast, if q fails to satisfy the condition (15), then the second-best BCE is to maintain investment in the good state: $\alpha_G = \gamma_G = 1$, while maximizing the probability of investment α_B in the bad state subject to the obedience constraint (12). The no investment constraint (13) will automatically be satisfied. In a game with strategic complements, this is achieved by coordinating investments, i.e. setting $\alpha_B = \gamma_B$ and satisfying (12) as an equality:

$$\psi\left(1+y_G\right)-\left(1-\psi\right)\left(1-q\right)\alpha_B\left(1-y_B\right)=0 \quad \Leftrightarrow \quad \alpha_B=\gamma_B=\frac{\psi\left(1+y_G\right)}{\left(1-\psi\right)\left(1-q\right)\left(1-y_B\right)}. \tag{16}$$

Now, we observe that the solution (16) requires the probabilities to differ across the states, or $\alpha_B = \gamma_B < \alpha_G = \gamma_G = 1$. It follows that this decision rule requires additional information, and hence an expansion of the information structure S for it to be decentralized as a Bayes Nash equilibrium. The necessary expansion is achieved by two additional signals t'_b, t'_g which lead to an expansion S^* and an associated likelihood function $\pi^*(t, t' | \theta)$ as displayed below:

$\pi^*(\cdot \| \theta_B)$	t_b, t'_b	t_g, t'_b	t_g, t'_g
t_b, t'_b	q	0	0
t_g, t'_b	0	r	0
t_g, t'_g	0	0	$1-q-r$

$\pi^*(\cdot \| \theta_G)$	t_b, t'_b	t_g, t'_g
t_b, t'_b	0	0
t_g, t'_g	0	1

.

We observe that the expansion preserves the public nature of the signals, in that the realizations remain perfectly correlated across the players. The additional signals confirm the original signals everywhere except for the pair (t_g, t'_b) which changes the posterior of each player to a probability one belief that the state is θ_B. In other words, the additional signals t'_b, t'_g "split" the posterior conditional on receiving t_g in the information structure S. We can readily compute the minimal probability that the public signal (t_g, t'_b) has to have so that in the associated BNE the players invest with probability one after receiving the signal $\left(t_g, t'_g\right)$, namely by requiring that the best response for investment is met as an equality in the BNE:

$$\psi(1+y_G) - (1-\psi)(1-q-r)(1-y_B) = 0 \Leftrightarrow r = 1 - q - \frac{\psi(1+y_G)}{(1-\psi)(1-y_B)}.$$

Strategic substitutes Next, we discuss the game with strategic substitutes, $y_j \leq 0$. While the basic equilibrium conditions remain unchanged, the information structures that decentralize the second-best BCE have very different properties with strategic substitutes. In particular, private rather than public signals become necessary to decentralize the decision rule σ as a Bayes Nash equilibrium.

To begin with, just as in the case of strategic complements, if the information structure S, as represented by q, is sufficiently accurate, then investment following the realization of the signal t_g can be achieved with probability one, this is the earlier condition (15). Similarly, if q fails to satisfy the condition (15) then the second-best BCE is to maintain investment in the good state: $\alpha_G = \gamma_G = 1$, while maximizing the probability of investment α_B in the bad state subject to the obedience constraint (12). But importantly, in a game with strategic substitutes, the obedience constraint is maintained by minimizing the probability of joint investments, hence *minimizing* γ_B. In terms of the decision rule $\sigma(\cdot, t_g)$ as represented in the matrix (10), we seek to place most probability off the diagonal, in which only one, but not both players, invest. If there is substantial slack in the obedience constraint (12), then the residual probability can lead to investment by both players, but if there is little slack, then it will require that the residual probability leads to no investment by either player, which suggests a second threshold for q, below the one established in (15).

Thus if condition (15) fails, then it is optimal to maximize α_B and minimize γ_B, where the later is constrained by the nonnegativity restrictions of (10): $\gamma_B = \max\{0, 2\alpha_B - 1\}$. Thus we want α_B to solve the obedience constraint for investment, (12), as an equality:

$$\psi(1+y_G) + (1-\psi)(1-q)(-\alpha_B + \max(2\alpha_B - 1, 0)\, y_B) \geq 0 \tag{17}$$

This leads to a strictly positive solution of γ_B, the probability of joint investment, as long as the probability q is not too low, or

$$1 - \frac{2\psi}{1-\psi}(1+y_G) \leq q \leq 1 - \frac{\psi}{1-\psi}\frac{1+y_G}{1-y_B}, \tag{18}$$

and the second-best decision rule given by:

$$\alpha_G = \gamma_G = 1, \alpha_B = \frac{1}{1-2y_B}\left(\frac{\psi(1+y_G)}{(1-\psi)(1-q)} - y_B\right),\ \gamma_B = 2\alpha_B - 1. \tag{19}$$

Finally, if q falls below the lower threshold established in (18), then the second-best decision rule σ prescribes investment only by one player, but never by both players simultaneously:

$$\alpha_G = \gamma_G = 1, \quad \alpha_B = \frac{\psi(1+y_G)}{(1-\psi)(1-q)},\ \gamma_B = 0. \tag{20}$$

As expected, we find that both the probability of investment by a player, given by α_B, as well as the probability of a joint investment, γ_B, are increasing in the accuracy q.

We ask again which expanded information structures decentralize these second-best decision rules. As $\gamma_B < \alpha_B$, the decision rule σ requires with positive probability investment by one player only. This can only be achieved by private signals that lead to distinct choices by the players with positive probability. The expansion can still be achieved with two additional signals, t'_b, t'_g, and as before the additional signals refine or split the posterior that each player held at t_g in the information structure S. But importantly, now the signal realizations cannot be perfectly correlated across the players anymore. Thus if q is not too low, i.e. condition (18) prevails, then the following information structure decentralizes the second-best BCE:

$\pi^*(\cdot\,\vert\theta_B)$	t_b, t'_b	t_g, t'_b	t_g, t'_g
t_b, t'_b	q	0	0
t_g, t'_b	0	0	r
t_g, t'_g	0	r	$1-q-2r$

$\pi^*(\cdot\,\vert\theta_G)$	t_b, t'_b	t_g, t'_g
t_b, t'_b	0	0
t_g, t'_g	0	1

; (21)

and by contrast if q is sufficiently low, i.e. below the lower bound of (18), then the expanded information structure below decentralizes the BCE:

$\pi^*(\cdot\,\vert\theta_B)$	t_b, t'_b	t_g, t'_b	t_g, t'_g
t_b, t'_b	q	0	0
t_g, t'_b	0	$1-q-2r$	r
t_g, t'_g	0	r	0

$\pi^*(\cdot\,\vert\theta_G)$	t_b, t'_b	t_g, t'_g
t_b, t'_b	0	0
t_g, t'_g	0	1

. (22)

In either case, the expansion requires private signals in the sense that conditional on receiving a given signal, either $\left(t_g, t'_g\right)$ or (t_g, t'_b) respectively, each player remains uncertain as to the signal received by the other player, i.e. either (t_g, t'_b) or $\left(t_g, t'_g\right)$. As required, the expanded information structure S^* preserves the likelihood distribution ψ of the initial information structure S.[2]

The above analysis focussed on second-best Bayes correlated equilibria that maximize welfare. We now visualize all symmetric Bayes correlated equilibria in a special case. We stay with a game of strategic substitutes, $y_j \leq 0$ and the illustrations below are computed for the prior probability of the good state, $\psi = 1/3$ and $z = 2, y_G = 0, y_B = -1/6$. Because there is never investment conditional on bad signals, it is enough the focus on the probabilities α_G and α_B that any player invests, conditional on good and bad states respectively, after observing the positive signal t_g.[3] Figures 1 through 3 show the set of all values of α_G and α_B corresponding to symmetric BCE for low, intermediate and high levels of accuracy q, namely $q = 1/5, 11/20$ and $4/5$, respectively.

For all values of $q \in [0, 1]$, the action profile that maximizes the sum of the payoffs is $\alpha_B = \alpha_G = 1$, the first-best action profile. Every Bayes Nash equilibrium under the given information structure S has to be located on the 45° line, as each player cannot distinguish between the states θ_B and θ_G conditional on t_g. In fact, the Bayes Nash equilibrium in the game with strategic substitutes is unique for all levels of q, and depending on the accuracy q, it is either a pure strategy equilibrium with no investment as in Figure 1, a mixed strategy equilibrium with positive probability of investment as in Figure 2, or a pure strategy equilibrium with investment as in Figure 3, respectively. By contrast, the second-best BCE, as computed by (20), always yields a strictly positive level of investment in the bad state θ_B, and one that is strictly higher than in any BNE, unless the BNE itself is a pure strategy equilibrium with investment (following t_g), see Figure 3.

If we consider a intermediate level of accuracy q, rather than a low level of accuracy q, as in Figure 2, then we find that there is unique mixed BNE which provides investment with positive probability following t_G. The BNE is therefore in the interior of the unit square of conditional investment probabilities (α_G, α_B). By contrast, the second-best BCE remains at the exterior of the unit square, and yields a strictly higher probability of investment in the bad state than the corresponding Bayes Nash equilibrium. Interestingly, the BNE is in the interior of the set of BCE, when expressed in the space of investment probabilities rather

[2] An interesting question that we do not explore in any systematic manner in this paper is what we can say about the relation between Bayes correlated equilibria and the expansions that are needed to support them as Bayes Nash equilibria. Milchtaich (2012) examines properties of devices needed to implement correlated equilibria, and tools developed in his paper might be useful for this task.

[3] A complete description of the BCE would also include the probabilities $\gamma_j, j = B, G$ of joint investment, but for the present purpose the two-dimensional graph of α_B and α_G shall suffice. The computation of the complete characterization is recorded in the Appendix.

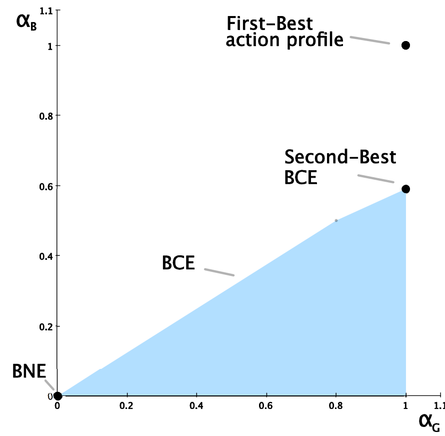

Figure 1: BNE and set of BCE with low accuracy: $q = 1/5$.

than an extreme point of the set of BCE.If the accuracy of the information structure increases even further,

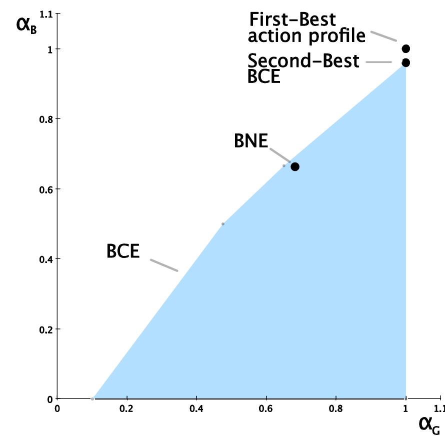

Figure 2: BNE and set of BCE with intermediate accuracy: $q = 11/20$.

see Figure 3, then conditional on receiving the positive signal t_G, it is sufficiently likely that the state is θ_G, that investment occurs with probability one even in the Bayes Nash equilibrium. Essentially, the high probability of θ_G (and resulting high payoffs from investment) more than offset the low probability of θ_B (and resulting low payoffs from investment).

This first set of illustrations depict the probabilities of investment conditional on the realization of the positive signal t_g *and* the state θ_j, $j = B, G$. But as we vary the accuracy q, we are changing the probability of the signal t_g, and hence the above figures do not directly represent the probabilities of investment β_j conditional on the state θ_j only, which are simply given by $\beta_B = (1 - q)\,\alpha_B$ and $\beta_G = \alpha_G$. The resulting sets of investment probabilities are depicted in Figure 4, for all three levels of q. The set of BCE is shrinking as the information structure S, as represented by q, becomes more accurate. This comparative

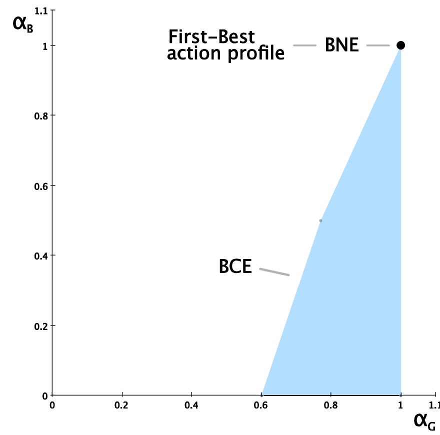

Figure 3: BNE and set of BCE with hiqh accuracy: $q = 4/5$.

static illustrates Theorem 2 in the next section. Because the set of BCE is shrinking, the best achievable BCE welfare is necessarily getting weakly lower with more information and, in this example, is strictly lower. On the other hand, as q increases, welfare in BNE will increase over some ranges.

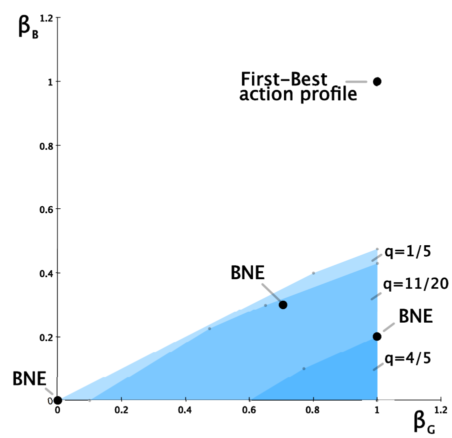

Figure 4: Set inclusion of BCE with increasing information

3 A Partial Order on Information Structures

We will define and describe statistical properties of, a partial order on information structures, which we call "individual sufficiency". We think that this order captures an intuitive statistical notion of "more informative" that might be important in a wide variety of contexts. In this Section, we define the ordering and discuss some of its properties. In the next Section, we give the main result of this paper, showing that this ordering characterizes when one information structure generates a smaller set of Bayes correlated equilibria by adding more incentive constraints. We postpone until Section 5 a detailed discussion of the relation to Blackwell's Theorem and other related work.

We say that one information structure, S, is "individually sufficient" for another, S', if there is a "combined information structure," S^*, in which both information structures can be embedded where, for every player i, player i's signal t_i' in information structure S' gives no additional information - beyond that contained in his signal t_i in information structure S - about the state, θ, and the signals of other players, t_{-i}, under S. We label the partial order "individual sufficiency" because, player by player, information structure S is sufficient for information structure S'.

The ordering depends only on the probability distribution over beliefs and higher-order beliefs about states that it induces. Thus the set of information structures which generate the same probability distribution over higher-order beliefs are an equivalence class under our order. Liu (2011) gives a characterization of this equivalence relation, so relative to Liu (2011), our contribution is to extend the equivalence relation to a partial order on all information structures.

Blackwell (1951), (1953) introduced a well known sufficiency ordering on experiments. Say that an information structure, S, with possibly many players and thus possibly multi-dimensional signals, is *sufficient* for another, S', if the pooled information of the players under S is sufficient in Blackwell's sense for the *pooled* information in S'. If there is only one player, our individual sufficiency order is equivalent to sufficiency and, in this sense, individual sufficiency is a many player generalization of Blackwell's order on experiments, or single player information structures. Lehrer, Rosenberg, and Shmaya (2010), (2013) introduced a family of partial orders on information structures which refine the pooled information version of sufficiency defined above, and establish important game theoretical properties of these orders, which we review in Section 5.4. We will show that individual sufficiency neither implies nor is implied by the many player version of sufficiency. Individual sufficiency is what you get if you relax the maintained assumption of (pooled information) sufficiency in Lehrer, Rosenberg, and Shmaya (2010), (2013).

Before formally defining our ordering, it is useful to highlight a statistical property that plays a key role in our definitions and results, as it has in earlier related work.

3.1 A Statistical Digression: Blackwell Triples

Suppose that we are interested in a triple of random variables, $(x, y, z) \in X \times Y \times Z$, and that we are given a probability distribution on the product space, $P \in \Delta(X \times Y \times Z)$. We will abuse notation by using P to refer both to marginal probabilities, writing $P(x)$ for $P(\{x\} \times Y \times Z)$ and $P(x, y)$ for $P(\{x\} \times \{y\} \times Z)$;

and conditional probabilities, writing

$$P(x|y,z) \triangleq \frac{P(x,y,z)}{P(y,z)},$$

if $P(y,z) > 0$; and

$$P(x|y) \triangleq \frac{P(x,y)}{P(y)},$$

if $P(y) > 0$. We will say that the probability of x conditional on y is independent of z (under P) if

$$P(x|y,z) = P(x|y),$$

for all $z \in Z$ whenever $P(y,z) > 0$. Now we have the following statistical fact concerning a triple of random variables.

Lemma 1 (Conditional Independence)
The following statements are equivalent if $P(x,y,z) > 0$ for every x, y, z.

1. *The probability $P(x|y,z)$ is independent of z.*

2. *The probability $P(z|y,x)$ is independent of x.*

3. *$P(x,y,z) = P(y)\,P(x|y)\,P(z|y)$.*

Proof. (3) immediately implies (1) and (2). To see that (1) implies (3), observe that if $P(y,z) > 0$,

$$\begin{aligned} P(y,z)\,P(x|y,z) &= P(y)\,P(z|y)\,P(x|y,z)\,, \text{ by definition} \\ &= P(y)\,P(z|y)\,P(x|y)\,, \text{ by (1).} \end{aligned}$$

A symmetric argument shows that (2) implies (3). ■

When these statements are true, we will say that the ordered triple (x,y,z) is a *Blackwell triple* (under P). Blackwell (1951) observed that the above relationship can be rephrased as saying that a Markov chain, namely $P(x|y,z) = P(x|y)$, is also a Markov chain in reverse, namely $P(z|y,x) = P(z|y)$.[4]

3.2 Three Equivalent Definitions of Individual Sufficiency

We defined "combined information structures" above in Definition 4. We say that information structure S is *individually sufficient* for information structure S' if there exists a combined information structure under which, for each i, $(t'_i, t_i, (t_{-i}, \theta))$ is a Blackwell triple. Thus more completely,

[4]For this reason, Torgersen (1991) p. 345, refers to the triple (x,y,z) as a "Markov triple". As the term "Markov triple" is commonly used to refer to solutions of Markov Diophantine equations, we prefer "Blackwell triple".

Definition 6 (Individual Sufficiency)

Information structure $S = (T, \pi)$ is individually sufficient for information structure $S' = (T', \pi')$ if there exists a combined information structure $S^* = (T^*, \pi^*)$ such that, for each i,

$$\Pr\left(t_i'|t_i, t_{-i}, \theta\right) \triangleq \frac{\sum_{t_{-i}'} \pi^*\left(\left(t_i, t_{-i}\right), \left(t_i', t_{-i}'\right)|\theta\right)}{\sum_{\widetilde{t}_i', t_{-i}'} \pi^*\left(\left(t_i, t_{-i}\right), \left(\widetilde{t}_i', t_{-i}'\right)|\theta\right)} \tag{23}$$

is independent of t_{-i} and θ.

Via Lemma 1 above, an equivalent way of defining individual sufficiency is that observing t_i' gives no additional information about t_{-i} and θ beyond that contained in t_i. In particular, (23) is equivalent to the claim that, for some (or every) full support prior on the state space Θ, $\psi \in \Delta_{++}(\Theta)$, and for each i,

$$\Pr\left(t_{-i}, \theta|t_i, t_i'\right) \triangleq \frac{\sum_{t_{-i}'} \psi(\theta)\, \pi^*\left(\left(t_i, t_{-i}\right), \left(t_i', t_{-i}'\right)|\theta\right)}{\sum_{\widetilde{t}_{-i}, \widetilde{\theta}, t_{-i}'} \psi\left(\widetilde{\theta}\right) \pi^*\left(\left(t_i, \widetilde{t}_{-i}\right), \left(t_i', t_{-i}'\right)|\widetilde{\theta}\right)} \tag{24}$$

is independent of t_i'. Finally, we can equivalently define individual sufficiency by asking if there exists a Markov kernel $\phi : T \times \Theta \to \Delta(T')$ such that if we start with the information structure S and apply the Markov kernel ϕ, we end up with a combined information structure under which, for each i, $(t_i', t_i, (t_{-i}, \theta))$ is a Blackwell triple.[5] Thus information structure S is *individually sufficient* for S' if there exists $\phi : T \times \Theta \to \Delta(T')$ satisfying

$$\sum_{t \in T} \pi(t|\theta)\, \phi\left(t'|t, \theta\right) = \pi'\left(t'|\theta\right) \tag{25}$$

for each t' and θ, and if

$$\sum_{t_{-i}' \in T_{-i}'} \phi\left(\left(t_i', t_{-i}'\right) | \left(t_i, t_{-i}\right), \theta\right) \tag{26}$$

is independent of t_{-i} and θ.

In particular, if the kernel ϕ satisfies the above independence, (26), for every player i, then we can define for every player i an individual (or marginal) kernel $\phi_i : T_i \to \Delta(T_i')$ by setting:

$$\phi_i\left(t_i'|t_i\right) \triangleq \sum_{t_{-i}' \in T_{-i}'} \phi\left(\left(t_i', t_{-i}'\right) | \left(t_i, t_{-i}\right), \theta\right). \tag{27}$$

[5] The equivalence is established by using the combined information structure $\pi^* : \Theta \to \Delta(T \times T')$ to generate the Markov kernel $\phi : \Theta \times T \to \Delta(T')$ by conditioning on θ and t, and conversely to use the Markov kernel ϕ to generate the combined information structure in the natural way. The conditional independence condition (23) and the conditional Markov property (26) express the same property either in terms of π^* or ϕ respectively.

It is useful to contrast the *individual sufficiency* with Blackwell's original notion of *sufficiency*, applied to the players' pooled information, that is the type profiles t and t'. Sufficiency can then be stated as the requirement that (t', t, θ) form a Blackwell triple. Using the celebrated Markov kernel statement of this requirement, information structure S is *sufficient* for information structure S' if there exists $\phi : T \to \Delta(T')$ such that

$$\sum_{t \in T} \pi(t|\theta)\, \phi(t'|t) = \pi'(t'|\theta), \tag{28}$$

for each t' and θ, and thus π' is frequently said to be *reproducible* from π and ϕ, see LeCam (1996).

Now, in the case of single player, $i = I = 1$, individual sufficiency reduces to sufficiency: in this case, condition (26) reduces to the requirement that $\phi(t'|t, \theta)$ is independent of θ and thus condition (25) becomes condition (28). Alternatively, using the language of triples, the condition of individual sufficiency reduces with a single player i to the requirement that (t'_i, t_i, θ) form a Blackwell triple.

By contrast, here we are interested in a many player version of sufficiency, and thus each individual player forms beliefs about the payoff state θ *and* the type profile t_{-i} of the other players. We thus augment the payoff state θ by the type profile t_{-i} of the other players, as player i is concerned about θ and t_{-i}. We are therefore lead to the stronger requirement that $(t'_i, t_i, (\theta, t_{-i}))$ form a Blackwell triple.[6]

Lehrer, Rosenberg, and Shmaya (2010), (2013) introduced a partial ordering on many player information structures, "non-communicating garbling", which is equivalent to individual sufficiency with the additional requirement that ϕ is independent of θ. Thus non-communicating garbling refines sufficiency in the many player case. Individual sufficiency does not imply sufficiency in the many player case. Individual sufficiency is the relaxation of non-communicating garbling obtained if sufficiency is not maintained in the many player case.[7]

Liu (2011) introduced a natural definition of a correlation device in a many player context: if you fix an information structure S and signal sets $(T'_i)_{i=1}^{I}$ that players might observe, Liu (2011) said that a mapping $\phi : T \times \Theta \to \Delta(T')$ was a correlation device if it satisfied (26). Thus we can re-state our definition of

[6] We note that there is a certain asymmetry in the triple $(t'_i, t_i, (\theta, t_{-i}))$, as it only refers to the type profile t_{-i} of the information structure S which is individually sufficient for S', and does not refer at all to t'_{-i}. In fact, the certainly more symmetric condition - that $(t'_i, t_i, (\theta, t_{-i}, t'_{-i}))$ form a Blackwell triple - is a natural alternative candidate to rank information structures in many player settings. But, as we will show, this stronger condition which requires independence with respect to t'_{-i} as well, is not necessary. In particular, if t'_i were informative about t'_{-i} beyond the information contained in t_i, and hence were to violate the stronger condition, but still satisfy the weaker condition of individual sufficiency (and hence would not contain information about θ and t_{-i}), then this additional information would not affect the beliefs and higher order beliefs of the player, but rather would encode redundant information, as represented by redundant types. In turn, we do not ask for this stronger condition as it will be shown to be irrelevant for the behavior in Bayes correlated equilibrium. We illustrate these subtle differences shortly in Subsection 3.4 by means of an example.

[7] They also introduce further refinements of the partial order. We discuss these orderings and their uses in Section 5.4.

individual sufficiency in the language of Liu (2011): information structure S is individually sufficient for information structure S' if there exists a combined experiment S^* such that the implied Markov kernel $\phi : T \times \Theta \rightarrow \Delta(T')$ is a correlating device.[8]

3.3 Individual Sufficiency and Higher Order Beliefs

Suppose that we fix an information structure $S = (T, \pi)$ and a full support prior $\psi \in \Delta_{++}(\Theta)$, and we write $\pi_\psi(t_{-i}, \theta | t_i)$ for the implied posterior beliefs,

$$\pi_\psi(t_{-i}, \theta | t_i) \triangleq \frac{\pi(t_i, t_{-i} | \theta)\, \psi(\theta)}{\sum_{t_{-i} \in T_{-i}} \sum_{\theta' \in \Theta} \pi(t_i, t_{-i} | \theta')\, \psi(\theta')}.$$

An information structure is *non-redundant* if no two signals give rise to the same posteriors, or $t_i \neq t_i' \Rightarrow \pi_\psi(\cdot | t_i) \neq \pi_\psi(\cdot | t_i')$. In the one player case, this reduces to the requirement that no two signals gave rise to the same posterior over states; Blackwell (1953) said that an information structure (experiment) was in *standard form* if this property held. Two information structures are said to be *higher-order belief equivalent* if they map to the same type space when redundancies are removed. This is equivalent to the requirement that they give rise to the same probability distribution over higher-order beliefs about Θ. These properties are formally defined and the equivalence is established in the Appendix. The following Lemma establishes that individual sufficiency only depends on non-redundant aspects of the information structure.

Lemma 2 (Unique Non-Redundant Information Structure)

1. *For every information structure S, there is a unique non-redundant information structure $\widehat{S}$ such that S and $\widehat{S}$ are individually sufficient for each other and higher-order belief equivalent*

2. *Any two information structures are individually sufficient for each other if and only if they are higher-order belief equivalent.*

Liu (2011) proved (in his Theorem 1) that two information structures - one of which is non-redundant - are higher-order belief equivalent if and only if there is a unique correlating device - in the sense of equation (26) - that maps the non-redundant information structure into the possibly redundant one. This implies that the non-redundant information structure is individually sufficient for the possibly redundant one. Since it is easy to show that the redundant information structure is sufficient for the non-redundant one, we have they are mutually individually sufficient. This implies part 1 of the Lemma. The second part can be proved by adapting arguments in Lehrer, Rosenberg, and Shmaya (2013); for completeness we report a proof in the Appendix. We also show a tight connection between the individual sufficiency ordering and the more primitive notion of "expansion" as a criterion for more information.

[8] We discuss how Liu (2011) uses this definition in Section 5.1.

Proposition 1

Information structure S is individually sufficient for information structure S' if and only if S is higher-order belief equivalent to an expansion of S'.

3.4 A Class of Binary Information Structures

We illustrate individual sufficiency and the results of this Section by giving two comparisons of information structures under the individual sufficiency ordering. Throughout this section, we shall suppose that there are two possible states, θ_0 and θ_1, and that each state is equally likely.

The first comparison illustrates the irrelevance of information which is "redundant" in the sense of Mertens and Zamir (1985). Examples such as this have been leading examples in the literature, see Dekel, Fudenberg, and Morris (2007), Ely and Peski (2006), Liu (2011). Let S be a "null" information structure where each player has only one possible signal which is always observed. Let S' be given by

$$\begin{array}{|c|c|c|} \hline \pi'(\cdot|\theta_0) & t'_0 & t'_1 \\ \hline t'_0 & \frac{1}{2} & 0 \\ \hline t'_1 & 0 & \frac{1}{2} \\ \hline \end{array} \quad \begin{array}{|c|c|c|} \hline \pi'(\cdot|\theta_1) & t'_0 & t'_1 \\ \hline t'_0 & 0 & \frac{1}{2} \\ \hline t'_1 & \frac{1}{2} & 0 \\ \hline \end{array}. \tag{29}$$

where each player observes one of two signals, t'_0 and t'_1; the above matrices describe the probabilities of different signal profiles, where player 1's signal gives the row, player 2's signal gives the column, the left and right hand matrices correspond to the distribution of signal profiles in states θ_0 and θ_1 respectively and the matrix entries correspond to probabilities.

For each player, each signal arises with probability $\frac{1}{2}$ in each state, so he does not learn about the state. But, collectively, the signals are positively correlated in state θ_0 and negatively correlated in state θ_1. Thus there is common certainty that each player assigns probability $\frac{1}{2}$ to each state. By Lemma 2, information structure S' and the null information structure S are *individually sufficient* for each other. On the other hand, with probability one, every realized profile $t = (t_i, t_j)$, is perfectly informative about the state θ. Thus information structure S' is *sufficient* for null information structure S but S is not sufficient for S'.

This first comparison may suggest that individual sufficiency can be checked by first removing redundancies and then checking "informativeness" player by player. The next comparison is intended to be the simplest possible to illustrate that this is not the case and that individual sufficiency is a more subtle relation. We will now compare two new information structures with the same signal sets and labels that we used previously. Let information structure S be given by:

$$\begin{array}{|l|c|c|}\hline \pi\left(\cdot\,|\theta_0\right) & t_0 & t_1 \\ \hline t_0 & \frac{1}{2} & 0 \\ \hline t_1 & 0 & \frac{1}{2} \\ \hline \end{array} \qquad \begin{array}{|l|c|c|}\hline \pi\left(\cdot\,|\theta_1\right) & t_0 & t_1 \\ \hline t_0 & 0 & 0 \\ \hline t_1 & 0 & 1 \\ \hline \end{array}\,. \tag{30}$$

Under information structure S, if the state is θ_0, with probability $\frac{1}{2}$, it is common knowledge that the state is θ_0 (and both players observe signal t_0); otherwise, both players observe t_1. This is a special case of the class of information structures studied in Section 2.3. Consider now a second binary information structure S' given by:

$$\begin{array}{|l|c|c|}\hline \pi'\left(\cdot\,|\theta_0\right) & t_0' & t_1' \\ \hline t_0' & \frac{1}{2} & \frac{1}{6} \\ \hline t_1' & \frac{1}{6} & \frac{1}{6} \\ \hline \end{array} \quad \begin{array}{|l|c|c|}\hline \pi'\left(\cdot\,|\theta_0\right) & t_0' & t_1' \\ \hline t_0' & \frac{1}{3} & 0 \\ \hline t_1' & 0 & \frac{2}{3} \\ \hline \end{array}\,. \tag{31}$$

In the information structure S', each player observes a signal with "accuracy" $\frac{2}{3}$ in either state: that is, if the state is θ_0, then each player observes t_0' with probability $\frac{2}{3}$; if the state is θ_1, then each player observes t_1' with probability $\frac{2}{3}$. But in state θ_1, the signals are perfectly correlated across players, whereas in state θ_0, the signals are less than perfectly correlated. We will show that S is not *sufficient* for S' and nor is S' sufficient for S. But S is *individually sufficient* for S'. This is true even though there are no redundancies. We will establish individual sufficiency and illustrate and motivate the ordering by noting how information structure S' can be transformed into information structure S by, first, giving players additional information and, second, removing redundancies.

We briefly argue why S is not sufficient for S' and S' is not sufficient for S, using the characterization of the state independent Markov kernel ϕ given in (28) above. If S were sufficient for S', then $\pi\left((t_1,t_1)\,|\theta_1\right)=1$ and $\pi'\left((t_1',t_1')\,|\theta_1\right)=\frac{2}{3}$ would imply that the probability $\phi\left((t_1',t_1')\,|\,(t_1,t_1)\right)$ would have to be $\frac{2}{3}$. But this would imply that the probability of (t_1',t_1') in state θ_0 would have to be at least $\frac{1}{2}\frac{2}{3}=\frac{1}{3}$, which is strictly larger than $\pi'\left((t_1',t_1')\,|\theta_0\right)=1/6$. Thus S cannot be sufficient for S'. Conversely, for S' to be sufficient for S, $\pi\left((t_1,t_1)\,|\theta_1\right)=1$ would imply that $\phi\left((t_1,t_1)\,|(t_0',t_0')\right)=\phi\left((t_1,t_1)\,|(t_1',t_1')\right)=1$, but this would imply that the probability of (t_1,t_1) in state θ_0 would have to be at least $\frac{1}{2}\times 1+\frac{1}{6}\times 1=\frac{2}{3}$, which is strictly larger than $\pi\left((t_1,t_1)\,|\theta_0\right)=1/2$.

We now show, however, that the S is individually sufficient for S'. We first describe the Markov kernel

$\phi : T_1 \times T_2 \times \Theta \rightarrow \Delta(T_1' \times T_2')$ mapping S into S' which establishes individual sufficiency as follows:

$\phi(t' \mid t, \theta)$	$t_0' t_0'$	$t_0' t_1'$	$t_1' t_0'$	$t_1' t_1'$
$t_0 t_0 \theta_0$	1	0	0	0
$\dots$				
$t_1 t_1 \theta_0$	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$\dots$				
$t_1 t_1 \theta_1$	$\frac{1}{3}$	0	0	$\frac{2}{3}$

. (32)

In the above Markov kernel $\phi(t' \mid t, \theta)$, we only specify the conditional probabilities of the above three profiles (and associated rows), $(t_0 t_0 \theta_0), (t_1 t_1 \theta_0), (t_1 t_1 \theta_1)$ that are necessary to verify that the reproducibility condition (25) is satisfied. As the conditional probability $\pi(t \mid \theta)$ of all other profiles is equal to zero, we are free to specify the remaining rows so that the conditional independence requirement (26) of the Markov kernel ϕ is satisfied as well. The only substantial requirement that arises from the independence requirement is that from each player's point of view, the sum (26) when he is type t_1 is independent of θ and the type of the other player, which is immediately verified using the kernel ϕ defined above.

Now let us use the example of S and S' to illustrate Proposition 1. We first use the Markov kernel ϕ to directly construct the relevant combined experiment, S^*, for S and S', and we omit, without loss of generality zero probability signal profiles:

$\pi^*(\cdot \mid \theta_0)$	$t_0' t_0'$	$t_0' t_1'$	$t_1' t_0'$	$t_1' t_1'$
$t_0 t_0$	$\frac{1}{2}$	0	0	0
$t_1 t_1$	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$\pi^*(\cdot \mid \theta_1)$	$t_0' t_0'$	$t_0' t_1'$	$t_1' t_0'$	$t_1' t_1'$
$t_1 t_1$	$\frac{1}{3}$	0	0	$\frac{2}{3}$

.

In this combined information structure S^*, each player has three possible types (or *combined types*) which are a combination of types in S and S' : $t^* \in \{t_0 t_0', t_1 t_0', t_1 t_1'\}$. We re-arrange the representation of the combined information structure S^* by describing for each type t^* in S^*, in the rows, the beliefs over the other player's combined type and state θ:

	$\theta_0 t_0 t_0'$	$\theta_0 t_0 t_1'$	$\theta_0 t_1 t_0'$	$\theta_0 t_1 t_1'$	$\theta_1 t_0 t_0'$	$\theta_1 t_0 t_1'$	$\theta_1 t_1 t_0'$	$\theta_1 t_1 t_1'$
$t_0 t_0'$	1	0	0	0	0	0	0	0
$t_1 t_0'$	0	0	0	$\frac{1}{3}$	0	0	$\frac{2}{3}$	0
$t_1 t_1'$	0	0	$\frac{1}{6}$	$\frac{1}{6}$	0	0	0	$\frac{2}{3}$

Now from the above table, we can extract the first and second-order beliefs of player i, respectively:

1st	θ_0	θ_1
t_0t_0'	1	0
t_1t_0'	$\frac{1}{3}$	$\frac{2}{3}$
t_1t_1'	$\frac{1}{3}$	$\frac{2}{3}$

,

2nd	$\theta_0, 0$	$\theta_0, \frac{2}{3}$	$\theta_1, \frac{2}{3}$
t_0t_0'	1	0	0
t_1t_0'	0	$\frac{1}{3}$	$\frac{2}{3}$
t_1t_1'	0	$\frac{1}{3}$	$\frac{2}{3}$

.

In the table of the second-order beliefs, the entry in each cell corresponds to the probability that the combined type (of a given row) assigns to the state being $\theta \in \{\theta_0, \theta_1\}$ and the other player assigning probability q to the state being θ_1 (of a given (θ_i, q) column). Thus t_1t_0' and t_1t_1' have identical beliefs and higher-order beliefs, and the same beliefs and higher-order beliefs as type t_1 in information structure S. We thus have completed the construction of an expansion of S', namely S^* that is higher-order belief equivalent to S. As established in Proposition 1, this is equivalent to S being individually sufficient for S'.

As we expand the information structure S' and combine it with some other information structure to obtain an expansion S^*, there is a natural sense in which we *augment* the information structure S' by adding extra signals. We finally explicitly describe how the extra information is generated in the combined experiment. Suppose that we first observed S' and then observed extra signals distributed according to $\phi' : T_1' \times T_2' \times \Theta \to \Delta (T_1 \times T_2)$:

$\phi'(\cdot\|\cdot)$	t_0t_0	t_1t_1
$\theta_0t_0't_0'$	1	0
$\theta_0t_0't_1'$	0	1
$\theta_0t_1't_0'$	0	1
$\theta_0t_1't_1'$	0	1
$\theta_1t_0't_0'$	0	1
$\theta_1t_1't_1'$	0	1

So suppose that the player initially observed signals according to the information structure S'. Then each player i observes an additional signal, according the above Markov kernel ϕ', and hence t_0 if the true state was θ_0 and both players had observed t_0'. Otherwise, each player i would observe signal t_1. This expanded information structure, which refines or "splits" the information structure S', is now the same as S^*, and by our earlier argument indeed higher-order belief equivalent to S.

4 Comparing Information Structures

Giving players more information will generate more obedience constraints and thus reduce in size the set of Bayes correlated equilibria. If "giving players more information" is interpreted to mean that we expand their information structures, allowing them to keep their previous signals and observe more, then this claim follows trivially from the definition and characterization of Bayes correlated equilibria in Section 2. In this Section, we strengthen this observation by showing that it is also true if by "giving player more information," we mean that we replace their information structure with one that is individually sufficient for it. And we prove a converse, showing that if an information structure, S, is not individually sufficient for another, S', then there exists a basic game G such that (G, S) has a Bayes correlated equilibrium that generates outcomes that could not arise under a Bayes correlated equilibrium of (G, S').

In order to compare outcomes across information structures, we will be interested in what can be said about actions and states if signals are not observed. We will call a mapping

$$\nu : \Theta \to \Delta(A), \tag{33}$$

a *random choice rule*, and say ν is *induced* by decision rule σ if it is the marginal of σ on A, so that

$$\nu(a|\theta) \triangleq \sum_{t \in T} \sigma(a|t, \theta) \pi(t|\theta), \tag{34}$$

for each $a \in A$ and $\theta \in \Theta$. Random choice rule ν is a Bayes correlated equilibrium random choice rule of (G, S) if it is induced by a Bayes correlated equilibrium decision rule σ of (G, S).

We now define a partial order on information structures that corresponds to shrinking the set of BCE random choice rules in all basic games. Writing $BCE(G, S)$ for the set of BCE random choice rules of (G, S), we say:

Definition 7 (Incentive Constrained)
Information structure S is more incentive constrained than information structure S' if, for all basic games G:

$$BCE(G, S) \subseteq BCE(G, S').$$

Theorem 2
Information structure S is individually sufficient for information structure S' if and only if S is more incentive constrained than S'.

An immediate corollary of this result is:

Corollary 1
If two information structures are individually sufficient for each other, then each is more incentive constrained than the other.

This corollary, but not the preceding theorem, could have been proved by adapting the apparently distinct arguments and results of either Liu (2011) or Lehrer, Rosenberg, and Shmaya (2013).[9]

In establishing the theorem, we will show constructively that if S is individually sufficient for S' and ν is a BCE random choice rule of (G, S), then we can use the BCE decision rule inducing ν and the Markov kernel establishing individual sufficiency to construct a BCE of (G, S') which induces ν; this argument adapts an argument in Lehrer, Rosenberg, and Shmaya (2013). However, the results and arguments in Liu (2011) and Lehrer, Rosenberg, and Shmaya (2013) do not help prove the converse.

To prove the converse, we consider, for any information structure S, a particular basic game G and a particular BCE random choice rule ν of (G, S). If S is more incentive constrained than S', that particular ν must also be a BCE random choice rule of (G, S'). We then show that our choice of G and ν imply that ν is a BCE random choice rule of (G, S'), there must exist a Markov kernel establishing that S is individually sufficient for S'. The basic game used is, roughly, one where players report their beliefs and higher-order beliefs about Θ, and have an incentive to do so truthfully, as in Dekel, Fudenberg, and Morris (2007). Thus there is a BCE random choice rule of (G, S) where they "truthfully" report the beliefs and higher-order beliefs of their original types in S (although of course there are in general also other BCE where they have more information than under S). We then show that for there to be a BCE of (G, S') which induces ν (and thus has players report the distribution of beliefs and higher-order beliefs corresponding to S), there must be a mapping from type profiles in S' to actions in G (which correspond to types in S) that allows us to represent how the random choice rule translates into a decision rule for (G, S'). Obedience constraints ensure that this mapping satisfies the properties that ensure the individually sufficiency of S for S'. Two complications not dealt with in this summary of the argument, but dealt with in the formal proof, are the possibility of redundant types in S and the fact that a finite action approximation of the "higher-order beliefs" game must be used.

Proof. Suppose that S is individually sufficient for S'. Take any basic game G and any BCE σ of (G, S). We will construct $\sigma' : T' \times \Theta \to \Delta(A)$ which is a BCE of (G, S') which gives rise to the same stochastic map as σ.

Write $V_i(a_i, a_i', t_i)$ for the expected utility for player i under distribution σ if he is type t_i, receives recommendation a_i but chooses action a_i', so that

$$V_i\left(a_i, a_i', t_i\right) \triangleq \sum_{a_{-i} \in A_{-i}, t_{-i} \in T_{-i}, \theta \in \Theta} \psi(\theta)\, \pi\left((t_i, t_{-i})\,|\theta\right) \sigma\left((a_i, a_{-i}) \,|\, (t_i, t_{-i}), \theta\right) u_i\left(\left(a_i', a_{-i}\right), \theta\right).$$

Now - by Definition 1 - for each $i = 1, ..., I$, $t_i \in T_i$ and $a_i \in A_i$, we have

$$V_i\left(a_i, a_i, t_i\right) \geq V_i\left(a_i, a_i', t_i\right), \tag{35}$$

[9] We discuss how in Sections 5.1 and 5.4 respectively.

for each $a_i' \in A_i$. Since S is individually sufficient for S', there exists a mapping $\phi : T \times \Theta \rightarrow \Delta(T')$ satisfying (25) and (26). Define $\sigma' : T' \times \Theta \rightarrow \Delta(A)$ by

$$\sigma'\left(a|t',\theta\right) = \frac{\sum_{t\in T} \pi\left(t|\theta\right)\sigma\left(a|t,\theta\right)\phi\left(t'|t,\theta\right)}{\pi'\left(t'|\theta\right)}, \tag{36}$$

for all $(a,t',\theta) \in A \times T' \times \Theta$ whenever $\pi(t'|\theta) > 0$ (and if $\pi(t'|\theta) = 0$, we are free to choose an arbitrary probability distribution $\sigma'(a|t',\theta)$). Hence, the mapping $\sigma(a|t,\theta)$ and $\sigma'(a|t',\theta)$ induce the same random choice function $\nu : \Theta \rightarrow \Delta(A)$. Symmetrically, write $V_i'(a_i, a_i', t_i')$ for the expected utility for player i under decision rule σ' if he is type t_i', receives recommendation a_i but chooses action a_i', so that

$$V_i'\left(a_i,a_i',t_i'\right) \triangleq \sum_{a_{-i}\in A_{-i},t_{-i}'\in T_{-i}',\theta\in\Theta} \psi\left(\theta\right)\pi'\left(\left(t_i',t_{-i}'\right)|\theta\right)\sigma'\left(\left(a_i,a_{-i}\right)|\left(t_i',t_{-i}'\right),\theta\right)u_i\left(\left(a_i',a_{-i}\right),\theta\right).$$

Now σ' satisfies the obedience condition (Definition 1) to be a correlated equilibrium of (G, S') if for each $i = 1, ..., I$, $t_i' \in T_i'$ and $a_i \in A_i$,

$$V_i'\left(a_i,a_i,t_i'\right) \geq V_i'\left(a_i,a_i',t_i'\right),$$

for all $a_i' \in A_i$. But

$$\begin{aligned}
& V_i'\left(a_i,a_i',t_i'\right) \\
= & \sum_{a_{-i}\in A_{-i},t_{-i}'\in T_{-i}',\theta\in\Theta} \psi\left(\theta\right)\pi'\left(\left(t_i',t_{-i}'\right)|\theta\right)\sigma'\left(\left(a_i,a_{-i}\right)|\left(t_i',t_{-i}'\right),\theta\right)u_i\left(\left(a_i',a_{-i}\right),\theta\right) \\
= & \sum_{a_{-i}\in A_{-i},t_{-i}'\in T_{-i}',\theta\in\Theta,t\in T} \psi\left(\theta\right)\pi\left(t|\theta\right)\sigma\left(\left(a_i,a_{-i}\right)|t,\theta\right)\phi\left(t'|t,\theta\right)u_i\left(\left(a_i',a_{-i}\right),\theta\right) \\
& \text{by the definition of } \sigma' \text{, see (36),} \\
= & \sum_{a_{-i}\in A_{-i},\theta\in\Theta,t\in T} \psi\left(\theta\right)\pi\left(t|\theta\right)\sigma\left(\left(a_i,a_{-i}\right)|t,\theta\right)u_i\left(\left(a_i',a_{-i}\right),\theta\right)\sum_{t_{-i}'\in T_{-i}'}\phi_i\left(\left(t_i',t_{-i}'\right)|t,\theta\right) \\
= & \sum_{a_{-i}\in A_{-i},\theta\in\Theta,t\in T} \psi\left(\theta\right)\pi\left(t|\theta\right)\sigma\left(\left(a_i,a_{-i}\right)|t,\theta\right)u_i\left(\left(a_i',a_{-i}\right),\theta\right)\phi_i\left(t_i'|t_i\right), \text{ by (26) and (27),} \\
= & \sum_{t_i\in T_i}\phi_i\left(t_i'|t_i\right)\left[\sum_{a_{-i}\in A_{-i},\theta\in\Theta,t_{-i}\in T_{-i}} \psi\left(\theta\right)\pi\left(\left(t_i,t_{-i}\right)|\theta\right)\sigma\left(\left(a_i,a_{-i}\right)|\left(t_i,t_{-i}\right),\theta\right)u_i\left(\left(a_i',a_{-i}\right),\theta\right)\right] \\
= & \sum_{t_i\in T_i}\phi_i\left(t_i'|t_i\right)V_i\left(a_i,a_i',t_i\right).
\end{aligned} \tag{37}$$

Now for each $i = 1, ..., I$, $t_i' \in T_i'$ and $a_i \in A_i$,

$$\begin{aligned}
V_i'\left(a_i,a_i,t_i'\right) &= \sum_{t_i\in T_i}\phi_i\left(t_i'|t_i\right)V_i\left(a_i,a_i,t_i\right), \text{ by (37)} \\
&\geq \sum_{t_i\in T_i}\phi_i\left(t_i'|t_i\right)V_i\left(a_i,a_i',t_i\right), \text{ by (35) for each } t_i\in T_i \\
&= V_i'\left(a_i,a_i',t_i'\right), \text{ by (37)}
\end{aligned}$$

for each $a'_i \in A_i$. Thus σ' is a BCE of (G, S'). By construction σ' and σ induce the random choice rule $\nu : \Theta \to \Delta(A)$. Since this argument started with an arbitrary BCE random choice rule ν of (G, S) and an arbitrary G, we have $BCE(G, S) \subseteq BCE(G, S')$ for all basic games G.

Now we will show that if S is non-redundant and S is more incentive constrained than S', then S is individually sufficient for S'. We deal with the case that S is redundant at the end of the proof.

Write $\lambda_i(t_i) \in \Delta(T_{-i} \times \Theta)$ for type t_i's beliefs:

$$\lambda_i(t_{-i}, \theta | t_i) \triangleq \frac{\psi(\theta)\, \pi((t_i, t_{-i}) | \theta)}{\sum\limits_{\widetilde{t}_{-i}, \widetilde{\theta}} \psi\left(\widetilde{\theta}\right) \pi\left((t_i, \widetilde{t}_{-i}) | \widetilde{\theta}\right)}.$$

Write Λ_i for the range of λ_i. Thus $\lambda_i : T_i \to \Lambda_i$. By non-redundancy of S, there is well-defined inverse map $\lambda_i^{-1} : \Lambda_i \to T_i$, so that $\lambda_i(t_i) = \xi_i$ if and only if $\lambda_i^{-1}(\xi_i) = t_i$.

For any combined information structure $\pi : \Theta \to \Delta(T \times T')$, write for the induced beliefs of player i given a type pair (t_i, t'_i):

$$\lambda_i^{\pi}\left(t_{-i}, \theta | t_i, t'_i\right) \triangleq \frac{\sum\limits_{t'_{-i}} \psi(\theta)\, \pi\left((t_i, t_{-i}), (t'_i, t'_{-i}) | \theta\right)}{\sum\limits_{\widetilde{t}_{-i}, \widetilde{\theta}} \sum\limits_{t'_{-i}} \psi\left(\widetilde{\theta}\right) \pi\left((t_i, \widetilde{t}_{-i}), (t'_i, \widetilde{t}'_{-i}) | \widetilde{\theta}\right)}.$$

Write Z for the set of combined information structures $\pi : \Theta \to \Delta(T \times T')$ and note that Z is a compact set.

The proof now proceeds by contrapositive. Thus suppose that S is not individual sufficient for S'. Then, by the property (24) of individual sufficiency, for every $\pi \in Z$, there exists, by non-redundancy, i, t_i and t'_i such that:

$$\lambda_i^{\pi}\left(\cdot | t_i, t'_i\right) \neq \lambda_i(\cdot | t_i).$$

Now define

$$\varepsilon \triangleq \frac{1}{2} \inf_{\zeta \in Z} \max_{i, t_i, t'_i} \left\| \lambda_i^{\pi}\left(\cdot | t_i, t'_i\right) - \lambda_i(\cdot | t_i) \right\|,$$

where $\|\cdot\|$ represents the Euclidean distance between vectors in $\mathbb{R}^{T_{-i} \times \Theta}$. The compactness of the set Z and the continuity of the finite collection of mappings $\lambda_i^{\pi}(\cdot | t_i, t'_i)$ with respect to π imply that $\varepsilon > 0$.

Now we will construct a basic game $G = \left((A_i, u_i)_{i=1}^{I}, \psi\right)$ and a random choice rule $\nu^* : \Theta \to \Delta(A)$ such that $\nu^* \in BCE(G, S)$ but $\nu^* \notin BCE(G, S')$. This will complete the proof, via contrapositive, that S being more incentive constrained than S' implies that S is individually sufficient for S'.

Recall that Λ_i (the range of λ_i) is a finite subset of $\Delta(T_{-i} \times \Theta)$. Let Ξ_i be any ε-grid of $\Delta(T_{-i} \times \Theta)$, i.e., a finite subset of $\Delta(T_{-i} \times \Theta)$ satisfying the property that, for all $\xi_i \in \Delta(T_{-i} \times \Theta)$, there exists $\xi'_i \in \Xi_i$ with $\left\|\xi_i - \xi'_i\right\| \leq \varepsilon$.

Now for every player i, let the set of actions be $A_i \triangleq \Lambda_i \cup \Xi_i$. An element $a_i \in A_i$ is therefore a vector of beliefs over the types t_{-i} and state θ, and we denote by $a_i(t_{-i}, \theta)$ the entry of the vector a_i that specifies the probability assigned to (t_{-i}, θ). We let the payoff function of each player i be:

$$u_i(a, \theta) \triangleq \begin{cases} 2a_i\left(\left(\lambda_j^{-1}(a_j)\right)_{j \neq i}, \theta\right) - \sum\limits_{\widetilde{t}_{-i} \in T_{-i}, \widetilde{\theta} \in \Theta} \left(a_i\left(\widetilde{t}_{-i}, \widetilde{\theta}\right)\right)^2 & \text{if} \quad a_j \in \Lambda_j, \forall j \neq i; \\ 0 & \text{if} \quad \text{otherwise.} \end{cases}$$

The action sets A_i and the utility functions u_i induce a basic game of belief elicitation, where the payoff function is a quadratic scoring rule.

Now suppose player i assigns probability 1 to his opponents choosing $a_{-i} \in \Lambda_{-i}$ and, in particular, for some $\xi_i \in \Delta(T_{-i} \times \Theta)$, assigns probability $\xi_i\left(\left(\lambda_j^{-1}(a_j)\right)_{j \neq i}, \theta\right)$ to his opponents choosing a_{-i} and the state being θ. The expected payoff to player i with this belief over $A_{-i} \times \Theta$ parameterized by $\xi_i \in \Delta(T_{-i} \times \Theta)$ is

$$\begin{aligned} & \sum_{t_{-i} \in T_{-i}, \theta \in \Theta} \xi_i(t_{-i}, \theta)\left(2a_i(t_{-i}, \theta) - \sum_{\widetilde{t}_{-i} \in T_{-i}, \widetilde{\theta} \in \Theta} \left(a_i\left(\widetilde{t}_{-i}, \widetilde{\theta}\right)\right)^2\right) \\ = \; & 2 \sum_{t_{-i} \in T_{-i}, \theta \in \Theta} \xi_i(t_{-i}, \theta)\, a_i(t_{-i}, \theta) - \sum_{\widetilde{t}_{-i} \in T_{-i}, \widetilde{\theta} \in \Theta} \left(a_i\left(\widetilde{t}_{-i}, \widetilde{\theta}\right)\right)^2 \\ = \; & 2 \sum_{t_{-i} \in T_{-i}, \theta \in \Theta} \xi_i(t_{-i}, \theta)\, a_i(t_{-i}, \theta) - \sum_{t_{-i} \in T_{-i}, \theta \in \Theta} \left(a_i(t_{-i}, \theta)\right)^2 \\ = \; & \left(\|\xi_i\|^2 - \|a_i - \xi_i\|^2\right). \end{aligned}$$

Thus player i with belief ξ_i has a best response to set a_i equal to one of the points in $A_i \subseteq \Delta(T_{-i} \times \Theta)$ with the shortest Euclidean distance to ξ_i.

Now the game (G, S) has - by construction - a "truth-telling" BCE where each type t_i always chooses action $\lambda_i(t_i)$. This gives rise to a random choice rule ν^* where

$$\nu^*(a|\theta) = \begin{cases} \pi(a|\theta), & \text{if} \quad a \in \{\lambda_i(t_i)\}_{i=1}^I \text{ for some } t \in T, \\ 0, & \text{if} \quad \text{otherwise.} \end{cases}$$

So ν^* is a BCE random choice rule of (G, S). For ν^* to be BCE of (G, S'), there must exist a combined information structure π and and associated decision rule $\sigma' : \Theta \times T' \to \Delta(T)$, defined by

$$\sigma'(t|t', \theta) = \frac{\pi(t, t'|\theta)}{\sum_{\widetilde{t}} \pi\left(\widetilde{t}, t'|\theta\right)}$$

is a BCE. But for any $\pi \in Z$, we showed by the failure of individual sufficiency that there exist i, t_i and t_i' with

$$\left\|\lambda_i^\pi(\cdot|t_i, t_i') - \lambda_i(\cdot|t_i)\right\| \geq 2\varepsilon.$$

But this implies a violation of obedience, since by construction of G, there exists an action $a_i \in A_i$ which is within ε of $\lambda_i^{\pi}(\cdot|t_i, t_i')$ and thus closer to $\lambda_i^{\pi}(\cdot|t_i, t_i')$ than $\lambda_i(\cdot|t_i)$, and so a player with type t_i' receiving action recommendation t_i would strictly prefer to deviate to a_i.

If S was not non-redundant, and S is more incentive constrained than S', let $\widehat{S}$ be the unique non-redundant information structure which is higher-order belief equivalent to S, as shown in Lemma 2. By the previous argument, $\widehat{S}$ is individually sufficient for S'. By Lemma 2, S is individually sufficient for $\widehat{S}$. Individual sufficiency is transitive (this is proved as Lemma 4 in the Appendix). So S is individually sufficient for S'. ■

5 Discussion

5.1 Incomplete Information Correlated Equilibrium

Aumann (1974), (1987) introduced a definition of correlated equilibrium for complete information games. A classic paper of Forges (1993) is titled "five legitimate definitions of correlated equilibrium in games with incomplete information." Her title and paper make the point - that we agree with - that there are many natural ways of extending the complete information definition to incomplete information settings and which definition makes sense depends on the purpose for which it is to be used. In this Section, we present a way of seeing how our definition of "Bayes correlated equilibrium" relates to the more closely related definitions of incomplete information correlated equilibrium, highlighting which definition is relevant for which purpose.

For a fixed basic game G and information structure S, a Bayes correlated equilibrium is a decision rule mapping signal profiles and payoff states to probability distributions over action profiles that satisfies obedience (2), requiring that a player who knows his signal and the action he is supposed to play has no incentive to deviate. We treat states symmetrically with signals and impose no restrictions on what is feasible. The role of the information structure, then, is only to impose incentive constraints on behavior by conditioning best responses, the obedience condition, on the signal realizations. Our motive to study the solution concept of Bayes correlated equilibrium is Theorem 1: the solution concept captures rational behavior given that players have access to the signals in the information structure, but may have additional information.

A decision rule σ is *belief invariant* if, for each player i, the probability distribution over player i's actions that it induces depends only on player i's type, and is independent of other players' types and the state. Writing $\sigma_i : T \times \Theta \rightarrow \Delta(A_i)$ for the probability distribution over player i's actions induced by σ,

$$\sigma_i\left(a_i | \left(t_i, t_{-i}\right), \theta\right) \triangleq \sum_{a_{-i}} \sigma\left(\left(a_i, a_{-i}\right) | \left(t_i, t_{-i}\right), \theta\right),$$

we have:

Definition 8 (Belief Invariant Decision Rule)
Decision rule σ is belief invariant for (G, S) if, for each player i, $\sigma_i\left(a_i | \left(t_i, t_{-i}\right), \theta\right)$ is independent of t_{-i} and θ.

An equivalent way of stating this property, appealing to the language of Section 3.1, is that if we look at the joint distribution over (a, t, θ) generated by any strictly positive prior on Θ, information structure π and decision rule σ, we have that for each player i, $(a_i, t_i, (t_{-i}, \theta))$ form a Blackwell triple. By Lemma 1, we then have that player i's beliefs about (t_{-i}, θ) conditional on t_i do not depend on a_i. In the language of mediation it says that the mediator's recommendation does not give a player any additional information about the state and other players' types. The condition of belief invariance was introduced in this form and so named by Forges (2006). If a decision rule σ is belief invariant for (G, S), then players have no less but also no more information under σ and S than under information structure S. If we impose belief invariance as well as obedience on a decision rule, we get a solution concept that was introduced in Liu (2011).

Definition 9 (Belief Invariant BCE)
Decision rule σ is a belief invariant Bayes correlated equilibrium of (G, S) if it is obedient and belief invariant for (G, S).

It captures the implications of common knowledge of rationality and that players know exactly the information contained in S, if the common prior assumption is maintained. As explained in Liu (2011), this solution concept can be seen as the common prior analogue of the solution concept of interim correlated equilibrium discussed by Dekel, Fudenberg, and Morris (2007). The set of Bayes correlated equilibria of (G, S) is the union of all belief invariant BCE of (G, S') for all information structures S' which are individually sufficient for S.

Liu (2011) showed that if two information structures are higher-order belief equivalent, then they have the same set of belief invariant Bayes correlated equilibria. This in turn implies that they have the same set of Bayes correlated equilibria, which was Corollary 1 of Theorem 2.

Much of the literature on incomplete information correlated equilibrium started from the premise that an incomplete information definition of correlated equilibrium should capture what could happen if players had access to a correlation device / mediator under the maintained assumption that the correlation device/mediator did not have access to information that was not available to the players. We can describe the assumption formally as:

Definition 10 (Join Feasible)
Decision rule σ is join feasible for (G, S) if $\sigma(a|t, \theta)$ is independent of θ.

Thus (a, t, θ) are a Blackwell triple; an equivalent statement is that the action recommendations to all players reveal no additional information about the state beyond the signal profile. If join feasibility but not belief invariance is assumed, we get another solution concept:

Definition 11 (Bayesian Solution)
Decision rule σ is a Bayesian solution of (G, S) if it is obedient and join feasible.

Join feasibility was implicitly assumed in Forges (1993) and other works, because it was assumed that type or signal profiles exhausted all uncertainty;[10] Lehrer, Rosenberg, and Shmaya (2010), (2013) explicitly impose this assumption. The Bayesian solution was named by Forges (1993) and it is the weakest version of incomplete information correlated equilibrium she studies. Imposing both join feasibility and belief invariance, we get a solution concept that has played an important role in the literature.

Definition 12 (Belief Invariant Bayesian Solution)
Decision rule σ is a belief invariant Bayesian solution of (G, S) if it is obedient, belief invariant and join feasible.

Forges (2006) introduced this name. The other incomplete information solution concepts for an incomplete information game in Forges (1993), (2006) - communication equilibrium, agent normal form correlated equilibrium and strategic form correlated equilibrium - are all strictly stronger than the belief invariant Bayesian solution, by imposing additional "truth-telling" constraints (for communication equilibrium), feasible correlation structure constraints (for agent normal form correlated equilibrium) and a combination of the two (for strategic form correlated equilibrium). Forges (1993) also discusses a "universal Bayesian solution" which corresponds to Bayes correlated equilibrium in the case where S is degenerate, i.e., there is no prior information structure (beyond the common prior over payoff states).

We conclude by mentioning two further ways of connecting Bayes correlated equilibrium with existing solution concepts. Information structure S satisfies *distributed certainty* if the players' pooled signals

[10] The issue is discussed in Section 4.5 of Forges (1993), where she notes how analyzing a "reduced form" game is not innocuous.

perfectly reveal the state, so that there exists a mapping $g: T \to \Theta$ such that $\pi(t|\theta) > 0 \Rightarrow g(t) = \theta$. If an information structure satisfies distributed certainty, then join feasibility is automatically satisfied and thus any Bayes correlated equilibrium is a Bayesian solution and any belief invariant Bayesian correlated equilibrium is a belief invariant Bayesian solution. Distributed certainty is satisfied in important economic settings, an important example being the case of private values.[11]

Another way to relate solution concepts is to think about adding a "dummy player" who knows the state but is otherwise irrelevant.[12] Suppose we started with a basic game G and information structure S. Suppose we write G_0 for the basic game where we add a dummy player who has no action choice and whose payoff is therefore irrelevant; and we write S_0 for the information structure where we add a dummy player who observes the state θ perfectly. There is a natural isomorphism between Bayes correlated equilibria of (G, S) and Bayesian solutions (ignoring the behavior of the dummy player) of (G_0, S_0). While we do not use this dummy player formulation because it is cumbersome, we will use this translation in Section 5.4 below to explain the relation of our results to the literature.

5.2 The One Player Special Case

Our results apply to the case of one player. In this case, they have natural and important interpretations. As we will discuss, versions of our results have appeared in the prior literature, although expressed and motivated in somewhat different ways. In this Section, we briefly discuss informally the one player analogues of our results and how they relate to three recent works.

In the one player case, a basic game reduces to a *decision problem*, mapping actions and states to a payoff of the *decision maker*. An information structure corresponds to an *experiment* in the sense of Blackwell (1951), (1953). A decision rule in now a mapping from state and signals to probability distributions over actions. A decision rule is a Bayes correlated equilibrium if it is obedient. To interpret obedience, consider a decision maker who observed a signal under the experiment and received an action recommendation chosen according to the decision rule. The decision rule is obedient if he would have an incentive to follow the recommendation. A decision rule is a Bayes Nash equilibrium if it could be generated by a decision maker who observed a signal and then chose an optimal action without receiving an additional action recommendation. Theorem 1 states that the set of Bayes correlated equilibria for a fixed decision

[11] Thus there are private values in our related work on price discrimination, Bergemann, Brooks, and Morris (2013). We assumed "distributed certainty" - but not private values - in our earlier work on robust mechanism design, Bergemann and Morris (2012), and the epistemic foundations we reviewed in Bergemann and Morris (2007) were also based on that assumption.

[12] We are grateful to Atsushi Kajii for suggesting that we pursue this dummy player interpretation of Bayes correlated equilibrium. The taxonomy of incomplete information correlated equilibrium concepts in Milchtaich (2012) discusses the possibility of treating nature as a player.

problem and experiment equals the set of decision rules from a decision maker choosing an optimal action with access to that experiment and possibly more information (an expanded information structure). Thus Bayes correlated equilibria capture all possible optimal behavior if the decision maker had access to the fixed experiment and perhaps some additional information.

To compare the present result to the existing results, consider the case where the original information structure is degenerate (there is only one signal which represents the prior over the states of the world). In this case, the set of Bayes correlated equilibria correspond to joint distributions of actions and states that could arise under rational choice by a decision maker with any information structure. We follow Chwe (2006) in studying the implications of obedience constraints without reference to how much information decision makers might have. Chwe (2006) identified implications for the covariance of actions and states that are implied by obedience conditions in general, and in particular decision problems (as well as games). Kamenica and Gentzkow (2011) consider a problem of "Bayesian persuasion". Suppose a "sender" could pick the experiment that the decision maker could observe. Kamenica and Gentzkow (2011) characterize the set of joint distributions over states and actions that the sender could induce through picking an experiment and having the decision maker choose optimally. This set is exactly what we label Bayes correlated equilibria. They can then analyze which (in our language) Bayes correlated equilibrium the sender would prefer to induce in a variety of applications. Kamenica and Gentzkow (2011) note how the linear algebraic characterization is related to the classic "concavification" or "splitting" arguments in the literature on repeated games with one-sided private information (Aumann and Maschler (1995)).

Caplin and Martin (2013) introduce a theoretical and empirical framework for analyzing imperfect perception. Suppose that we observe a joint distribution over actions and states, but are uncertain what information the players have observed (in their laboratory experiments, they do not know how much attention subjects have paid to information about the state provided to them). When is it that joint distribution are consistent with rational behavior for some utility function? They argue that there is a utility function that rationalizes behavior only if the distribution satisfies "No Improving Action Switches" (NIAS) inequalities, which correspond to the obedience constraints for the decision problem (with a degenerate experiment). While their primary motivation is to identify if there is a utility function consistent with a joint distribution of actions and states (e.g., one observed in experiments), they also illustrate the set of what we call "Bayes correlated equilibria" by characterizing the set of joint distributions that could arise for a fixed utility function ("robust predictions").

In the one player case, we argued above that our individual sufficiency ordering on information structures reduces to the classical sufficiency ordering on experiments of Blackwell (1951), (1953). In the one player case, an experiment, S, is "more incentive constrained" than another, S', if, for every decision problem, the set of random choice rules (joint distributions over actions and states) that are induced by Bayes correlated equilibria with experiment S are a subset of those induced with experiment S'. Thus our Theorem 2 shows that more information in the sense of Blackwell (1951), (1953) leads to more binding incentive constraints and thus *reduces* the set of possible outcomes. The idea that more information reduces the set of incentive compatible outcomes appears in the mechanism design literature in a variety of contexts. Theorem 2 applied to the one player case seems fairly obvious; in particular, we report in footnote 13 in the next subsection an elegant proof in the case of one player. However, we do not know a basic reference for the observation, beyond the version reported above.

We emphasize that Theorem 2 applied to the one player case is *not* Blackwell's Theorem. We discuss Blackwell's Theorem and the relation to our work in the next subsection.

5.3 Feasibility and Blackwell's Theorem

Our Theorem 2 relates together a *statistical* ordering (individual sufficiency) and a *incentive ordering* (more incentive constrained). More information leads to a *smaller* set of Bayes correlated equilibria because it adds incentive constraints. Information is unambiguously "bad" in the sense of reducing the set of possible outcomes. Lack of information is never a constraint on what is feasible for players because the solution concept of Bayes correlated equilibrium imposes no feasibility constraints on players' behavior.

On the other hand, we would argue that Blackwell's Theorem relates a statistical ordering to a *feasibility* ordering. In the one player case, more information is "good" in the sense of leading to more feasible joint distributions of actions and states and thus (in the one person case) to higher ex-ante utility. Incentive constraints do not bind, because there is a single decision maker. In this section, we will report a result which relates our statistical ordering to a feasibility ordering in the many player case. The approach and result is a straightforward variation on the work of Lehrer, Rosenberg, and Shmaya (2010), so we report the result without formal proof.

Say that basic game G has common interests if $u_1 = u_2 = ... = u_I = u^*$. Fix a common interest basic game G and an information structure S. Recall from Definition 8 that a decision rule is belief invariant for (G, S) if, for each player, the distribution of his action depends only on his type and is independent of others' types and the state. Let $v(G, S)$ be the highest possible ex-ante utility that is attained by any player under belief invariant decision rule:

$$v(G,S) \triangleq \max_{\{\sigma:T\times\Theta\rightarrow\Delta(A)|\sigma \text{ is belief invariant for } (G,S)\}} \sum_{a,t,\theta} \psi(\theta)\,\pi(t|\theta)\,\sigma(a|t,\theta)\,u(a,\theta). \tag{38}$$

Thus we are asking what is the highest (common) payoff that players could obtain if they were able to correlate their behavior but could only do so using correlation devices in the sense of Liu (2011) under which a player's action recommendation gives him no additional information about others' types and the state. Here the information structure is constraining (through belief invariance) the set of joint distributions over actions and states that can arise. Say that an information structure S is more valuable than S' if, in every common interest basic game G, there is a belief invariant decision rule for (G, S) that gives a higher common ex-ante payoff than any belief invariant decision rule for (G, S').

Definition 13 *Information structure S is more valuable than information structure S' if, for every common interest basic game G, $v(G, S) \geq v(G, S')$.*

Now we have:

Theorem 3 *Information structure S is individually sufficient for information structure S' if and only if S is more valuable than S'.*

Notice that obedience constraints do not arise in any of the properties used to state this theorem. In that sense, the Theorem relates a statistical ordering to a feasibility ordering and does not make reference to incentive compatibility constraints. But also notice that, since the game has common interests, the belief invariant decision rule that is the argmax of expression (38) will automatically satisfy obedience. Recall from Definition 9 that a decision rule is a belief invariant Bayes correlated equilibrium of (G, S) if it satisfies belief invariance and obedience. Thus $v(G, S)$ is also the ex-ante highest common payoff that can be obtained in a belief invariant Bayes correlated equilibrium.

In the special case of one player, Theorem 3 clearly reduces to the classic statement of Blackwell's theorem favored by economists; we discuss how, in the many player case, it follows from argument of Lehrer, Rosenberg, and Shmaya (2010) in the next sub-Section.

We can sketch a direct proof of the harder direction of Theorem 3. Manipulations of definitions shows that S is individually sufficient for information structure S' if and only if the set of random choice rules induced by belief invariant decision rules for (G, S) is larger than that set for (G, S'). In other words, for any action sets for the players, S supports a larger set of feasible random choice rules than S'. Since these sets are compact and convex, the separating hyperplane theorem implies we can choose a common utility function such that ex-ante expected utility is higher under S than under S'.

Theorems 2 and 3 together imply an equivalence between S is more valuable than S' and S is more incentive constrained than S'. We do not know a direct proof of the equivalence of these two partial orders in the many player case.[13]

[13] In the one player case only, we do know an elegant argument that if S is more incentive constrained than S', then S is more valuable than S'. Suppose S is not more valuable than S'. Then there exists a common interest basic game G with

5.4 Other Relations on Information Structures and Their Uses

There is a large literature that has studied the impact of changing information structures on equilibrium outcomes in games. We will review some of that literature and its relation to our work here. The following perspective on this literature will organize our discussion. Intuitively, more information sometimes enlarges the set of things that can happen in a game, by making more outcomes feasible. Sometimes, more information reduces the set of things that can happen in a game by adding incentive constraints. In general, there may be a tension between these effects. In this paper, we propose the "individual sufficiency" ordering as a natural partial order on information structures capturing a natural notion of more information. In our main result, by focussing on Bayes correlated equilibria, we abstract from the equilibrium enlarging effect of information, due to the feasibility considerations and focus on the equilibrium reducing effect, due to the incentive considerations.

We first describe the results from Lehrer, Rosenberg, and Shmaya (2010), (2013) that are closest to ours, and identify the exact connections. We described in Section 3.2, an ordering on information structures introduced by Lehrer, Rosenberg, and Shmaya (2010), (2013) - non-communicating garbling - which is stronger than both individual sufficiency and sufficiency.[14] The results that are closest to ours are those for the belief invariant Bayesian solution, which we defined in Section 5.1 and is the weakest solution concept that they focus on.

Lehrer, Rosenberg, and Shmaya (2010) analyzes common interest basic games. They consider belief invariant Bayesian solution of (G, S) which gives players the highest common ex-ante utility. Theorem 4.5 shows that the maximum utility is higher in (G, S) than in (G, S') for all common interest games G if and only if S' is a non-communicating garbling of S. Our Theorem 3 is an easy corollary of Theorem 4.5 of Lehrer, Rosenberg, and Shmaya (2010): if we relax join feasibility in the solution concept, we get belief invariant Bayes correlated equilibrium and the corresponding relaxation of the information structure ordering gives individual sufficiency.

Lehrer, Rosenberg, and Shmaya (2013) consider general basic games. Part (c) of Theorem 2.8 shows that two information structures are non-communicating garblings of each other if and only, in every basic game, they have the same set of belief invariant Bayesian solution random choice rules. It is a straightforward analogue of their results - which could be proved by treating nature as a dummy player and applying the connections described in Section 5.1 - that two information structures are individually sufficient for each

$v(G, S') > v(G, S)$. This implies that there exists a Bayes correlated equilibrium of (G, S) that gives the player ex ante utility $v(G, S)$. But in the one player case, every Bayes correlated equilibrium must give the one player ex ante utility greater than $v(G, S')$ (this is not true with many players). Thus there is a BCE random choice rule of (G, S) which is not a BCE random choice rule of (G, S'). Thus S is not more incentive constrained than S'.

[14] One can also show by example that it is strictly stronger than requiring both individual sufficiency and sufficiency.

other if and only if, in every basic game, they have the same set of belief invariant BCE random choice rules. Since the set of BCE random choice rules for (G, S) equals the set of belief invariant Bayes correlated equilibria random choice rules for (G, S') for all information structures S' for which S is individually sufficient, this implies Corollary 1 above. The analogous claim to our main Theorem 2 would be that if S' is a non-communicating garbling of S, then the set of belief invariant Bayesian solution random choice rules of (G, S) is a subset of those for (G, S') for all basic games G. However, this claim is not true. The reason - previewed above - is that the Bayes correlated equilibrium solution concept imposes only incentive constraints and no feasibility conditions, so information can only reduce the set of equilibria. However, the belief invariant Bayesian solution imposes join feasibility and belief invariance, conditions that become less demanding the more information there is. Thus the classical conflict between incentive and feasibility requirements becomes relevant.[15]

While Lehrer, Rosenberg, and Shmaya (2010), (2013) are the closest works to ours, there is a large literature on the value of information in games, and we now discuss that work and its relation. Hirshleifer (1971) noted why information might be damaging in a many player context because it removed options to insure ex ante. Our result on the incentive constrained ordering can be seen as a formalization of the idea behind the observation of Hirshleifer (1971): we give a general statement of how information creates more incentive constraints and thus reduces the set of incentive compatible outcomes.

We have highlighted the dual roles of information which are common to the one player and many player cases: increasing feasible outcomes and reducing incentive compatible ones. Neyman (1991) emphasized that within a fixed overall information structure, under Bayes Nash equilibrium, a player was better off with more information. Thus if some of player i's signals are more informative than others, then player i is better off in equilibrium conditional on receiving the more informative signals. In this case, more information makes more outcomes feasible and, because other players do not know if he is more informed or not, does not tighten their incentive constraints.

Gossner and Mertens (2001) consider Bayes Nash equilibrium and zero sum games. They show that two information structures imply the same value if and only if they are higher-order belief equivalent. And they show that a sufficient condition for a player to have a higher value is that he has more information or his opponent has less information. Peski (2008) shows that the sufficient conditions are also necessary. That is, for a fixed information structure, the set of information structures where a player will have a higher value in all zero sum games consists of those where he is more informed and his opponent is less

[15] Lehrer, Rosenberg, and Shmaya (2010), (2013) also propose even stronger orderings on information structures ("independent garbling" and "coordinated garbling") and show that their results on common interest basic games in Lehrer, Rosenberg, and Shmaya (2010) and general basic games in Lehrer, Rosenberg, and Shmaya (2013) extend in a natural way to more refined solution concepts (Bayes Nash equilibrium and agent normal form correlated equilibrium, respectively).

informed. The proof of this result involves an appeal to the separating hyperplane theorem to show that if the condition on information structures is not satisfied, it is possible to construct a zero sum basic game where the player has a lower value. In our main result, we must similarly construct a basic game showing a failure of the incentive constrained ordering if the statistical relation fails. The argument, and in particular the construction of the critical games, seem quite different, however.

Gossner (2000) considers Bayes Nash equilibrium and general games and characterizes when one information structure supports more BNE outcomes than another. While the bulk of his work focusses on complete information games, in Section 6 and Theorem 19 he considers incomplete information games. His definition that one information structure S' is a *faithful interpretation* of another S translates in our language to the requirement that they are higher-order belief equivalent and there is a profile of Markov kernels which are independently mapping each player signals S_i into signals in S'_i. He shows that S supports more BNE outcomes than S' in all games if and only if S' is a faithful interpretation of S. Thus this ordering ranks an information structure higher if it gives more "correlation possibilities", but holds fixed beliefs and higher-order beliefs. By contrast, individual sufficiency abstracts from "correlation possibilities" and depends non-trivially on beliefs and higher-order beliefs about payoffs.

6 Appendix

6.1 Binary Game with Strategic Substitutes

We record the computation of the set of BCE for a specific subset of binary games as introduced in Section 2.3 and illustrated in Figure 1 - 4. The parameter are set to be $z = 2$, $y_B = -1/6$ and $y_G = 0$, thus a game of strategic substitutes, and the prior probability ψ of θ_G is $\psi = 1/3$. The resulting payoff matrices are:

θ_B	I	N
I	$1-\frac{1}{6}, 1-\frac{1}{6}$	$-1, 2$
N	$2, -1$	$0, 0$

θ_G	I	N
I	$3, 3$	$1, 2$
N	$2, 1$	$0, 0$

.

This game is best response equivalent, and hence BCE equivalent to the following game, which we shall use in the subsequent computations:

θ_B	I	N
I	$-\frac{7}{6}, -\frac{7}{6}$	$-1, 0$
N	$0, -1$	$0, 0$

θ_G	I	N
I	$1, 1$	$1, 0$
N	$0, 1$	$0, 0$

. (39)

The asymmetry in the externality y across states, namely $y_B = -1/6$ and $y_G = 0$, facilitates the computation as $y_G = 0$ renders some case distinctions unnecessary but does not affect the qualitative features of the BCE set. We shall compute the set of BCE for three different information structures, indexed by q, with $q = 1/4, 11/20, 4/5$.

Obedience Constraints The set of BCE is characterized completely by the obedience constraints (12) and (13), given the parametrized decision choice function (10). Given the payoff matrices of (39), these reduce to investment obedience constraint:

$$\frac{1}{3}\alpha_G + \frac{2}{3}(1-q)\left(-\alpha_B - \frac{1}{6}\gamma_B\right) \geq 0; \tag{40}$$

and the no investment obedience constraint:

$$\frac{1}{3}(1-\alpha_G) + \frac{2}{3}(1-q)\left(-(1-\alpha_B) - \frac{1}{6}(\alpha_B - \gamma_B)\right) \leq 0. \tag{41}$$

These constraints can be re-written as

$$\alpha_B + \frac{1}{6}\gamma_B \leq \frac{1}{2(1-q)}\alpha_G;$$

and

$$\alpha_B + \frac{1}{6}\gamma_B \leq \frac{1}{2(1-q)}\alpha_G - \frac{1}{2(1-q)} + \left(1 + \frac{1}{6}\alpha_B\right).$$

We observe that the value of γ_G drops out of the obedience constraints altogether as $y_G = 0$.

BNE Problem We find the BNE by using the above obedience conditions.

There is a BNE with always invest, $\alpha_G = \gamma_G = \alpha_B = \gamma_B = 1$, if and only if

$$\frac{1}{3} + \frac{2}{3}(1-q)\left(-\frac{7}{6}\right) \geq 0 \Leftrightarrow q \geq \frac{4}{7}.$$

There is a BNE with never invest, $\alpha_G = \gamma_G = \alpha_B = \gamma_B = 0$, if and only if

$$\frac{1}{3} + \frac{2}{3}(1-q)(-1) \leq 0 \Leftrightarrow q \leq \frac{1}{2}.$$

If

$$\frac{1}{2} < q < \frac{4}{7},$$

then there is a unique BNE, with $\alpha_B = \alpha_G = \alpha$, as the players receive the same information t_G in both states, θ_B and θ_G, and by the independence of the choice probabilities across player, the BNE has to have $\gamma_G = \alpha_G^2 = \gamma_B = \alpha_B^2 = \alpha^2$. Thus we have:

$$\frac{1}{3}\alpha + \frac{2}{3}(1-q)\left(-\alpha - \frac{1}{6}\alpha^2\right) = 0 \Leftrightarrow \alpha = \frac{3}{1-q} - 6. \tag{42}$$

BCE Problem We are interested in finding the values of α_G and α_B consistent with the obedience constraints. We can, without loss of generality, let γ_B be as small as possible (this slackens *both* constraints) subject to the nonnegativity constraints of the probabilities in (10). These smallest values depend on whether α_B is greater than or less than $\frac{1}{2}$.

Case 1: $\alpha_B \geq \frac{1}{2}$ In this case, $\gamma_B = 2\alpha_B - 1$ by the nonnegativity constraints of (10), and the investment constraint becomes:

$$\alpha_B \leq \frac{3}{8(1-q)}\alpha_G + \frac{1}{8}, \tag{43}$$

whereas the no investment constraint becomes:

$$\alpha_B \leq \frac{3}{7(1-q)}\alpha_G + 1 - \frac{3}{7(1-q)}. \tag{44}$$

Case 2: $\alpha_B \leq \frac{1}{2}$ In this case, $\gamma_B = 0$ and the investment constraint becomes

$$\alpha_B \leq \frac{1}{2(1-q)}\alpha_G, \tag{45}$$

and the no investment constraint becomes

$$\alpha_B \leq \frac{3}{5(1-q)}\alpha_G + \frac{6}{5} - \frac{3}{5(1-q)}. \tag{46}$$

Combining the Cases With the above case distinction for α_B, we have to join the investment constraints and the no investment constraints at the critical points. Beginning with the incentive constraints, we find that the inequalities (43) and (45) solve as equalities at the critical value of $\alpha_B = 1/2$:

$$\frac{1}{2} = \frac{3}{8(1-q)}\alpha_G + \frac{1}{8} \Leftrightarrow \alpha_G = 1 - q$$

and

$$\frac{1}{2} = \frac{1}{2(1-q)}\alpha_G \Leftrightarrow \alpha_G = 1 - q,$$

respectively, and thus the combined investment constraint is:

$$\alpha_B \leq \begin{cases} \frac{1}{2(1-q)}\alpha_G, & \text{if} \quad \alpha_G \leq 1 - q; \\ \frac{3}{8(1-q)}\alpha_G + \frac{1}{8}, & \text{if} \quad \alpha_G > 1 - q. \end{cases} \tag{47}$$

Similarly, we find that the inequalities (44) and (46) solve as equalities at the critical value of $\alpha_B = 1/2$:

$$\frac{1}{2} = \frac{3}{7(1-q)}\alpha_G + 1 - \frac{3}{7(1-q)} \Leftrightarrow \alpha_G = \frac{7}{6}q - \frac{1}{6},$$

and

$$\frac{1}{2} = \frac{3}{5(1-q)}\alpha_G + \frac{6}{5} - \frac{3}{5(1-q)} \Leftrightarrow \alpha_G = \frac{7}{6}q - \frac{1}{6}.$$

So the combined no investment constraint is:

$$\alpha_B \leq \begin{cases} \frac{3}{5(1-q)}\alpha_G + \frac{6}{5} - \frac{3}{5(1-q)}, & \text{if} \quad \alpha_G \leq \frac{7}{6}q - \frac{1}{6}; \\ \frac{3}{7(1-q)}\alpha_G + 1 - \frac{3}{7(1-q)}, & \text{if} \quad \alpha_G > \frac{7}{6}q - \frac{1}{6}. \end{cases} \tag{48}$$

Case Distinction Now, from the combined investment and no investment constraints, we can determine, which one of the constraints is generating the lowest upper bound of α_B as a function of α_G as a function of the information structure, that is of the accuracy of q.

We end up with four relevant regions: two in which only one of the two constraints, (47) and (48), is binding, and two in which the relevant constraints switch as α_G increases.

If $q \leq 1/2$, then the investment constraint (in terms of α_B) is always below the no investment constraint and hence binding. If $q \geq 4/7$, then the no investment constraint is always below the investment constraint, and hence binding. These values coincide with the distinction between pure and mixed strategy BNE above.

The final case distinction is determined by whether the kink in the combined constraints, (47) and (48), arises first in the investment or the no investment constraint:

$$1 - q \geq \frac{7}{6}q - \frac{1}{6} \Leftrightarrow q \leq \frac{7}{13}.$$

Case 1: Very Inaccurate Information: $q \leq \frac{1}{2}$ In this case, the investment constraint is always below the no investment constraint and hence binding. Thus we have that α_G can take any value and α_B any value that satisfies the above inequality (47). For example, if $q = 1/5$, we have

$$\alpha_B \leq \begin{cases} \frac{5}{8}\alpha_G, & \text{if } \alpha_G \leq \frac{4}{5}; \\ \frac{16}{32}\alpha_G + \frac{1}{8}, & \text{if } \alpha_G > \frac{4}{5}. \end{cases}$$

Thus the set of BCE feasible (α_G, α_B) is the convex hull of:

$$\left((0,0), \left(\frac{4}{5}, \frac{1}{2}\right), \left(1, \frac{5}{8}\right), (1,0)\right),$$

and is visually represented in Figure 1.

Case 2: Inaccurate Information: $\frac{1}{2} \leq q \leq \frac{7}{13}$ In this case, the investment constraint binds for low values of α_G and the no investment constraint binds for high values of α_G. We do not represent this case separately as it similar to the next case

Case 3: Accurate Information: $\frac{7}{13} \leq q \leq \frac{4}{7}$ In this case, the no investment constraint binds for low values of α_G and the investment constraint binds for high values of α_G. The switching point occurs for a value of $\alpha_B \geq \frac{1}{2}$. In particular, we have the matching condition:

$$\frac{3}{8(1-q)}\alpha_G + \frac{1}{8} = \frac{3}{7(1-q)}\alpha_G + 1 - \frac{3}{7(1-q)} \Leftrightarrow \alpha_G = \frac{49q-25}{3},$$

and we have $\alpha_G \geq 2q - 1$ and thus:

$$\alpha_B \leq \begin{cases} \frac{3}{5(1-q)}\alpha_G + \frac{6}{5} - \frac{3}{5(1-q)}, & \text{if } 2q-1 \leq \alpha_G \leq \frac{7}{6}q - \frac{1}{6}; \\ \frac{3}{7(1-q)}\alpha_G + 1 - \frac{3}{7(1-q)}, & \text{if } \frac{7}{6}q - \frac{1}{6} \leq \alpha_G \leq \frac{49q-25}{3}; \\ \frac{3}{8(1-q)}\alpha_G + \frac{1}{8}, & \text{if } \frac{49q-25}{3} \leq \alpha_G \leq 1. \end{cases}$$

For example, if $q = \frac{11}{20}$, we have that $\alpha_G \geq \frac{1}{10}$ and

$$\alpha_B \leq \begin{cases} \frac{4}{3}\alpha_G - \frac{2}{15}, & \text{if } 0 \leq \alpha_G \leq \frac{19}{40}; \\ \frac{20}{21}\alpha_G + \frac{1}{21}, & \text{if } \frac{19}{40} \leq \alpha_G \leq \frac{13}{20}; \\ \frac{5}{6}\alpha_G + \frac{1}{8}, & \text{if } \frac{13}{20} \leq \alpha_G \leq 1. \end{cases}$$

Thus the set of BCE feasible (α_G, α_B) is the convex hull of

$$\left(\left(\frac{1}{10}, 0\right), \left(\frac{19}{40}, \frac{1}{2}\right), \left(\frac{13}{20}, \frac{2}{3}\right), \left(1, \frac{23}{24}\right), (1,0)\right),$$

the mixed BNE is $(2/3, 2/3)$ and the graphical representation is given in Figure 2.

Case 4: Very Accurate Information: $\frac{4}{7} \leq q$ In this case, the no investment constraint is always below the investment constraint and therefore the no investment constraint always binds. Thus we have that $\alpha_G \geq 2q - 1$ and (48) holds. For example, if $q = \frac{4}{5}$, we have that $\alpha_G \geq \frac{3}{5}$ and

$$\alpha_B \leq \begin{cases} 3\alpha_G - \frac{9}{5}, & \text{if } \frac{3}{5} \leq \alpha_G \leq \frac{23}{30}; \\ \frac{15}{7}\alpha_G - \frac{8}{7}, & \text{if } \frac{23}{30} \leq \alpha_G \leq 1. \end{cases} \tag{49}$$

Thus the set of BCE is the set of feasible (α_G, α_B) given by the convex hull of

$$\left(\left(\frac{3}{5}, 0\right), \left(\frac{23}{30}, \frac{5}{6}\right), (1,1), (1,0)\right)$$

and a graphical visualization is given by Figure 3, and completes the analysis of the binary game.

6.2 Omitted Definitions and Proofs

Definition 14 (Higher Order Belief Equivalent)

1. *Information structure S is non-redundant if, for every i and $t_i, t_i' \in T_i$, there exists $t_{-i} \in T_{-i}$ and $\theta \in \Theta$ such that $\pi_\psi(t_{-i}, \theta | t_i) \neq \pi_\psi(t_{-i}, \theta | t_i')$ for some (or all) $\psi \in \Delta_{++}(\Theta)$.*

2. *Two information structures $S^1 = \left((T_i^1)_{i=1}^I, \pi^1\right)$ and $S^2 = \left((T_i^2)_{i=1}^I, \pi^2\right)$ are higher-order belief equivalent if there exists a non-redundant information structure $S^* = \left((T_i^*)_{i=1}^I, \pi^*\right)$ such that there exist, for each $i = 1,..,I$ and $k = 1,2$, $f_i^k : T_i^k \to T_i^*$such that:*

   (a) for each $k = 1,2$, $t^* \in T^*$ and $\theta \in \Theta$:

$$\pi^k\left(\left\{t^k \middle| f^k\left(t^k\right) = t^*\right\} \middle| \theta\right) = \pi^*\left(t^* | \theta\right); \tag{50}$$

   (b) for each $k = 1,2$, $i = 1,.,I$, $t_i \in T_i^k$, $t_{-i}^* \in T_{-i}^*$ and $\theta \in \Theta$

$$\pi_\psi^k\left(\left\{t_{-i}^k \middle| f_{-i}^k\left(t_{-i}^k\right) = t_{-i}^*\right\}, \theta \middle| t_i\right) = \pi_\psi^*\left(t_{-i}^*, \theta | f_i^k(t_i)\right). \tag{51}$$

We now present a formal argument that the notion of higher-order belief equivalence presented in Definition 14 is equivalent to a definition in terms of the hierarchical belief types of Mertens and Zamir (1985).

Fix Θ. Let $X^0 = \Theta$, and define $X^k = \Theta \times \left[\Delta\left(X^{k-1}\right)\right]^{I-1}$. An element of $\left(\Delta\left(X^k\right)\right)_{k=0}^\infty \triangleq H$ is called a hierarchy (of beliefs). For notational simplicity, we will work with a uniform prior on Θ (other full support priors will lead to shifts in posteriors over Θ but not change the higher-order belief equivalence). Fix an information structure $S = \left((T_i)_{i=1}^I, \pi \right)$. For each i and $t_i \in T_i$, write $\widehat{\pi}_i^1 [t_i] \in \Delta(\Theta) = \Delta\left(X^0\right)$ for his posterior under a uniform prior on Θ, so

$$\widehat{\pi}_i^1 [t_i](\theta) = \frac{\sum\limits_{t_{-i} \in T_{-i}} \pi((t_i, t_{-i}) | \theta)}{\sum\limits_{\theta' \in \Theta, t_{-i} \in T_{-i}} \pi\left((t_i, t_{-i}) | \theta'\right)}.$$

Write $\widehat{\pi}_i^2 [t_i] \in \Delta\left(\Theta \times (\Delta(\Theta))^{I-1}\right) = \Delta\left(X^1\right)$ for his belief over Θ and the first order beliefs of other players, so

$$\widehat{\pi}_i^2 [t_i]\left(\theta, \pi_{-i}^1\right) = \frac{\sum\limits_{\left\{t_{-i} \in T_{-i} | \widehat{\pi}_j^1(t_j) = \pi_j^1 \text{ for each } j \neq i\right\}} \pi((t_i, t_{-i}) | \theta)}{\sum\limits_{\theta' \in \Theta, t_{-i} \in T_{-i}} \pi\left((t_i, t_{-i}) | \theta'\right)}.$$

Proceeding inductively for $k \geq 2$, write $\widehat{\pi}_i^k (t_i) \in \Delta\left(X^{k-1}\right)$ for his belief over Θ and the $(k-1)$th order beliefs of other players, so

$$\widehat{\pi}_i^k [t_i]\left(\theta, \pi_{-i}^{k-1}\right) = \frac{\sum\limits_{\left\{t_{-i} \in T_{-i} | \widehat{\pi}_j^{k-1}(t_j) = \pi_j^{k-1} \text{ for each } j \neq i\right\}} \pi((t_i, t_{-i}) | \theta)}{\sum\limits_{\theta' \in \Theta, t_{-i} \in T_{-i}} \pi\left((t_i, t_{-i}) | \theta'\right)}.$$

Now we can define $\widehat{\pi}_i : T_i \to H$ by

$$\widehat{\pi}_i [t_i] = \left(\widehat{\pi}_i^1 [t_i], \widehat{\pi}_i^2 [t_i], ...\right)$$

and $\widehat{\pi} : T \to H^I$ by

$$\widehat{\pi}[t] = (\widehat{\pi}_i [t_i])_{i=1}^I.$$

Now we can identify information structure S with a probability distribution $\chi_S \in \Delta\left(H^I\right)$ defined by:

$$\chi_S\left((\pi_i)_{i=1}^I\right) \triangleq \frac{1}{\#\Theta} \sum_{\left\{\theta, t \middle| \widehat{\pi}[t] = (\pi_i)_{i=1}^I\right\}} \pi(t|\theta).$$

Lemma 3 (Higher Order Belief Characterization)
The following statements are equivalent:

1. *Information structures S^1 and S^2 are higher-order belief equivalent;*

2. $\chi_{S^1} = \chi_{S^2}$.

Proof. We argue that (1) implies (2) by induction. Fix $t_i^* \in T_i^*$. By (51),

$$f_i^k (t_i) = t_i^* \Rightarrow \widehat{\pi}_i^{k,1} [t_i] = \widehat{\pi}_i^{*,1} [t_i^*].$$

Now suppose that

$$f_i^k(t_i) = t_i^* \Rightarrow \widehat{\pi}_i^{k,l}[t_i] = \widehat{\pi}_i^{*,l}[t_i^*].$$

By (51), we have

$$f_i^k(t_i) = t_i^* \Rightarrow \widehat{\pi}_i^{k,l+1}[t_i] = \widehat{\pi}_i^{*,l+1}[t_i^*].$$

But since the premise of the inductive step holds for $l = 1$, we have that for all l

$$f_i^k(t_i) = t_i^* \Rightarrow \widehat{\pi}_i^{k,l}[t_i] = \widehat{\pi}_i^{*,l}[t_i^*].$$

and thus

$$f_i^k(t_i) = t_i^* \Rightarrow \widehat{\pi}_i^{k}[t_i] = \widehat{\pi}_i^{*}[t_i^*].$$

Since each non-redundant type t_i^* maps to a distinct hierarchy in H, this establishes (50) implies (2).

Now suppose that (2) holds. Let T_i^* =range$\left(\widehat{\pi}_i^1\right)$ =range$\left(\widehat{\pi}_i^2\right)$. Let $f_i^k(t_i) = \widehat{\pi}_i^k(t_i)$. By construction, properties (50) and (51) hold with respect to information structure $S^* = \left(\left(T_i^*\right)_{i=1}^I, \pi^*\right)$. ■

Proof of Lemma 2. Part (1) of the lemma is stated in Theorem 1 in Liu (2011). Now if information structures $S^1 = \left(T^1, \pi^1\right)$ and $S^2 = \left(T^2, \pi^2\right)$ are higher-order belief equivalent, we can show that S^1 is individually sufficient for S^2 by letting

$$\phi\left(t^2|t^1, \theta\right) \triangleq \begin{cases} \dfrac{\pi^2\left(t^2|\theta\right)}{\sum\limits_{\left\{\widetilde{t}^2|f^2\left(\widetilde{t}^2\right)=f^2\left(t^2\right)\right\}} \pi^2\left(\widetilde{t}^2|\theta\right)}, & \text{if} \quad f^2\left(t^2\right) = f^1\left(t^1\right), \\ 0, & \text{otherwise.} \end{cases}$$

One can similarly show that S^2 is individually sufficient for S^1.

Now suppose that S^1 and S^2 are individually sufficient for each other. If either S^1 or S^2 are redundant, we can replace them with their (by part (1)) unique non-redundant versions, and they will remain mutually individually sufficient. So it is enough to show that if S^1 and S^2 are individually sufficient for each other and non-redundant, then they are higher-order belief equivalent. Write ϕ^1 and ϕ^2 for the Markov kernels establishing that, respectively, S^1 is individually sufficient for S^2 and S^2 is individually sufficient for S^1. Define $\widehat{\phi} : T^1 \times \Theta \to \Delta\left(T^1\right)$ by

$$\widehat{\phi}\left(\widetilde{t}^1\left|t^1, \theta\right.\right) \triangleq \sum_{t^2 \in T^2} \phi^1\left(t^2\left|t^1, \theta\right.\right)\phi^2\left(\widetilde{t}^1\left|t^2, \theta\right.\right)$$

for all $t^1, \widetilde{t}^1 \in T^1$ and $\theta \in \Theta$. It inherits the properties that

$$\begin{aligned}\pi^1\left(t^1|\theta\right) &= \sum_{\widetilde{t^1}\in T^1}\pi^1\left(\widetilde{t^1}|\theta\right)\widehat{\phi}\left(t^1\left|\widetilde{t^1},\theta\right.\right) \text{ and}\\ \widehat{\phi}_i\left(\widetilde{t}_i^1\left|t_i^1\right.\right) &= \sum_{\widetilde{t}_{-i}^1\in T_{-i}^1}\widehat{\phi}\left(\left(\widetilde{t}_i^1,\widetilde{t}_{-i}^1\right)\left|\left(t_i^1,t_{-i}^1\right),\theta\right.\right)\end{aligned}$$

is independent of $\left(t_{-i}^1,\theta\right)$.

Define a partition of T_i^1 by

$$P_i\left(t_i\right)\triangleq\left\{\widetilde{t}_i\in T_i^1\left|\begin{array}{l}\text{there exists }\left(t_i^k\right)_{k=1}^K\text{ with}\\ \widehat{\phi}_i\left(\widetilde{t}_i^{k+1}\left|t_i^k\right.\right)>0\text{ for each }k=1,...,K-1\\ \text{and }\left(t_i^k,t_i^K\right)=\text{either }\left(t_i,\widetilde{t}_i\right)\text{ or }\left(\widetilde{t}_i,t_i\right)\end{array}\right.\right\}.$$

Now

$$\begin{aligned}\sum_{\widehat{t}_{-i}\in P_{-i}(t_{-i})}\pi\left(\left(t_i,\widehat{t}_{-i}\right)|\theta\right) &= \sum_{\widetilde{t}\in T}\pi\left(\widetilde{t}|\theta\right)\sum_{\widehat{t}_{-i}\in P_{-i}(t_{-i})}\widehat{\phi}\left(\left(t_i,\widehat{t}_{-i}\right)\left|\left(\widetilde{t}_i,\widetilde{t}_{-i}\right),\theta\right.\right)\\ &= \sum_{\widetilde{t}_i\in P_i(t_i)}\sum_{\widetilde{t}_{-i}\in P_{-i}(t_{-i})}\pi\left(\left(\widetilde{t}_i,\widetilde{t}_{-i}\right)\right|\theta\right)\sum_{\widehat{t}_{-i}\in P_{-i}(t_{-i})}\widehat{\phi}\left(\left(t_i,\widehat{t}_{-i}\right)\left|\left(\widetilde{t}_i,\widetilde{t}_{-i}\right),\theta\right.\right)\\ &= \sum_{\widetilde{t}_i\in P_i(t_i)}\sum_{\widetilde{t}_{-i}\in P_{-i}(t_{-i})}\pi\left(\left(\widetilde{t}_i,\widetilde{t}_{-i}\right)\right|\theta\right)\sum_{\widehat{t}_{-i}\in T_{-i}}\widehat{\phi}\left(\left(t_i,\widehat{t}_{-i}\right)\left|\left(\widetilde{t}_i,\widetilde{t}_{-i}\right),\theta\right.\right)\\ &= \sum_{\widetilde{t}_i\in P_i(t_i)}\sum_{\widetilde{t}_{-i}\in P_{-i}(t_{-i})}\pi\left(\left(\widetilde{t}_i,\widetilde{t}_{-i}\right)\right|\theta\right)\widehat{\phi}_i\left(t_i\left|\widetilde{t}_i\right.\right)\\ &= \sum_{\widetilde{t}_i\in P_i(t_i)}\widehat{\phi}_i\left(t_i\left|\widetilde{t}_i\right.\right)\sum_{\widetilde{t}_{-i}\in P_{-i}(t_{-i})}\pi\left(\left(\widetilde{t}_i,\widetilde{t}_{-i}\right)\right|\theta\right).\end{aligned}$$

Thus for any $\psi\in\Delta_{++}\left(\Theta\right)$,

$$\psi\left(\theta\right)\sum_{\widehat{t}_{-i}\in P_{-i}(t_{-i})}\pi\left(\left(t_i,\widehat{t}_{-i}\right)|\theta\right)=\psi\left(\theta\right)\sum_{\widetilde{t}_i\in P_i(t_i)}\widehat{\phi}_i\left(t_i\left|\widetilde{t}_i\right.\right)\sum_{\widetilde{t}_{-i}\in P_{-i}(t_{-i})}\pi\left(\left(\widetilde{t}_i,\widetilde{t}_{-i}\right)\right|\theta\right).$$

Writing

$$\lambda_i\left(t_i\right)=\sum_{t_{-i}\in T_{-i},\theta\in\Theta}\psi\left(\theta\right)\pi\left(\left(t_i,t_{-i}\right)|\theta\right),$$

we have

$$\sum_{\widetilde{t}_{-i}\in P_{-i}(t_{-i})}\pi_\psi\left(\widetilde{t}_{-i},\theta|t_i\right)=\frac{1}{\lambda_i\left(t_i\right)}\sum_{\widetilde{t}_i\in P_i(t_i)}\widehat{\phi}_i\left(t_i\left|\widetilde{t}_i\right.\right)\lambda_i\left(\widetilde{t}_i\right)\sum_{\widetilde{t}_{-i}\in P_{-i}(t_{-i})}\pi_\psi\left(\widetilde{t}_{-i},\theta|\widetilde{t}_i\right).$$

This condition states that posteriors over $\left(P_{-i}\left(t_{-i}\right),\theta\right)$ for t_i are a weighted sum of posteriors for $\widetilde{t}_i\in P_i\left(t_i\right)$. This implies that all have the same beliefs. If the information structure is non-redundant, this implies that each $\widehat{\phi}_i$ must be the identity function. But this implies that ϕ^1 and ϕ^2 are identities and thus S^1 and S^2 are higher-order belief equivalent. ■

Proof of Proposition 1. Suppose that S is individually sufficient for S'. Thus there exists $\phi : T \times \Theta \to \Delta(T')$ such that

$$\sum_t \pi(t|\theta)\, \phi(t'|t,\theta) = \pi'(t'|\theta) \tag{52}$$

for each t' and θ, and

$$\sum_{t'_{-i}} \phi\left(\left(t'_i, t'_{-i}\right) | \left(t_i, t_{-i}\right), \theta\right) \tag{53}$$

is independent of t_{-i} and θ. Let $S^* = (T^*, \pi^*)$ be the combined information structure with $T_i^* = T_i \times T'_i$ for each i and

$$\pi^*\left(t, t'|\theta\right) = \pi(t|\theta)\, \phi\left(t'|t,\theta\right) \tag{54}$$

for each t, t' and θ.

We will first show that S is individually sufficient for S^*. To do so, define $\phi^* : T \times \Theta \to \Delta(T^*)$ by

$$\phi^*(t^*|t,\theta) = \phi^*\left(\left(\widetilde{t}, t'\right)|t,\theta\right) = \begin{cases} \phi(t'|t,\theta), & \text{if } \ \widetilde{t} = t, \\ 0, & \text{if } \ \widetilde{t} \neq t, \end{cases} \tag{55}$$

for each $t^* = \left(\widetilde{t}, t'\right) \in T^*$, t and θ. Observe that

$$\begin{aligned} \sum_t \pi(t|\theta)\, \phi\left(\left(\widetilde{t}, t'\right)|t,\theta\right) &= \pi\left(\widetilde{t}|\theta\right) \phi\left(t'|\widetilde{t},\theta\right), \text{ by (55)} \\ &= \pi^*\left(\widetilde{t}, t'|\theta\right), \text{ by (54).} \end{aligned}$$

for each $\widetilde{t}$, t' and θ. Also observe that

$$\begin{aligned} \sum_{t^*_{-i}} \phi^*\left(\left(t^*_i, t^*_{-i}\right) | \left(t_i, t_{-i}\right), \theta\right) &= \sum_{\widetilde{t}_{-i}, t'_{-i}} \phi^*\left(\left(\left(\widetilde{t}_i, t'_i\right), \left(\widetilde{t}_{-i}, t'_{-i}\right)\right) | \left(t_i, t_{-i}\right), \theta\right) \\ &= \begin{cases} \sum\limits_{t'_{-i} \in T_{-i}} \phi^*\left(\left(t'_i, t'_{-i}\right) | \left(t_i, t_{-i}\right), \theta\right), & \text{if } \ \widetilde{t} = t, \\ 0, & \text{if } \ \widetilde{t} \neq t, \end{cases} \end{aligned}$$

is independent of t_{-i} and θ by (53).

We will now show that S^* is individually sufficient for S. To do so, define $\widehat{\phi} : T^* \times \Theta \to \Delta(T)$ by

$$\widehat{\phi}(t|t^*,\theta) = \widehat{\phi}\left(t|\left(\widetilde{t}, t'\right), \theta\right) = \begin{cases} 1, & \text{if } \ \widetilde{t} = t, \\ 0, & \text{if } \ \widetilde{t} \neq t, \end{cases} \tag{56}$$

for each $t^* = \left(\widetilde{t}, t'\right) \in T^*$, t and θ. Observe that

$$\begin{aligned}
\sum_{t^8 \in T^*} \pi^*\left(t^*|\theta\right) \widehat{\phi}\left(t|t^*, \theta\right) &= \sum_{\left(\widetilde{t}, t'\right) \in T^*} \pi^*\left(\widetilde{t}, t'|\theta\right) \widehat{\phi}\left(t|\left(\widetilde{t}, t'\right), \theta\right) \\
&= \sum_{t' \in T'} \pi^*\left(t, t'|\theta\right), \text{ by (56)} \\
&= \sum_{t' \in T'} \pi\left(t|\theta\right) \phi\left(t'|t, \theta\right), \text{ by (54)} \\
&= \pi\left(t|\theta\right)
\end{aligned}$$

for each t and θ. Also observe that

$$\begin{aligned}
\sum_{t_{-i}} \widehat{\phi}\left(\left(t_i, t_{-i}\right) | \left(t_i^*, t_{-i}^*\right), \theta\right) &= \sum_{t_{-i}} \widehat{\phi}\left(\left(t_i, t_{-i}\right) | \left(\left(\widetilde{t}_i, t_i'\right), \left(\widetilde{t}_{-i}, t_{-i}'\right)\right), \theta\right) \\
&= \begin{cases} 1, & \text{if } \widetilde{t}_i = t_i, \\ 0, & \text{if } \widetilde{t}_i \neq t_i, \end{cases}
\end{aligned}$$

is independent of t_{-i} and θ.

We have now shown that if S is individually sufficient for S' then there exists an expansion of S', S^*, such that S and S^* are mutually individually sufficient. By Lemma 2, S and S^* are higher-order belief equivalent.

Conversely, suppose that S is higher-order belief equivalent to an expansion of S'. Let us call that expansion $S^* = \left(\left(T_i' \times T_i^+\right)_{i=1}^I, \pi^*\right)$. By Lemma 2, S is individually sufficient for S^*. Thus there exists $\phi^* : T \times \Theta \to \Delta\left(T^*\right)$ such that

$$\sum_t \pi\left(t|\theta\right) \phi^*\left(t^*|t, \theta\right) = \pi^*\left(t^*|\theta\right) \tag{57}$$

for each t^* and θ, and

$$\sum_{t_{-i}^*} \phi^*\left(\left(t_i^*, t_{-i}^*\right) | \left(t_i, t_{-i}\right), \theta\right) \tag{58}$$

is independent of t_{-i} and θ. Define $\phi : T \times \Theta \to \Delta\left(T'\right)$ by

$$\phi\left(t'|t, \theta\right) = \sum_{t^+} \phi^*\left(\left(t', t^+\right)|t, \theta\right) \tag{59}$$

for each t, t' and θ. Now

$$\begin{aligned}
\sum_t \pi\left(t|\theta\right) \phi\left(t'|t, \theta\right) &= \sum_{t^+} \sum_t \pi\left(t|\theta\right) \phi^*\left(\left(t', t^+\right)|t, \theta\right), \text{ by (59)} \\
&= \sum_{t^+} \pi^*\left(\left(t', t^+\right)|\theta\right), \text{ by (59)} \\
&= \pi'\left(t'|\theta\right), \text{ because } S^* \text{ is an expansion of } S'
\end{aligned}$$

for each t' and θ. Also

$$\begin{aligned}\sum_{t'_{-i}} \phi\left(\left(t'_i, t'_{-i}\right) | \left(t_i, t_{-i}\right), \theta\right) &= \sum_{t^+} \sum_{t'_{-i}} \phi^* \left(\left(t'_i, t^+_{-i}\right), \left(t'_i, t^+_{-i}\right) | \left(t_i, t_{-i}\right), \theta\right) \\ &= \sum_{t^+_i} \sum_{t^*_{-i}} \phi^* \left(\left(t'_i, t^+_i\right), t^*_{-i} | \left(t_i, t_{-i}\right), \theta\right)\end{aligned}$$

which is independent of t_{-i} and θ by (58). ∎

In the proof of Theorem 2, we appeal to the transitivity of individual sufficiency.

Lemma 4 (Transitivity of Individual Sufficiency)
If S is individually sufficient for S' and S' is individually sufficient for S'', then S is individually sufficient for S''.

Proof of Lemma 4. If S is sufficient for S', then there exists $\phi : T \times \Theta \to \Delta(T')$ such that

$$\sum_t \pi(t|\theta) \phi\left(t'|t, \theta\right) = \pi'\left(t'|\theta\right)$$

for each t' and θ, and

$$\phi_i\left(t'_i|t_i\right) \equiv \sum_{t'_{-i}} \phi\left(\left(t'_i, t'_{-i}\right) | \left(t_i, t_{-i}\right), \theta\right)$$

is independent of t_{-i} and θ. If S' is individually sufficient for S'', then there exists $\phi' : T' \times \Theta \to \Delta(T'')$ such that

$$\sum_{t'} \pi'\left(t'|\theta\right) \phi'\left(t''|t', \theta\right) = \pi''\left(t''|\theta\right)$$

for each t'' and θ, and

$$\phi'_i\left(t''_i|t'_i\right) \equiv \sum_{t''_{-i}} \phi'\left(\left(t''_i, t''_{-i}\right) | \left(t'_i, t'_{-i}\right), \theta\right)$$

is independent of t'_{-i} and θ. Define $\phi^* : T \times \Theta \to \Delta(T'')$ by

$$\phi^*\left(t''|t, \theta\right) = \sum_{t' \in T'} \phi\left(t'|t, \theta\right) \phi'\left(t''|t', \theta\right).$$

Now

$$\begin{aligned}\phi^*_i\left(t''_i|t_i\right) &\equiv \sum_{t''_{-i}} \phi^*\left(\left(t''_i, t''_{-i}\right) | \left(t_i, t_{-i}\right), \theta\right) \\ &= \sum_{t''_{-i}} \sum_{t' \in T'} \phi\left(t'|t, \theta\right) \phi'\left(t''|t', \theta\right) \\ &= \sum_{t' \in T'} \phi\left(t'|t, \theta\right) \phi'_i\left(t''_i|t'_i\right) \\ &= \sum_{t'_i \in T'_i} \phi_i\left(t'_i|t_i\right) \phi'_i\left(t''_i|t'_i\right)\end{aligned}$$

which is independent of t_{-i} and θ. ■

References

Aumann, R. (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics*, 1, 67–96.

Aumann, R. (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55, 1–18.

Aumann, R., and M. Maschler (1995): *Repeated Games with Incomplete Information*. MIT.

Bergemann, D., B. Brooks, and S. Morris (2013): "The Limits of Price Discrimination," Discussion Paper 1896R, Cowles Foundation for Research in Economics, Yale University.

Bergemann, D., and S. Morris (2007): "Belief Free Incomplete Information Games," Discussion Paper 1629, Cowles Foundation for Research in Economics, Yale University.

——— (2012): *Robust Mechanism Design*. World Scientific Publishing, Singapore.

——— (2013): "Robust Predictions in Games with Incomplete Information," *Econometrica*, 81, 1251–1308.

Blackwell, D. (1951): "Comparison of Experiments," in *Proc. Second Berkeley Symp. Math. Statist. Probab.*, pp. 93–102. University of California Press, Berkeley.

——— (1953): "Equivalent Comparison of Experiments," *Annals of Mathematics and Statistics*, 24, 265–272.

Calvó-Armengol, A. (2006): "The Set of Correlated Equilibria of 2x2 Games," Discussion paper, Universitat Autònoma de Barcelona.

Caplin, A., and D. Martin (2013): "A Testable Theory of Imperfect Perception," New York University.

Chwe, M. (2006): "Incentive Compatibility Implies Signed Covariance," Discussion paper, UCLA.

Dekel, E., D. Fudenberg, and S. Morris (2007): "Interim Correlated Rationalizability," *Theoretical Economics*, 2, 15–40.

Ely, J. C., and M. Peski (2006): "Hierarchies of Belief and Interim Rationalizability," *Theoretical Economics*, 1, 19–65.

Forges, F. (1993): "Five Legitimate Definitions of Correlated Equilibrium in Games with Incomplete Information," *Theory and Decision*, 35, 277–310.

——— (2006): "Correlated Equilibrium in Games with Incomplete Information Revisited," *Theory and Decision*, 61, 329–344.

Gossner, O. (2000): "Comparison of Information Structures," *Games and Economic Behavior*, 30, 44–63.

Gossner, O., and J. Mertens (2001): "The Value of Information in Zero-Sum Games," Universite Paris-Nanterre and CORE, University Catholique de Louvain.

Hirshleifer, J. (1971): "The Private and Social Value of Information and the Reward to Inventive Activity," *American Economic Review*, 61, 561–574.

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

LeCam, L. (1996): "Comparison of Experiments - A Short Review," *IMS Lecture Notes*, 30, 127–138.

Lehrer, E., D. Rosenberg, and E. Shmaya (2010): "Signaling and Mediation in Games with Common Interest," *Games and Economic Behavior*, 68, 670–682.

——— (2013): "Garbling of Signals and Outcome Equivalence," *Games and Economic Behavior*, 81, 179–191.

Liu, Q. (2011): "Correlation and Common Priors in Games with Incomplete Information," Discussion paper, Columbia University.

Mertens, J., and S. Zamir (1985): "Formalization of Bayesian Analysis for Games with Incomplete Information," *International Journal of Game Theory*, 14, 1–29.

Milchtaich, I. (2012): "Implementability of Correlated and Communication Equilibrium Outcomes in Incomplete Information Games. Part I: Definitions and Results," Discussion paper, Bar Ilan University.

Neyman, A. (1991): "The Positive Value of Information," *Games and Economic Behavior*, 3, 350–355.

Peski, M. (2008): "Comparison of Information Structures in Zero-Sum Games," *Games and Economic Behavior*, 62, 732–735.

Torgersen, E. (1991): *Comparison of Statistical Experiments*. Cambridge University Press, Cambridge.