Can Hidden Variables Explain Correlation?*

Adam Brandenburger† Amanda Friedenberg‡

First version: June 2004
Current version: October 2005

Abstract

Correlations arise naturally in non-cooperative games, e.g., in the equivalence between undominated and optimal strategies in games with more than two players. But the non-cooperative assumption is that players do not coordinate their strategy choices, so where do these correlations come from? In coin tossing, correlated assessments are usually understood as reflecting the observer's ignorance of a hidden variable–viz., the coin's parameter or bias. We take a similar view of correlation in games. The hidden variables are the players' characteristics that aren't already described in the matrix–viz., their hierarchies of beliefs (i.e., their beliefs about what strategies will be played, their beliefs about other players' beliefs, etc.). A player may then consider other players' strategies to be correlated, if he considers their hierarchies to be correlated. We analyze this hidden-variable view of game-theoretic correlation.

1 Introduction

Correlation is basic to game theory. For example, consider the equivalence between undominated strategies–strategies that are not strongly dominated–and strategies that are optimal under some measure on the strategy profiles of the other players. As is well known, for this to hold in games with more than two players, the measure may need to be dependent (i.e., correlated).

Figure 1.1 is similar to Examples 2.5 and 2.6 in Aumann [1, 1974], which introduced the study of correlation. Here, Ann chooses the row, Bob chooses the column, Charlie chooses the matrix, and the payoffs shown are to Charlie. The strategy Y is optimal under a measure that puts probability

---

*Our thanks to Geir Asheim, Pierpaolo Battigalli, Ethan Bueno de Mesquita, Philip Dawid, Yossi Feinberg, Jerry Green, Qingmin Liu, Paul Milgrom, Andres Perea, John Pratt, Michael Schwarz, Jeroen Swinkels, Gus Stuart, Daniel Yamins, and audiences at the 2005 Econometric Society World Congress, 2005 SAET Conference, University College London Conference in Honor of Ken Binmore (August 2005), University of Leipzig, Notre Dame, Pittsburgh, Stanford, and Washington University for important input. Financial support from the Stern School of Business and the Olin School of Business is gratefully acknowledged. chvec-10-15-05

†Address: Stern School of Business, New York University, New York, NY 10012, adam.brandenburger@stern.nyu.edu, www.stern.nyu.edu/∼abranden

‡Address: Olin School of Business, Washington University, St. Louis, MO 63130, friedenberg@wustl.edu, www.olin.wustl.edu/faculty/friedenberg

$\frac{1}{2}$ on (U, L) and probability $\frac{1}{2}$ on (D, R), and so is undominated. But there is no product measure under which Y is optimal.[1]

	L	R
U	3	3
D	0	0

X

	L	R
U	2	0
D	0	2

Y

	L	R
U	0	0
D	3	3

Z

Figure 1.1

Where does a correlated assessment–such as probability $\frac{1}{2}$ on (U, L) and probability $\frac{1}{2}$ on (D, R)–come from? After all, the non-cooperative assumption is that players do not coordinate their strategy choices. Alternatively put, there is no 'physical' correlation across players.

Coin tossing provides an analogy. While there is no physical correlation across tosses, correlated assessments are possible. These are usually thought of as reflecting the observer's ignorance of a hidden variable–namely, the coin's parameter or bias. The usual requirement is that the tosses are i.i.d. conditional on a parameter on which the observer has a prior.

Here we take a similar view of correlation in games. For a given game, the hidden variables must be players' characteristics that are not described in the matrix. In keeping with the epistemic literature, we take these characteristics to be the players' hierarchies of beliefs about the play of the game–i.e., their beliefs about what strategies other players will choose, their beliefs about other players' beliefs, etc. We'll see that a player may then have a correlated assessment about the strategies other players will choose, if he has a correlated assessment about their hierarchies of beliefs.

Below, we formalize the idea that the hidden variables in a game are the players' hierarchies of beliefs about the strategies played in the game. We show that correlations arise in non-cooperative game theory that cannot be explained by these hidden variables. We go on to interpret this finding.

2 An Example

Return to the decision problem in Figure 1.1 and turn it into a game by adding payoffs for Ann and Bob.[2] We want to formalize the idea that the hidden variables are the players' hierarchies of beliefs over the strategies chosen in the game. In line with the epistemic literature, we describe these hierarchies via a type structure.

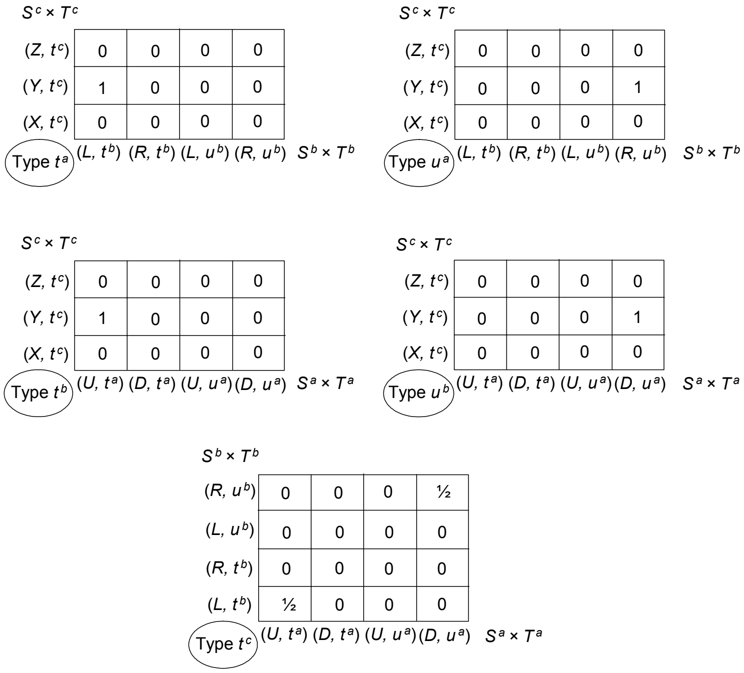
In Figure 2.1, there are two types of Ann, viz. t^a and u^a, two types of Bob, viz. t^b and u^b, and one type of Charlie, viz. t^c. Each type is associated with a measure on the strategies and types of other players. In the usual way, each type induces a hierarchy of beliefs about the strategies played

---

[1] Let p be the probability on U, and q be the probability on L. It is straightforward to check that $\max\{3p, 3(1-p)\} > 2pq + 2(1-p)(1-q)$.

[2] For now, any payoffs can be added.

in the game.



Type t^a

$S^c \times T^c$	(L, t^b)	(R, t^b)	(L, u^b)	(R, u^b)
(Z, t^c)	0	0	0	0
(Y, t^c)	1	0	0	0
(X, t^c)	0	0	0	0

$S^b \times T^b$

Type u^a

$S^c \times T^c$	(L, t^b)	(R, t^b)	(L, u^b)	(R, u^b)
(Z, t^c)	0	0	0	0
(Y, t^c)	0	0	0	1
(X, t^c)	0	0	0	0

$S^b \times T^b$

Type t^b

$S^c \times T^c$	(U, t^a)	(D, t^a)	(U, u^a)	(D, u^a)
(Z, t^c)	0	0	0	0
(Y, t^c)	1	0	0	0
(X, t^c)	0	0	0	0

$S^a \times T^a$

Type u^b

$S^c \times T^c$	(U, t^a)	(D, t^a)	(U, u^a)	(D, u^a)
(Z, t^c)	0	0	0	0
(Y, t^c)	0	0	0	1
(X, t^c)	0	0	0	0

$S^a \times T^a$

Type t^c

$S^b \times T^b$	(U, t^a)	(D, t^a)	(U, u^a)	(D, u^a)
(R, u^b)	0	0	0	½
(L, u^b)	0	0	0	0
(R, t^b)	0	0	0	0
(L, t^b)	½	0	0	0

$S^a \times T^a$

Figure 2.1

For example, type t^c assigns probability $\frac{1}{2}$ to (U, t^a, L, t^b) and probability $\frac{1}{2}$ to (D, u^a, R, u^b), and so has a first-order belief that assigns probability $\frac{1}{2}$ to Ann's playing U and Bob's playing L, and probability $\frac{1}{2}$ to Ann's playing D and Bob's playing R. Type t^a has a first-order belief that assigns probability 1 to (L, Y), and type u^a has a first-order belief that assigns probability 1 to (R, Y). Type t^b has a first-order belief that assigns probability 1 to (U, Y), and type u^b has a first-order belief that assigns probability 1 to (D, Y). So type t^c has a second-order belief that assigns: (i) probability $\frac{1}{2}$ to "Ann's playing U and assigning probability 1 to (L, Y)", and "Bob's playing L and assigning probability 1 to (U, Y)"; and (ii) probability $\frac{1}{2}$ to "Ann's playing D and assigning probability 1 to (R, Y)", and "Bob's playing R and assigning probability 1 to (D, Y)". And so on.

Also notice that the two types of Ann, viz. t^a and u^a, induce distinct hierarchies of beliefs about the play of the game. While type t^a has a first-order belief that assigns probability 1 to (L, Y), type u^a has a first-order belief that assigns probability 1 to (R, Y). Similarly, types t^b and u^b induce different hierarchies of beliefs.

We point to two particular features of the example:

Conditional Independence (CI) The measure associated with each type assesses the other players' strategy choices as independent, conditional on their hierarchies of beliefs. Consider, for example, type t^c of Charlie and the associated measure $\lambda^c(t^c)$. Then conditioning on the hierarchies of beliefs of Ann and Bob, we have

$$\lambda^c(t^c)(U, L|t^a, t^b) = \lambda^c(t^c)(U|t^a, t^b) \times \lambda^c(t^c)(L|t^a, t^b),$$

since types induce distinct hierarchies. Similarly for the hierarchies associated with types u^a, u^b. So type t^c of Charlie assesses Ann's and Bob's strategies as independent conditional on their hierarchies. CI ensures that the correlations in the game come only from–more precisely, are believed to come only from–the hidden variables, and aren't 'physical.' Do note that Charlie doesn't assess Ann's and Bob's strategies as (unconditionally) independent. For example: $\lambda^c(t^c)(U, L) = \frac{1}{2} \neq \lambda^c(t^c)(U) \times \lambda^c(t^c)(L) = \frac{1}{2} \times \frac{1}{2}$.

Sufficiency (SUFF) We have

$$\lambda^c(t^c)(U \mid t^a, t^b) = \lambda^c(t^c)(U \mid t^a),$$

and similarly for D. If Charlie knows Ann's hierarchy of beliefs, and comes to learn Bob's hierarchy of beliefs too, this won't change his (Charlie's) assessment of Ann's choice. Likewise

$$\lambda^c(t^c)(L \mid t^a, t^b) = \lambda^c(t^c)(L \mid t^b),$$

and similarly for R. If Charlie knows Bob's hierarchy of beliefs, and comes to learn Ann's hierarchy of beliefs too, this won't change his assessment of Bob's choice. These conditions capture the idea that, from Charlie's perspective, Ann can't condition her own choice on more than she knows (viz., her own hierarchy), and Bob can't condition his own choice on more than he knows (viz., his own hierarchy).[3]

Taken together, CI and SUFF capture the idea that the hidden variables in a game are precisely the players' hierarchies of beliefs (about the strategies the players choose). In particular, we will see that under these conditions, if a player has a correlated assessment about the strategies chosen by the other players, he must have a correlated assessment about their hierarchies of beliefs. In Section 5 we define CI and SUFF, and give a formal statement of this result.

3 The Question

We want to understand the implications of this hidden-variable view of correlations for the analysis of games. In short, we want to know how CI and SUFF affect a game.

Consider Figure 3.1, and again associated type structure in Figure 2.1. For Ann, the strategy-type pairs (U, t^a) and (D, u^a) are rational–each strategy maximizes Ann's expected payoff under the corresponding type. Similarly, (L, t^b), (R, u^b) are rational for Bob; and (Y, t^c) is rational for

---

[3] The sufficiency conditions reflect the game–i.e., multi-player–context. There are different hidden variables for the different players, and we have to 'attach' the right variable to the right player.

Charlie. Also, each type of each player assigns positive probability only to rational strategy-type pairs of the other players. That is, each player believes the other players are rational. By induction, each of these strategy-type pairs is therefore consistent with **rationality and common belief of rationality** (**RCBR**). (For formal definitions, see Section 6 below.)

	L	R
U	1,1,3	1,0,3
D	0,1,0	0,0,0

X

	L	R
U	1,1,2	0,0,0
D	0,0,0	1,1,2

Y

	L	R
U	1,1,0	1,0,0
D	0,1,3	0,0,3

Z

Figure 3.1

The type structure in Figure 2.1 satisfies CI and SUFF. We see that under these conditions, the prediction of RCBR is quite different from rationalizability (Bernheim [6, 1984], Pearce [18, 1984]).[4] While all strategy profiles are consistent with RCBR under these conditions, only (U, L, X) is rationalizable.[5]

What then is the prediction of CI, SUFF, and RCBR? A natural candidate is the strategies that survive iterated strong dominance–or the **iteratively undominated** (**IU**) strategies. In Figure 3.1, all strategies are IU. More generally, RCBR is characterized by IU. (See Proposition 6.1 for the formal statement.) So, the question is whether adding CI and SUFF changes the answer. Certainly, under these conditions, the strategies played at a state where there is RCBR must be IU. But, we will show that the converse fails. Specifically: *There is a game G and an IU strategy in G that cannot be played under CI, SUFF, and RCBR.*

Sections 4-7 below are devoted to formulating and proving this result. The remainder of the paper discusses implications for the analysis of games. Section 8 lays out the general relationship between: (i) the strategies we focus on, viz. the strategies consistent with RCBR when the players' hierarchies of beliefs are hidden variables; (ii) the rationalizable strategies; (iii) the IU strategies; and (iv) the correlated-equilibrium strategies (Aumann [1, 1974], [2, 1987]). Section 9 relates our analysis to a Savage Small-Worlds [19, 1954, pp.82-91] vs. Classical view of game theory, and also discusses how it differs from an analysis with physical correlation. We conclude in Section 10 with an open question.

4 Interactive Probability Structures

Given a Polish space Ω, let $\mathcal{B}(\Omega)$ be the Borel σ-algebra on Ω. Also, write $\mathcal{M}(\Omega)$ for the space of all Borel probability measures on Ω, where $\mathcal{M}(\Omega)$ is endowed with the topology of weak convergence (and so is again Polish).

---

[4] We follow the original definitions in [6, 1984] and [18, 1984]: The rationalizable strategies are those that survive iterated deletion of strategies that are not optimal under any product measure.

[5] So, Charlie can only get an (expected) payoff of 3–and not 2 as is possible from Y.

Given sets $X^1, \ldots, X^n$, write $X = \times_{i=1}^n X^i$, $X^{-i} = \times_{j \neq i} X^j$, and $X^{-i-j} = \times_{k \neq i,j} X^k$. (Throughout, we adopt the convention that in a product X, if some $X^i = \emptyset$ then $X^i = \emptyset$ for all i.) An n-**player strategic-form game** is given by $G = \left\langle S^1, \ldots, S^n; \pi^1, \ldots, \pi^n \right\rangle$, where S^i is player i's finite strategy set and $\pi^i : S \to \mathbb{R}$ is i's payoff function. Extend π^i to $\mathcal{M}(S)$ in the usual way.

Definition 4.1 *Fix an n-player strategic-form game G. An* $(S^1, \ldots, S^n)$-***based interactive probability structure*** *is a structure*

$$\Phi = \langle S^1, \ldots, S^n; T^1, \ldots, T^n; \lambda^1, \ldots, \lambda^n \rangle,$$

where each T^i *is a Polish space and each* $\lambda^i : T^i \to \mathcal{M}(S^{-i} \times T^{-i})$ *is continuous. Members of* T^i *are called* ***types*** *of player i. Members of* $S \times T$ *are called* ***states****.*

Associated with each type t^i of each player i in a structure Φ is a hierarchy of beliefs about the strategies played.[6] To see this, inductively define sets Y_m^i, by setting $Y_1^i = S^{-i}$ and

$$Y_{m+1}^i = Y_m^i \times \times_{j \neq i} \mathcal{M}(Y_m^j).$$

Now define continuous maps $\rho_m^i : S^{-i} \times T^{-i} \to Y_m^i$ inductively by

$$\begin{aligned} \rho_1^i \left(s^{-i}, t^{-i}\right) &= s^{-i}, \\ \rho_{m+1}^i \left(s^{-i}, t^{-i}\right) &= (\rho_m^i \left(s^{-i}, t^{-i}\right), (\delta_m^j \left(t^j\right))_{j \neq i}), \end{aligned}$$

where $\delta_m^j = \underline{\rho}_m^j \circ \lambda^j$ and, for each $\mu \in \mathcal{M}\left(S^{-j} \times T^{-j}\right)$, $\underline{\rho}_m^j(\mu)$ is the image measure under ρ_m^j. (Appendix A shows that these maps are indeed continuous and so are well-defined.) Define a continuous (Appendix A) map $\delta^i : T^i \to \times_{m=1}^{\infty} \mathcal{M}\left(Y_m^i\right)$ by $\delta^i\left(t^i\right) = \left(\delta_1^i\left(t^i\right), \delta_2^i\left(t^i\right), \ldots\right)$. In words, $\delta^i\left(t^i\right)$ is simply the hierarchy of beliefs (about strategies) induced by type t^i.

For each player i, define a map $\delta^{-i} : T^{-i} \to \times_{j \neq i} \times_{m=1}^{\infty} \mathcal{M}\left(Y_m^j\right)$ by

$$\delta^{-i}\left(t^1, \ldots, t^{i-1}, t^{i+1}, \ldots, t^n\right) = \left(\delta^1\left(t^1\right), \ldots, \delta^{i-1}\left(t^{i-1}\right), \delta^{i+1}\left(t^{i+1}\right), \ldots, \delta^n\left(t^n\right)\right).$$

Since each δ^j is continuous, δ^{-i} is continuous.

5 Conditional Independence and Sufficiency

Fix a structure Φ, and a player $i = 1, \ldots, n$. For each $j \neq i$, define random variables $\overrightarrow{s}_i^j$ and $\overrightarrow{t}_i^j$ on $S^{-i} \times T^{-i}$ by $\overrightarrow{s}_i^j = \mathrm{proj}_{S^j}$ and $\overrightarrow{t}_i^j = \mathrm{proj}_{T^j}$. (Here proj denotes the projection map.) Let $\overrightarrow{s}_i$ and $\overrightarrow{t}_i$ be random variables on $S^{-i} \times T^{-i}$ with $\overrightarrow{s}_i = \mathrm{proj}_{S^{-i}}$ and $\overrightarrow{t}_i = \mathrm{proj}_{T^{-i}}$. Also, define the composite maps $\eta_i^j = \delta^j \circ \overrightarrow{t}_i^j$ and $\eta^{-i} = \delta^{-i} \circ \overrightarrow{t}_i$.

---

[6] The formulation below closely follows Mertens-Zamir [15, 1985, Section 2] and Battigalli-Siniscalchi [3, 1999, Section 3].

Write $\sigma(\overrightarrow{s}_i^j)$ (resp. $\sigma(\overrightarrow{s}_i)$, $\sigma(\eta_i^j)$, $\sigma(\eta^{-i})$) for the σ-algebra on $S^{-i} \times T^{-i}$ generated by $\overrightarrow{s}_i^j$ (resp. $\overrightarrow{s}_i$, η_i^j, η^{-i}). Similarly, let $\sigma(\overrightarrow{s}_i^j : j \neq i)$ (resp. $\sigma(\eta_i^j : j \neq i)$) be the σ-algebra on $S^{-i} \times T^{-i}$ generated by the random variables $\overrightarrow{s}_i^1, \ldots, \overrightarrow{s}_i^{i-1}, \overrightarrow{s}_i^{i+1}, \ldots, \overrightarrow{s}_i^n$ (resp. $\eta_i^1, \ldots, \eta_i^{i-1}, \eta_i^{i+1}, \ldots, \eta_i^n$). Note that $\sigma(\overrightarrow{s}_i^j : j \neq i) = \sigma(\overrightarrow{s}_i)$ and $\sigma(\eta_i^j : j \neq i) = \sigma(\eta^{-i})$. (See Dellacherie-Meyer [10, 1978, p.9].)

Fix a type $t^i \in T^i$, an event $E \in \mathcal{B}\left(S^{-i} \times T^{-i}\right)$ and a sub σ-algebra $\mathcal{S}$ of $\mathcal{B}\left(S^{-i} \times T^{-i}\right)$. Write $\lambda^i(t^i)(E||\mathcal{S}) : S^{-i} \times T^{-i} \to \mathbb{R}$ for a version of conditional probability of $\lambda^i(t^i)$ given E and $\mathcal{S}$.

Definition 5.1 *The random variables* $\overrightarrow{s}_i^1, \ldots, \overrightarrow{s}_i^{i-1}, \overrightarrow{s}_i^{i+1}, \ldots, \overrightarrow{s}_i^n$ *are* $\lambda^i(t^i)$***-conditionally independent given the random variable*** η^{-i} *if, for all* $j \neq i$ *and* $E^j \in \sigma(\overrightarrow{s}_i^j)$*,*

$$\lambda^i\left(t^i\right)\left(\bigcap_{j \neq i} E^j || \sigma(\eta^{-i})\right) = \prod_{j \neq i} \lambda^i\left(t^i\right)\left(E^j || \sigma(\eta^{-i})\right) \quad a.s.$$

Say the type t^i *satisfies* ***conditional independence (CI)*** *if* $\overrightarrow{s}_i^1, \ldots, \overrightarrow{s}_i^{i-1}, \overrightarrow{s}_i^{i+1}, \ldots, \overrightarrow{s}_i^n$ *are* $\lambda^i(t^i)$*-conditionally independent given* η^{-i}*.*

Definition 5.2 *The random variable* η_i^j *is* $\lambda^i(t^i)$***-sufficient for the random variable*** $\overrightarrow{s}_i^j$ *if, for each* $j \neq i$ *and* $E^j \in \sigma(\overrightarrow{s}_i^j)$*,*

$$\lambda^i\left(t^i\right)\left(E^j || \sigma(\eta^{-i})\right) = \lambda^i\left(t^i\right)\left(E^j || \sigma(\eta_i^j)\right) \quad a.s.$$

Say the type t^i *satisfies* ***sufficiency (SUFF)*** *if, for each* $j \neq i$*,* η_i^j *is* $\lambda^i(t^i)$*-sufficient for* $\overrightarrow{s}_i^j$*.*

In words, Definition 5.1 says that a type t^i satisfies CI if, conditional on knowing the hierarchies of beliefs of the other players $j \neq i$, type t^i's assessment of their strategies is independent. For SUFF, suppose type t^i knows player j's hierarchy of beliefs, and comes to learn the hierarchies of beliefs of players $k \neq i, j$. Type t^i satisfies SUFF if this new information doesn't change t^i's assessment of j's strategy.

Taken together, CI and SUFF capture the idea that the hidden variables in a game are the players' hierarchies of beliefs. In particular, the proposition below says that under CI and SUFF, if a type t^i of player i assesses other players' hierarchies as independent, then t^i assesses their strategies as independent. Equivalently, if t^i assesses other players' strategies as correlated, then t^i must assess their hierarchies as correlated.[7]

Proposition 5.1 *Suppose type* t^i *satisfies CI and SUFF. If* $\eta_i^1, \ldots, \eta_i^{i-1}, \eta_i^{i+1}, \ldots, \eta_i^n$ *are* $\lambda^i(t^i)$*-independent, then the random variables* $\overrightarrow{s}_i^1, \ldots, \overrightarrow{s}_i^{i-1}, \overrightarrow{s}_i^{i+1}, \ldots, \overrightarrow{s}_i^n$ *are* $\lambda^i(t^i)$*-independent.*

---

[7] See the appendices for proofs not given in the text.

6 RCBR and Iterated Dominance

Definition 6.1 *Say $\left(s^i, t^i\right) \in S^i \times T^i$ is **rational** if*

$$\sum_{s^{-i} \in S^{-i}} \pi^i\left(s^i, s^{-i}\right) \operatorname{marg}_{S^{-i}} \lambda^i\left(t^i\right)\left(s^{-i}\right) \geq \sum_{s^{-i} \in S^{-i}} \pi^i\left(r^i, s^{-i}\right) \operatorname{marg}_{S^{-i}} \lambda^i\left(t^i\right)\left(s^{-i}\right)$$

for every $r^i \in S^i$. Let R_1^i be the set of all rational pairs (s^i, t^i).

Definition 6.2 *Say $E \subseteq S^{-i} \times T^{-i}$ is **believed under** $\lambda^i\left(t^i\right)$ if E is Borel and $\lambda^i\left(t^i\right)(E) = 1$. Let*

$$B^i(E) = \{t^i \in T^i : E \text{ is believed under } \lambda^i(t^i)\}.$$

For $m > 1$, define R_m^i inductively by

$$R_{m+1}^i = R_m^i \cap [S^i \times B^i\left(R_m^{-i}\right)].$$

Definition 6.3 *If $(s^1, t^1, \ldots, s^n, t^n) \in R_{m+1}$, say there is **rationality and** m**th-order belief of rationality (R**m**BR)** at this state. If $(s^1, t^1, \ldots, s^n, t^n) \in \bigcap_{m=1}^{\infty} R_m$, say there is **rationality and common belief of rationality (RCBR)** at this state.*

We now turn to the definition of strong dominance:

Definition 6.4 *Fix $X^1, \ldots, X^n$, where each $X^i \subseteq S^i$. Say $s^i \in X^i$ is **(strongly) dominated with respect to** X if there exists $\sigma^i \in \mathcal{M}\left(S^i\right)$ with $\sigma^i(X^i) = 1$ and $\pi^i\left(\sigma^i, s^{-i}\right) > \pi^i\left(s^i, s^{-i}\right)$ for each $s^{-i} \in X^{-i}$. Otherwise, say s^i is **undominated with respect to** X.*

The following result is standard:

Lemma 6.1 *Fix $X^1, \ldots, X^n$, where each $X^i \subseteq S^i$. The strategy $s^i \in X^i$ is undominated with respect to X if and only if there exists $\mu \in \mathcal{M}\left(S^{-i}\right)$, with $\mu(X^{-i}) = 1$, such that $\pi^i\left(s^i, \mu\right) \geq \pi^i\left(r^i, \mu\right)$ for every $r^i \in X^i$.*

Define sets S_m^i inductively by $S_0^i = S^i$, and

$$S_{m+1}^i = \{s^i \in S_m^i : s^i \text{ is undominated with respect to } S_m\}.$$

Note there is an M such that $\cap_{m=0}^{\infty} S_m^i = S_M^i \neq \emptyset$, for all i. A strategy $s^i \in S_M^i$ (resp. strategy profile $s \in S_M$) is called **iteratively undominated** (**IU**).

We'll make use the following concept (Pearce [18, 1984]):

Definition 6.5 *Fix a game $G = \left\langle S^1, \ldots, S^n; \pi^1, \ldots, \pi^n\right\rangle$ and subsets $Q^i \subseteq S^i$, for $i = 1, \ldots, n$. The set $Q \subseteq S$ is a **best-response set (BRS)** if, for every i and each $s^i \in Q^i$, there is a $\mu(s^i) \in \mathcal{M}(S^{-i})$ with $\mu(s^i)(Q^{-i}) = 1$, such that $\pi^i\left(s^i, \mu(s^i)\right) \geq \pi^i\left(r^i, \mu(s^i)\right)$ for every $r^i \in S^i$.*

Standard facts about BRS's are (i) the set S_M of IU profiles is a BRS, and (ii) every BRS is contained in S_M.

Next is the baseline result that RCBR is characterized by the IU strategies:[8]

Proposition 6.1 *Consider a game* $G = \left\langle S^1, \ldots, S^n; \pi^1, \ldots, \pi^n \right\rangle$.

- *(i) Fix a structure* $\left\langle S^1, \ldots, S^n; T^1, \ldots, T^n; \lambda^1, \ldots, \lambda^n \right\rangle$*, and suppose there is RCBR at the state* $\left(s^1, t^1, \ldots, s^n, t^n\right)$. *Then the strategy profile* $\left(s^1, \ldots, s^n\right)$ *is IU in* G*.*
- *(ii) There is a structure* $\left\langle S^1, \ldots, S^n; T^1, \ldots, T^n; \lambda^1, \ldots, \lambda^n \right\rangle$ *such that, for each IU strategy profile* $\left(s^1, \ldots, s^n\right)$*, there is a state* $\left(s^1, t^1, \ldots, s^n, t^n\right)$ *at which there is RCBR.*

7 Main Result

Here we prove our main result:

Theorem 7.1

- *(i) There is a game* G *and an IU strategy* s^i *of* G*, such that the following holds: For any structure* Φ*, there does not exist a state at which each type satisfies CI, RCBR holds, and* s^i *is played.*
- *(ii) There is a game* G *and an IU strategy* s^i *of* G*, such that the following holds: For any* Φ*, there does not exist a state at which each type satisfies SUFF, RCBR holds, and* s^i *is played.*

By contrast with Proposition 6.1, the conditions of CI and RCBR (resp. SUFF and RCBR) are not characterized by the IU strategies.

	L	*C*	*R*
U	0,0,2	0,0,2	0,1,2
M	0,0,0	0,0,0	0,1,0
D	1,0,2	1,0,2	1,1,2

X

	L	*C*	*R*
U	1,1,1	0,1,0	0,1,0
M	1,0,0	0,0,1	0,1,0
D	1,0,0	1,0,0	1,1,0

Y

	L	*C*	*R*
U	0,0,0	0,0,0	0,1,0
M	0,0,2	0,0,2	0,1,2
D	1,0,2	1,0,2	1,1,2

Z

Figure 7.1

Lemma 7.1 *In the game* G*:*

- *(i) the strategy* U *(resp.* M*) is optimal under* $\mu \in \mathcal{M}\left(S^b \times S^c\right)$ *if and only if* $\mu\left(L, Y\right) = 1$*;*
- *(ii) the strategy* L *(resp.* C*) is optimal under* $\mu \in \mathcal{M}\left(S^a \times S^c\right)$ *if and only if* $\mu\left(U, Y\right) = 1$*;*

---

[8] Brandenburger-Dekel [8, 1987] and Tan-Werlang [21, 1988] show related results. Proposition 2.1 in [8, 1987] demonstrates an equivalence between common knowledge of rationality and IU. Theorem 5.1 in [21, 1988] shows that (in a universal structure) RmBR yields strategy profiles that survive $(m+1)$ rounds of iterated dominance. ([21, 1988] also states a converse (Theorem 5.3) and references the proof to the unpublished version of the paper.)

(iii) the strategy Y is optimal under $\mu \in \mathcal{M}\left(S^a \times S^b\right)$ if and only if $\mu(U,L) = \mu(M,C) = \frac{1}{2}$ (moreover, this measure is not independent).

Proof. Parts (i) and (ii) are immediate.

For part (iii), note that Y is optimal under $\mu \in \mathcal{M}\left(S^a \times S^b\right)$ if and only if

$$\mu(U,L) + \mu(M,C) \geq \max\{2(1-\mu(M)), 2(1-\mu(U))\},$$

where we write M for the set $\{M\} \times S^b$ and U for the set $\{U\} \times S^b$. Since $1 \geq \mu(U,L) + \mu(M,C)$, it follows that

$$1 \geq \max\{2(1-\mu(M)), 2(1-\mu(U))\},$$

or $\mu(M) \geq \frac{1}{2}$ and $\mu(U) \geq \frac{1}{2}$. Since M and U are disjoint, we get $\mu(M) = \frac{1}{2}$ and $\mu(U) = \frac{1}{2}$. From this, $\mu(U,L) + \mu(M,C) = 1$. But $\mu(U) \geq \mu(U,L)$ and $\mu(M) \geq \mu(M,C)$, and so $\mu(U,L) = \mu(M,C) = \frac{1}{2}$.

Finally, notice that μ is not independent, since $\frac{1}{2} = \mu(U,L) \neq \mu(U) \times \mu(L) = \frac{1}{2} \times \frac{1}{2}$. ■

Corollary 7.1 *The IU set in G is $\{U,M,D\} \times \{L,C,R\} \times \{X,Y,Z\}$.*

Proof. We've just established the optimality of U, M, L, C, and Y. The optimality of D, R, X, and Z is clear. So, using Lemma 6.1, it is readily verified that $S_1 = S$. Thus $S_m = S$ for all m, by induction. ■

In particular then, IU allows Charlie to play Y, and to have an expected payoff of 1. By Proposition 6.1(ii), there is a structure Φ and a state at which there is RCBR and Charlie plays Y. However, we'll see below (Corollary 7.5), this cannot happen when we add CI. Moreover, Charlie must then have an (expected) payoff of 2 not 1.

It will be convenient to introduce some notation. Fix a player i. For $j \neq i$, we write $\left[s^j\right]$ for the subset $\left\{s^j\right\} \times S^{-i-j} \times T^{-i}$ of $S^{-i} \times T^{-i}$. We also write $\left[t^j\right]$ for the subset $S^{-i} \times \left\{u^j \in T^j : \delta^j\left(u^j\right) = \delta^j\left(t^j\right)\right\} \times T^{-i-j}$ of $S^{-i} \times T^{-i}$. Note, $\left[t^j\right] = (\eta_i^j)^{-1}(\delta^j\left(t^j\right))$.

Corollary 7.2 *Fix a structure Φ for G, with $(Y, t^c) \in R_1^c$. Then*

$$\lambda^c(t^c)(E) = \lambda^c(t^c)([U] \cap [L] \cap E) + \lambda^c(t^c)([M] \cap [C] \cap E)$$

for any event E in $S^a \times T^a \times S^b \times T^b$. Moreover, if $\lambda^c(t^c)(E) = 1$, then

$$\lambda^c(t^c)([U] \cap [L] \cap E) = \lambda^c(t^c)([M] \cap [C] \cap E) = \tfrac{1}{2}.$$

Proof. The first part is immediate from $\lambda^c(t^c)([U] \cap [L]) = \lambda^c(t^c)([M] \cap [C]) = \frac{1}{2}$ (Lemma 7.1), and the fact that $[U] \cap [M] = \emptyset$. The second part follows from $[U] \cap [L] \cap E \subseteq [U] \cap [L]$ and $[M] \cap [C] \cap E \subseteq [M] \cap [C]$. ■

Lemma 7.2 *Fix a structure* Φ *for* G. *Suppose* $(Y, t^c) \in \bigcap_m R^c_m$ *where* t^c *satisfies CI. Then there are* $(t^a, t^b), (u^a, u^b) \in T^a \times T^b$, *with* $\left(U, t^a, L, t^b\right), \left(M, u^a, C, u^b\right) \in \bigcap_m (R^a_m \times R^b_m)$ *and either* $\delta^a(t^a) \neq \delta^a(u^a)$ *or* $\delta^b(t^b) \neq \delta^b(u^b)$ *(or both).*

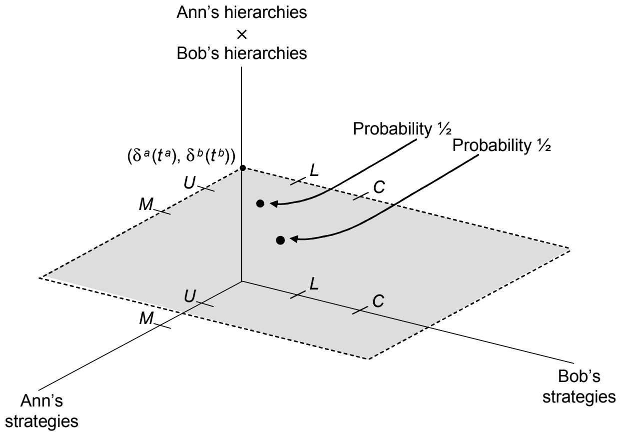


Figure 7.2

Here is the idea of the proof. We know that the given type of Charlie must assign probability $\frac{1}{2}$ to each of the events $[U] \cap [L]$ and $[M] \cap [C]$. Also, Charlie assigns probability 1 to the event "RCBR with respect to Ann and Bob" (formally, the event $\bigcap_m (R^a_m \times R^b_m)$). This means that if the lemma is false, we get a picture like Figure 7.2, where Charlie assigns probability $\frac{1}{2}$ to each of the two indicated points in some $(\delta^a(t^a), \delta^b(t^b))$-plane. But CI requires that Charlie's conditional measure, conditioned on any such horizontal plane, be a product measure, so we have a contradiction. The formal argument follows.

Proof of Lemma 7.2. Fix Φ with $(Y, t^c) \in \bigcap_m R^c_m$, where t^c satisfies CI. Then $\lambda^c(t^c)\left(R^a_m \times R^b_m\right) = 1$ for all m, so that $\lambda^c(t^c)\left(\bigcap_m (R^a_m \times R^b_m)\right) = 1$. Corollary 7.2 in then gives

$$\lambda^c(t^c)\left([U] \cap [L] \cap \textstyle\bigcap_m (R^a_m \times R^b_m)\right) = \lambda^c(t^c)\left([M] \cap [C] \cap \textstyle\bigcap_m (R^a_m \times R^b_m)\right) = \tfrac{1}{2}.$$

Suppose the lemma is false. Fix $\left(U, t^a, L, t^b\right) \in \bigcap_m (R^a_m \times R^b_m)$. Then, for any $\left(M, u^a, C, u^b\right) \in \bigcap_m (R^a_m \times R^b_m)$, $\delta^a(u^a) = \delta^a(t^a)$ and $\delta^b\left(u^b\right) = \delta^b\left(t^b\right)$. So we have

$$\begin{gathered}[U] \cap [L] \cap \textstyle\bigcap_m (R^a_m \times R^b_m) \subseteq [U] \cap [L] \cap [t^a] \cap \left[t^b\right], \\ [M] \cap [C] \cap \textstyle\bigcap_m (R^a_m \times R^b_m) \subseteq [M] \cap [C] \cap [t^a] \cap \left[t^b\right].\end{gathered}$$

But then $[U] \cap [M] = \emptyset$ implies $\lambda^c(t^c)([t^a] \cap \left[t^b\right]) = 1$. By Corollary 7.2 we then have

$$\begin{gathered}\lambda^c(t^c)\left([U] \cap [L] \cap [t^a] \cap \left[t^b\right]\right) = \tfrac{1}{2}, \\ \lambda^c(t^c)\left([M] \cap [C] \cap [t^a] \cap \left[t^b\right]\right) = \tfrac{1}{2}.\end{gathered}$$

From this,

$$\frac{1}{2} = \lambda^c(t^c)([U] \cap [L]) = \lambda^c(t^c)([U] \cap [L] \mid [t^a] \cap [t^b]) =$$
$$\lambda^c(t^c)([U] \mid [t^a] \cap [t^b]) \times \lambda^c(t^c)([L] \mid [t^a] \cap [t^b]) =$$
$$\lambda^c(t^c)([U]) \times \lambda^c(t^c)([L]),$$

where the second line uses CI. But Corollary 7.2 gives $\lambda^c(t^c)([U]) = \lambda^c(t^c)([L]) = \frac{1}{2}$, a contradiction. ■

Proposition 7.1 *Fix a game* $\langle S^1, \ldots, S^n; \pi^1, \ldots, \pi^n \rangle$ *and BRS* $\times_{i=1}^n Q^i \subseteq S$ *satisfying: For every* i *and each* $s^i \in Q^i$*, there is a unique* $\mu(s^i) \in \mathcal{M}(S^{-i})$ *under which* s^i *is optimal. Fix also a structure* Φ*. Then for every* i *and all* m *the following hold:*

(i) If $(s^{-i}, t^{-i}), (s^{-i}, u^{-i}) \in R_m^{-i} \cap (Q^{-i} \times T^{-i})$ *then* $\rho_m^i(s^{-i}, t^{-i}) = \rho_m^i(s^{-i}, u^{-i})$*.*

(ii) If $(s^i, t^i), (r^i, u^i) \in R_m^i \cap (Q^i \times T^i)$ *and* $\mu(s^i) = \mu(r^i)$ *then* $\delta_n^i(t^i) = \delta_n^i(u^i)$ *for all* $n \leq m$*.*

We first give the idea of the proof, and then the formal argument. Take strategies s^i and r^i of player i. Since they have the same unique measure under which they are optimal, and player i is rational, the associated types t^i and u^i must have the same marginal on S^{-i}. That is, the first-order beliefs $\delta_1^i(t^i) = \delta_1^i(u^i)$ must coincide. Next, since player i believes each other player j is rational, if i assigns positive probability to a strategy-type pair (s^j, t^j) of player j, it must be t^j has the unique marginal on S^{-j} that makes s^j optimal.[9] This is equally true if player i is type t^i or u^i. That is, the second-order beliefs $\delta_2^i(t^i) = \delta_2^i(u^i)$ must coincide. And so on (by induction). Essentially, the argument is that since strategies have unique 'support measures,' rationality pins down the first-order beliefs, rationality and belief in rationality pins down the second-order beliefs, etc.

Proof of Proposition 7.1. By induction on m.

Begin with $m = 1$: Part (i) is immediate from the fact that $\rho_1^i(s^{-i}, t^{-i}) = \rho_1^i(s^{-i}, u^{-i}) = s^{-i}$. For part (ii), fix $(s^i, t^i), (r^i, u^i) \in R_1^i \cap (Q^i \times T^i)$ with $\mu(s^i) = \mu(r^i)$. By definition, $\underline{\rho}_1^i(\lambda^i(t^i)) = \mathrm{marg}_{S^{-i}} \lambda^i(t^i)$ and $\underline{\rho}_1^i(\lambda^i(u^i)) = \mathrm{marg}_{S^{-i}} \lambda^i(u^i)$. Since $(s^i, t^i), (r^i, u^i) \in R_1^i$,

$$\mathrm{marg}_{S^{-i}} \lambda^i(t^i) = \mu(s^i) = \mu(r^i) = \mathrm{marg}_{S^{-i}} \lambda^i(u^i).$$

Now assume the lemma is true for m. Begin with part (i). Suppose $(s^{-i}, t^{-i}), (s^{-i}, u^{-i}) \in R_{m+1}^{-i} \cap (Q^{-i} \times T^{-i})$. The induction hypothesis applied to part (i) gives $\rho_m^i(s^{-i}, t^{-i}) = \rho_m^i(s^{-i}, u^{-i})$. Also, the induction hypothesis applied to part (ii) gives $\delta_m^j(t^j) = \delta_m^j(u^j)$ for each $j \neq i$. With this, $\rho_{m+1}^i(s^{-i}, t^{-i}) = \rho_{m+1}^i(s^{-i}, u^{-i})$, establishing part (i) for $(m + 1)$.

---

[9] The formal proof applies to general (infinite) type spaces.

Turn to part (ii). Suppose $\left(s^i, t^i\right), \left(r^i, u^i\right) \in R^i_{m+1} \cap \left(Q^i \times T^i\right)$ and $\mu\left(s^i\right) = \mu\left(r^i\right)$. Then $\left(s^i, t^i\right), \left(r^i, u^i\right) \in R^i_m \cap \left(Q^i \times T^i\right)$, and so the induction hypothesis applied to part (ii) gives $\delta^i_n\left(t^i\right) = \delta^i_n\left(u^i\right)$ for all $n \leq m$. As such, it suffices to show $\delta^i_{m+1}\left(t^i\right) = \delta^i_{m+1}\left(u^i\right)$.

Fix an event E in Y^i_{m+1}, and a point $\left(s^{-i}, t^{-i}\right) \in \operatorname{Supp} \lambda^i\left(t^i\right) \cap \left(\rho^i_{m+1}\right)^{-1}(E)$. Then, for each $\left(s^{-i}, u^{-i}\right) \in \operatorname{Supp} \lambda^i\left(t^i\right) \cup \operatorname{Supp} \lambda^i\left(u^i\right)$, it must be that $\left(s^{-i}, u^{-i}\right) \in \left(\rho^i_{m+1}\right)^{-1}(E)$. To see this, first notice that, by Corollary C1 in Appendix C, $\operatorname{Supp} \lambda^i\left(t^i\right) \cup \operatorname{Supp} \lambda^i\left(u^i\right) \subseteq R^{-i}_{m+1}$. Also note that since $\left(s^i, t^i\right), \left(s^i, u^i\right) \in R^i_1$,

$$\operatorname{marg}_{S^{-i}} \lambda^i\left(t^i\right) = \mu\left(s^i\right) = \operatorname{marg}_{S^{-i}} \lambda^i\left(u^i\right).$$

Since $\mu\left(s^i\right)\left(Q^{-i}\right) = 1$, it follows that $\operatorname{Supp} \lambda^i\left(t^i\right) \cup \operatorname{Supp} \lambda^i\left(u^i\right) \subseteq Q^{-i} \times T^{-i}$. So $\left(s^{-i}, t^{-i}\right), \left(s^{-i}, u^{-i}\right) \in R^{-i}_{m+1} \cap \left(Q^{-i} \times T^{-i}\right)$. With this, part (i) (for $(m+1)$) of the lemma gives $\rho^i_{m+1}\left(s^{-i}, t^{-i}\right) = \rho^i_{m+1}\left(s^{-i}, u^{-i}\right)$. Thus $\left(s^{-i}, u^{-i}\right) \in \left(\rho^i_{m+1}\right)^{-1}(E)$, as required.

Using this, we can now write

$$\begin{aligned}
\lambda^i\left(t^i\right)\left(\left(\rho^i_{m+1}\right)^{-1}(E)\right) &= \lambda^i\left(t^i\right)\left(\left(\rho^i_{m+1}\right)^{-1}(E) \cap \operatorname{Supp} \lambda^i\left(t^i\right)\right) = \\
&\sum_{s^{-i} \in \operatorname{proj}_{S^{-i}}\left(\rho^i_{m+1}\right)^{-1}(E)} \lambda^i\left(t^i\right)\left(\left\{s^{-i}\right\} \times \left\{t^{-i} : \left(s^{-i}, t^{-i}\right) \in \left(\rho^i_{m+1}\right)^{-1}(E) \cap \operatorname{Supp} \lambda^i\left(t^i\right)\right\}\right) = \\
&\sum_{s^{-i} \in \operatorname{proj}_{S^{-i}}\left(\rho^i_{m+1}\right)^{-1}(E)} \lambda^i\left(t^i\right)\left(\left\{s^{-i}\right\} \times \left\{t^{-i} : \left(s^{-i}, t^{-i}\right) \in \operatorname{Supp} \lambda^i\left(t^i\right)\right\}\right) = \\
&\sum_{s^{-i} \in \operatorname{proj}_{S^{-i}}\left(\rho^i_{m+1}\right)^{-1}(E)} \operatorname{marg}_{S^{-i}} \lambda^i\left(t^i\right)\left(s^{-i}\right).
\end{aligned}$$

A corresponding argument shows that

$$\lambda^i\left(u^i\right)\left(\left(\rho^i_{m+1}\right)^{-1}(E)\right) = \sum_{s^{-i} \in \operatorname{proj}_{S^{-i}}\left(\rho^i_{m+1}\right)^{-1}(E)} \operatorname{marg}_{S^{-i}} \lambda^i\left(u^i\right)\left(s^{-i}\right).$$

Now note

$$\operatorname{marg}_{S^{-i}} \lambda^i\left(t^i\right) = \mu\left(s^i\right) = \mu\left(r^i\right) = \operatorname{marg}_{S^{-i}} \lambda^i\left(u^i\right),$$

establishing $\delta^i_{m+1}\left(t^i\right) = \delta^i_{m+1}\left(u^i\right)$, as required. ■

Corollary 7.3 *Fix a game* $\left\langle S^1, \ldots, S^n; \pi^1, \ldots, \pi^n\right\rangle$ *and BRS* $\times^n_{i=1} Q^i \subseteq S$ *satisfying: For every* i *and each* $s^i \in Q^i$*, there is a unique* $\mu(s^i) \in \mathcal{M}(S^{-i})$ *under which* s^i *is optimal. Fix also a structure* Φ*. If* $(s^i, t^i), (r^i, u^i) \in \bigcap_m R^i_m \cap (Q^i \times T^i)$ *and* $\mu(s^i) = \mu(r^i)$*, then* $\delta^i(t^i) = \delta^i(u^i)$*.*

Proof. Suppose instead that $\delta^i\left(t^i\right) \neq \delta^i\left(u^i\right)$. Then there exists m such that $\delta^i_m\left(t^i\right) \neq \delta^i_m\left(u^i\right)$. Since $\left(s^i, t^i\right), \left(r^i, u^i\right) \in R^i_m$ this contradicts Proposition 7.1. ■

Corollary 7.4 *Let $Q^a = \{U, M\}$, $Q^b = \{L, C\}$, $Q^c = \{Y\}$ in the game G. Fix a structure Φ. For each i, if $(s^i, t^i), (r^i, u^i) \in \bigcap_m R^i_m \cap (Q^i \times T^i)$, then $\delta^i(t^i) = \delta^i(u^i)$.*

Proof. Immediate from Corollaries 7.1 and 7.3. ■

Proof of Theorem 7.1(i). Fix a structure Φ for G. Corollary 7.4 implies that if $(U, t^a), (M, u^a) \in \bigcap_m R^a_m$, then $\delta^a(t^a) = \delta^a(u^a)$. Likewise, if $\left(L, t^b\right), \left(C, u^b\right) \in \bigcap_m R^b_m$, then $\delta^b(t^b) = \delta^b(u^b)$. With this, Lemma 7.2 implies that if $(Y, t^c) \in \bigcap_m R^c_m$, then t^c does not satisfy CI. But Y is an IU strategy, by Lemma 7.1. Setting $s^i = Y$ establishes the theorem. ■

Corollary 7.5 *Fix a structure Φ for G, and a state $(s^a, t^a, s^b, t^b, s^c, t^c)$ at which there is RCBR. If t^c is CI, then $s^a = D$, $s^b = R$, and $s^c = X$ or Z.*

To prove part (ii) of Theorem 7.1, we use the game G' in Figure 7.3. (This is the same as the game G of Figure 7.1, except for swapping player 2's payoffs in (U, C, Y) and (M, C, Y).)

	L	C	R
U	0,0,2	0,0,2	0,1,2
M	0,0,0	0,0,0	0,1,0
D	1,0,2	1,0,2	1,1,2

X

	L	C	R
U	1,1,1	0,0,0	0,1,0
M	1,0,0	0,1,1	0,1,0
D	1,0,0	1,0,0	1,1,0

Y

	L	C	R
U	0,0,0	0,0,0	0,1,0
M	0,0,2	0,0,2	0,1,2
D	1,0,2	1,0,2	1,1,2

Z

Figure 7.3

The following is immediate from Corollary 7.3 above.

Corollary 7.6 *Let $Q^a = \{U, M\}$, $Q^b = \{L, C\}$, $Q^c = \{Y\}$ in the game G'. For $i = 1, 3$, if $(s^i, t^i), (r^i, u^i) \in \bigcap_m R^i_m \cap (Q^i \times T^i)$, then $\delta^i(t^i) = \delta^i(u^i)$. For $i = 2$, if $(s^i, t^i), (r^i, u^i) \in \bigcap_m R^i_m \cap (Q^i \times T^i)$, then $\delta^i(t^i) = \delta^i(u^i)$ only if $s^i = r^i$.*

Proof. Same as for Corollary 7.4, except that while L is optimal only under the measure that assigns probability one to (U, Y), now C is optimal only under the measure that assigns probability one to (M, Y). So if $\left(L, t^b\right), \left(C, u^b\right) \in \bigcap_m R^b_m$, then $t^b \neq u^b$. ■

Proof of Theorem 7.1(ii). Fix a structure Φ for G'. By Corollary 7.6, there are t^a, t^b, u^b, with $\delta^b(t^b) \neq \delta^b(u^b)$, such that

$$[U] \cap [L] \cap \textstyle\bigcap_m (R^a_m \times R^b_m) \subseteq [U] \cap [L] \cap [t^a] \cap \left[t^b\right],$$
$$[M] \cap [C] \cap \textstyle\bigcap_m (R^a_m \times R^b_m) \subseteq [M] \cap [C] \cap [t^a] \cap \left[u^b\right].$$

Paralleling the argument in the proof of Lemma 7.2, we then have

$$\lambda^c(t^c)\left([U] \cap [L] \cap [t^a] \cap \left[t^b\right]\right) = \tfrac{1}{2},$$
$$\lambda^c(t^c)\left([M] \cap [C] \cap [t^a] \cap \left[u^b\right]\right) = \tfrac{1}{2},$$

since $[U] \cap [M] = \emptyset$. Paralleling Corollary 7.2, we get that for any event E

$$\lambda^c(t^c)(E) = \lambda^c(t^c)([U] \cap [L] \cap [t^a] \cap [t^b] \cap E) + \lambda^c(t^c)([M] \cap [C] \cap [t^a] \cap [u^b] \cap E).$$

Setting $E = [t^a]$, $[U] \cap [t^a] \cap [t^b]$, and $[t^a] \cap [t^b]$, yields respectively

$$\begin{aligned} \lambda^c(t^c)([t^a]) &= 1, \\ \lambda^c(t^c)([U] \cap [t^a] \cap [t^b]) &= \frac{1}{2}, \\ \lambda^c(t^c)([t^a] \cap [t^b]) &= \frac{1}{2}. \end{aligned}$$

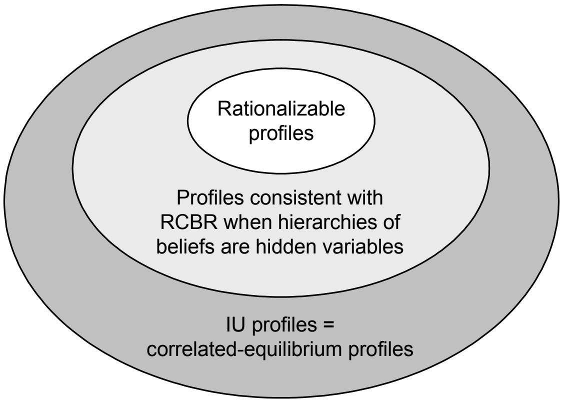
But then

$$\frac{1}{2} = \lambda^c(t^c)([U]) = \lambda^c(t^c)([U] \mid [t^a]) \neq$$
$$\lambda^c(t^c)([U] \mid [t^a] \cap [t^b]) = \frac{\lambda^c(t^c)([U] \cap [t^a] \cap [t^b])}{\lambda^c(t^c)([t^a] \cap [t^b])} = \frac{\frac{1}{2}}{\frac{1}{2}} = 1,$$

so that t^c doesn't satisfy SUFF. ■

8 General Relationships

Fix a game G. We have the general relationships shown in Figure 8.1:



The strategies in the middle set–viz., the strategies consistent with RCBR under our hidden-variable analysis–are the strategies consistent with RCBR, CI, and SUFF. We now relate this middle set to the inner and outer sets.

Begin with the inner set: Any rationalizable strategy is consistent with RCBR when the hidden variables are hierarchies of beliefs:

Proposition 8.1 *Fix a game* $G = \langle S^1, \ldots, S^n; \pi^1, \ldots, \pi^n \rangle$. *There is an associated structure* $\langle S^1, \ldots, S^n; T^1, \ldots, T^n; \lambda^1, \ldots, \lambda^n \rangle$ *such that, for each rationalizable strategy profile* $(s^1, \ldots, s^n)$, *there is a state* $(s^1, t^1, \ldots, s^n, t^n)$ *at which RCBR holds, and each type satisfies CI and SUFF.*

This result should be distinguished from the following: Fix a game G and associated structure Φ. Suppose that for each player i and type t^i, the marginal on S^{-i} of the measure $\lambda^i(t^i)$ is independent. Then: (i) if there is RCBR at the state $(s^1, t^1, \ldots, s^n, t^n)$, the strategy profile $(s^1, \ldots, s^n)$ is rationalizable in G; and (ii) the types in $(t^1, \ldots, t^n)$ satisfy CI. Certainly (i) is true (Proposition E1). But (ii) may be false, as Example E1 shows. The reason is that, within a given structure Φ, independence need not imply conditional independence (as is standard in probability theory).

In Figure 8.1, the inner set may be strictly contained in the middle set. The game of Figure 3.1 and the associated structure in Figure 2.1 was an example of this. (The 'problem' was the reverse of the above–conditional independence held but not independence.)

Now the relationship between the middle and outer sets in Figure 8.1. Proposition 6.1(i) gives the general relationship: Fix a game and a state of an associated structure. If RCBR holds, the strategies played are IU–so, certainly, this is the case if each type satisfies CI and SUFF. The games in Section 7 showed that the inclusion may be strict.

The same result gives the relationship between the middle set and (subjective) correlated equilibrium (Aumann [1, 1974], [2, 1987]). Brandenburger-Dekel [8, 1987] showed that the subjective correlated-equilibrium strategies are the same as the IU strategies.[10] So, our main result can equally be stated as: The strategies consistent with RCBR, CI, and SUFF may be strictly contained in the set of (subjective) correlated-equilibrium strategies.

Of course, for any given game, the inclusions in Figure 8.1 need not be strict. In the game of Figure 7.1, the rationalizable strategies coincide with the strategies consistent with RCBR, CI, and SUFF. In the game of Figure 3.1, the strategies consistent with RCBR, CI, and SUFF coincide with the IU strategies.

A natural question is whether we can identify a class of games where the middle and outer sets in Figure 8.1 coincide. Here is one such class. Fix a game $G = \langle S^1, \ldots, S^n; \pi^1, \ldots, \pi^n \rangle$. Recall that the IU set S_M of G is a BRS: For every i and each $s^i \in S^i_M$, there is a $\mu(s^i) \in \mathcal{M}(S^{-i})$ with $\mu(s^i)(S^{-i}_M) = 1$, under which s^i is optimal. Say a game satisfies the **injectivity condition** if the measures $\mu(s^i)$ can be chosen so that $\mu(r^i) \neq \mu(s^i)$ if $r^i \neq s^i$, for $r^i, s^i \in S^i_M$.

Proposition F2 shows that if a game satisfies this condition, then the middle and outer sets in Figure 8.1 coincide. Proposition F3 notes that the injectivity condition is generic in the matrix. So, generically in the matrix, all IU strategies can arise under our RCBR, CI, and SUFF conditions.

This said, genericity in the matrix is usually viewed as too strong a condition: it is well understood that many games of applied interest are non-generic (even in the tree). (See the discussions in Mertens [13, 1989, pp.582] and Marx-Swinkels [12, 1997, pp.224-225].) For this reason, we believe it

---

[10] More precisely, the equivalence involves "a posteriori equilibrium" ([1, 1974, Section 8])–a definition of subjective correlated equilibrium that ensures null events are treated properly.

is more illuminating to identify structural conditions on a game under which a particular statement–such as equality of the middle and outer sets in Figure 8.1–holds. Injectivity is one such condition. No doubt, there are other conditions of interest.

9 Discussion

We now give an explanation of the relationships in Figure 8.1. We start with the relationship between the rationalizable profiles (the inner set) and the profiles consistent with RCBR when the hierarchies are the hidden variables (the middle set).

9.1 Relationship to Rationalizability

The key difference between the two sets is the view taken on correlation vs. independence across players' hierarchies of beliefs.

One natural source of correlation is signals–so-called correlating devices, sunspots, etc.–the players observe outside the game. Another is 'intrinsic' correlations across the players' characteristics. The rationalizable profiles arise when we take the Classical view on these possibilities. Kohlberg-Mertens [11, 1986, p.1005] describe this:

> We adhere to the classical point of view that the game under consideration fully describes the real situation.... No random event (not described in the extensive form) can be observed by a player, except if it is completely independent of any random event observed by any other player and of the moves of nature in the tree.... Even before the start of the game such observations have to be forbidden (except random variables that are common knowledge to all players, if also the analysis is done conditionally to those random variables).

Mertens [14, 2003, p.389] writes:

> [S]taying clear as much as possible from any aspect of correlation, one should have a scenario like the following in mind: a player is picked at random, and put in a cubicle with a computer terminal as only link to the outside world. He first gets the following briefing from the terminal: "This is an experiment. Other players have been picked at random somewhere else in the world, and put in cubicles like yours, with all terminals connected to the same computer. You all receive this same briefing, next you will have to play the following game."

So, the Classical view has either no external signals or at most independent signals. Also, players are selected independently, presumably ruling out intrinsic correlations. In this scenario, we should assume that each player assesses the other players' hierarchies of beliefs as independent.[11] Assuming CI and SUFF, Proposition 5.1 implies that each player then assesses the other players' strategies

---

[11] Of course, Kohlberg-Mertens [11, 1986] and Mertens [14, 2004] are concerned with equilibrium analysis, but the Classical view can also be understood within the epistemic approach.

as independent. Under RCBR, the strategies played will therefore be rationalizable (proved as Proposition E1).[12]

But one can certainly argue that the Classical approach is unrealistic. There are external signals, quite possibly correlated, because there is always a history before the start of the game as described. Equally, intrinsic correlations are natural, because who the players are may be a ‘shorthand’ for their experiences (i.e., signals) prior to the game. This idea is in line with the Savage’s Small-Worlds idea in decision theory [19, 1954, pp.82-91]. Under it, correlation across the players’ hierarchies of beliefs is the norm. This leads to the strategy profiles we’ve identified in this paper–i.e., the middle set in Figure 8.1.

Note, under our approach, signals affect the play of the game insofar as they affect the structure of the game–via the players’ hierarchies of beliefs.[13] They do not affect the play of the game directly. In the next subsection we’ll argue that this latter kind of situation must involve physical correlation across players.

9.2 Relationship to IU

Next is the relationship between the profiles consistent with RCBR when the hierarchies are the hidden variables (the middle set in Figure 8.1), and the IU profiles (the outer set).

Go back to the game of Figure 7.1, and consider the associated structure in Figure 9.1.

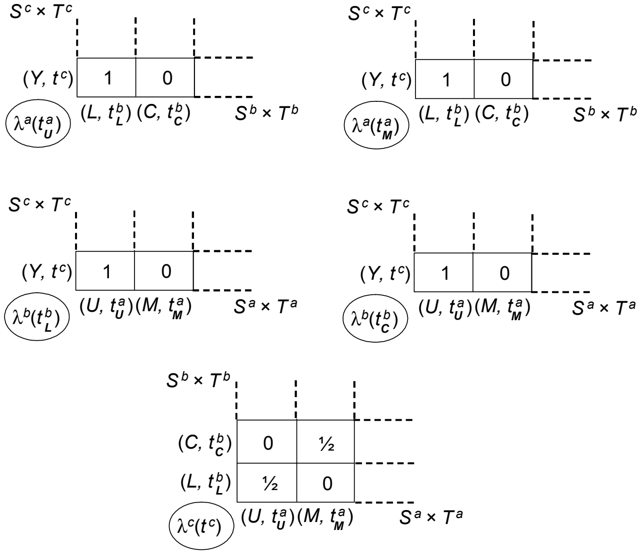


Figure 9.1

---

[12] Proposition 3.1 in Brandenburger-Dekel [8, 1987] shows an equivalence, given an independence condition on types, between common knowledge of rationality and rationalizability.

[13] Signals could affect other aspects of structure, such as payoff functions. So, if there is uncertainty over the payoff functions, signals could matter this way, too. See Appendix H.

There are two types of Ann, labelled t_U^a and t_M^a. Intuitively, type t_U^a (resp. t_M^a) is to be thought of as 'associated with' strategy U (resp. M). Likewise, there are two types of Bob, labelled t_L^b and t_C^b (one for each of his strategies L and C). Each type assigns positive probability only to rational strategy-type pairs of the other players. So, RCBR holds at the state $(U, t_U^a, L, t_L^b, Y, t^c)$, for example.

But CI doesn't hold. Types t_U^a and t_M^a (resp. t_L^b and t_C^b) induce the same hierarchy of beliefs. When we look at the measure $\lambda^c(t^c)$ conditioned on Ann's and Bob's (unique) hierarchies, we don't get a product measure. This is why Charlie can now play Y, unlike in Section 7.

What if we redefine CI to require independence conditional on Ann's and Bob's types rather than their hierarchies? If we similarly redefine SUFF? Now type t^c will satisfy (modified) CI and SUFF. This situation is general: Under RCBR, and CI and SUFF defined with respect to types not hierarchies, we can get all the IU profiles.

This is proved in Appendix F. We give the idea of the construction. For each player i, take the type set T^i to be a copy of the set of player i's IU strategies, viz. S_M^i. Use the fact that the IU set $\times_{i=1}^n S_M^i$ is a BRS: for each $s^i \in S_M^i$, take a measure $\mu\left(s^i\right)$, with $\mu\left(s^i\right)\left(S_M^{-i}\right) = 1$, under which s^i is optimal. The measure $\lambda^i\left(t^i\right)$ associated with a type $t^i = s^i$ is then as shown in Figure 9.2:

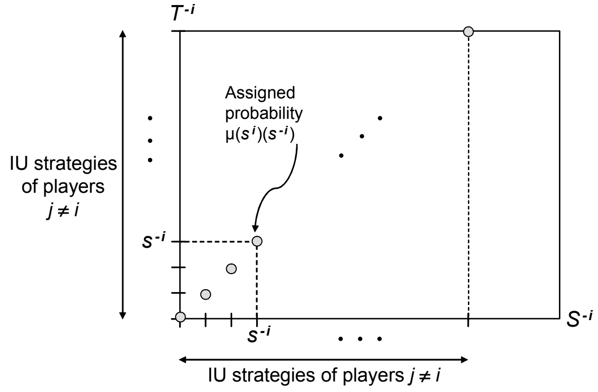


Figure 9.2

Note that the marginal on S^{-i} is exactly $\mu(s^i)$, so that the pair $(s^i, t^i) = (s^i, s^i)$ is rational. Since we do the same construction for every player $j \neq i$, the points along the diagonal are all rational, so t^i believes the other players are rational. Again, this holds for all players, so by induction, RCBR will hold. Moreover, notice that the conditional of $\lambda^i\left(t^i\right)$, conditioned on a profile $t^{-i} = s^{-i}$ of types of the players $j \neq i$, is concentrated on the single strategy profile s^{-i}. This tells us that (modified) CI will be satisfied. Similarly for (modified) SUFF.[14]

Summing up, the key idea is that for each player i, every strategy s^i is 'associated with' a different type t^i $(= s^i)$. Modified CI and SUFF follow directly from this.

---

[14] In the structure of Figure 9.1, we excluded the IU strategies D, R, X, Z from our construction of type sets. In general, the construction may need to include all IU strategies.

What is the meaning of this result? In the construction, the sets of hidden variables T^i aren't the players' hierarchies of beliefs–but rather their IU strategies, since $T^i = S^i_M$. Moreover, the construction is tight. To see this, return to Figure 9.1, and check that if, say, there was only one type for Ann–vs. the two types t^a_U and t^a_M–modified SUFF would fail. (If there was also only one type for Bob, modified CI would fail too.) But under RCBR, types t^a_U and t^a_M must induce the same hierarchy of beliefs. So, they are intrinsically the same, differing only in the strategies with which they are labelled.

We conclude that to get the IU strategies in a game, the strategies themselves may need to be the hidden variables.

What does this mean? Player i has a correlated assessment over the strategies of players $j \neq i$. He tries to explain this correlation via hidden variables. But the only way to do so, we have seen, may be by taking the hidden variables to be the variables in which we're trying to explain correlation in the first place. To us, this means there is physical correlation across the players. More precisely, player i must conclude that players $j \neq i$ are physically correlating–or coordinating–their strategies.

Refer back to Figure 8.1: We conclude that the difference between the set of strategies consistent with RCBR when the hidden variables are hierarchies of beliefs (the middle set), and the IU profiles (the outer set), is that the outer set may involve physical correlation in a game–which the middle set rules out.

As in Section 9.1, players can observe signals outside the game. This time though, signals (also) directly affect the play of the game. In Figure 9.1, Charlie could think that Ann will have type t^a_U and Bob will have type t^b_L, if they jointly observe one signal, while Ann will have type t^a_M and Bob will have type t^b_C, if they jointly observe another signal. But notice that this time, signals don't affect the structure of the game (the players' hierarchies of beliefs)–they directly affect the play of the game. With or without signals, we see this situation as one of physical correlation.

In Appendix H we make some additional comments on the relationship between the profiles consistent with RCBR, CI, and SUFF, and the IU profiles.

10 Conclusion

If we want to look at the entire set of IU strategies in a game, we may have to allow physical correlation across players. Under a strict interpretation of non-cooperative theory, this shouldn't be allowed. Physical correlation is incompatible with the non-cooperative set-up, in which players aren't allowed to coordinate their strategy choices.

Again, there is an analogy to coin tossing. Imagine a coin is tossed twice. Not all probability assessments reflect an observer's ignorance of the coin's parameter. For example, the assessment $p(H,T) = p(T,H) = \frac{1}{2}$ cannot arise this way.[15] Letting μ be the prior on the coin's parameter,

---

[15] This example is from Diaconis [9, 1977].

this would require

$$\int_0^1 p^2 \, \mathrm{d}\mu\left(p\right) = \int_0^1 \left(1 - p\right)^2 \, \mathrm{d}\mu\left(p\right) = 0,$$

which is impossible. Clearly, we could get this assessment, even with conditional independence, if we are allowed to condition on the outcome of the first toss. But then the outcome of the first toss becomes a characteristic of the coin, which wouldn't normally be allowed. (It says the coin remembers!)

Back to game theory, even if we take a looser interpretation of non-cooperative theory and allow coordination across players, there is still a puzzle with IU. In the game of Figure 7.1, Charlie now plays Y because he thinks Ann and Bob are coordinating their strategy choices. But notice that Charlie doesn't coordinate with Ann or Bob (or both) himself. This is asymmetric. We analyze each player as a separate decision maker, but allow each player to think of the other players as coordinating.

There is a clearly consistent way to proceed. As in this paper, we do non-cooperative game theory without physical correlation, by specifying the players' hierarchies of beliefs and doing the analysis relative to them. This leaves an open question. Can we give a (hierarchy-free) characterization of the strategies that can be played under our conditions of RCBR, CI, and SUFF? We don't know the answer.

Appendix A Proofs for Section 4

Lemma A1 *Fix Polish spaces A, B and a continuous map $f : A \to B$. Let $g : \mathcal{M}(A) \to \mathcal{M}(B)$ be given by $g(\mu) = \mu \circ f^{-1}$ for each $\mu \in \mathcal{M}(A)$. Then g is continuous.*

Proof. We need to show that the inverse image of every closed set in $\mathcal{M}(B)$ is closed in $\mathcal{M}(A)$. Let E be a closed set in $\mathcal{M}(B)$, then we want that: Fix a sequence of measures μ_n in $g^{-1}(E)$, where μ_n converges weakly to μ (in $\mathcal{M}(A)$). Then $\mu \in g^{-1}(E)$. To show this, it suffices to show that $g(\mu_n)$ converges weakly to $g(\mu)$ (in $\mathcal{M}(B)$). If so, $g(\mu) \in E$ and so $\mu \in g^{-1}(E)$.

So: Fix an open set U in B. Then $f^{-1}(U)$ is open in A. By the Portmanteau Theorem, $\liminf \mu_n\left(f^{-1}(U)\right) \geq \mu\left(f^{-1}(U)\right)$. But this says that $\liminf g(\mu_n)(U) \geq g(\mu)(U)$, and so, by the Portmanteau Theorem again, $g(\mu_n)$ converges weakly to $g(\mu)$. ■

Proposition A1 *The maps $\rho^i_m : S^{-i} \times T^{-i} \to Y^i_m$ and $\delta^i_m : T^i \to \mathcal{M}\left(Y^i_m\right)$ are continuous.*

Proof. First note that $\rho^i_1 = \mathrm{proj}_{S^{-i}}$, and so is certainly continuous. So $\underline{\rho}^i_1$ is continuous, by Lemma A1, and thus δ^i_1 is continuous.

Assume ρ^i_m and δ^i_m are continuous, for all i. Fix a rectangular open set $U \times \times_{j \neq i} V^j \subseteq Y^i_{m+1} = Y^i_m \times \times_{j \neq i} \mathcal{M}(Y^j_m)$. Notice

$$\left(\rho^i_{m+1}\right)^{-1}\left(U \times \times_{j \neq i} V^j\right) = \left(\rho^i_m\right)^{-1}(U) \cap \textstyle\bigcap_{j \neq i}[S^{-i} \times (\delta^j_m)^{-1}(V^j)].$$

Thus $\left(\rho^i_{m+1}\right)^{-1}\left(U \times \times_{j \neq i} V^j\right)$ is open since each set on the right-hand side is open. Since the rectangular sets form a basis, this shows that ρ^i_{m+1} is continuous. Again, by Lemma A1, for each i, $\underline{\rho}^i_{m+1}$ is then continuous, and so each δ^i_{m+1} is continuous. ■

Since each δ^i_m is continuous, it follows that the map δ^i is continuous. (See, e.g., Munkres [17, 1975, Theorem 8.5].)

Appendix B Proofs for Section 5

Let $X^1, \ldots, X^m, Y^1, \ldots, Y^m$ be Polish spaces and set

$$\Omega = X^1 \times \cdots \times X^m \times Y^1 \times \cdots \times Y^m.$$

For each j, define Borel measurable maps $f^j : \Omega \to X^j$ and $g^j : \Omega \to Y^j$ by $f^j = \mathrm{proj}_{X^j}$ and $g^j = \mathrm{proj}_{Y^j}$. Also let $g : \Omega \to \prod_j Y^j$ be given by $g = \mathrm{proj}_Y$. Fix a Borel probability measure μ on Ω, an event E in $\mathcal{B}(\Omega)$, and a sub σ-algebra $\mathcal{S}$ of $\mathcal{B}(\Omega)$. Write $\mu(E \| \mathcal{S}) : \Omega \to \mathbb{R}$ for (a version of) the conditional probability of E given $\mathcal{S}$.

Say the random variables $f^1, \ldots, f^m$ are **μ-conditionally independent given** g if, for all $E^j \in \sigma(f^j)$, $j = 1, \ldots, m$,

$$\mu(\textstyle\bigcap_j E^j \| \sigma(g)) = \prod_j \mu(E^j \| \sigma(g)) \quad \text{a.s.} \tag{B1}$$

Say the random variable g^j is **μ-sufficient for** f^j if, for all $E^j \in \sigma(f^j)$

$$\mu(E^j \| \sigma(g)) = \mu(E^j \| \sigma(g^j)) \quad \text{a.s.} \tag{B2}$$

Proposition B1 *Suppose the random variables $f^1, \ldots, f^m$ are μ-conditionally independent given g, and that, for each $j = 1, \ldots, m$, the random variable g^j is μ-sufficient for f^j. If the random variables $g^1, \ldots, g^m$ are μ-independent, then $f^1, \ldots, f^m$ are μ-independent.*

For the proof, we will use:

Lemma B1 *Fix some $E \in \mathcal{B}(\Omega)$ and some $\omega, \tilde{\omega} \in \Omega$ with $\omega = (x^1, \ldots, x^m, y^1, \ldots, y^m)$, $\tilde{\omega} = (\tilde{x}^1, \ldots, \tilde{x}^m, \tilde{y}^1, \ldots, \tilde{y}^m)$. If $y^j = \tilde{y}^j$ then*

$$\mu(E \| \sigma(g^j))_\omega = \mu(E \| \sigma(g^j))_{\tilde{\omega}}.$$

Proof. Suppose not, i.e. $\mu(E \| \sigma(g^j))$ is not constant on the set $\prod_k X^k \times \{y^j\} \times \prod_{k \neq j} Y^k$. Since $\mu(E^j \| \sigma(g^j))$ is measurable with respect to $\sigma(g^j)$, there must be nonempty disjoint subsets of $\prod_k X^k \times \{y^j\} \times \prod_{k \neq j} Y^k$ that are contained in $\sigma\left(g^j\right)$. This contradicts the definition of $\sigma\left(g^j\right)$. ■

Proof of Proposition B1. Assume the random variables $g^1, \ldots, g^m$ are μ-independent. Then $\operatorname{marg}_{Y^1 \times \cdots \times Y^m} \mu$ is a product measure. (See Billingsley [7, 1995, p.262].) We will use this fact to show that the random variables $f^1, \ldots, f^m$ are μ-independent.

For each $j = 1, \ldots, m$, fix $E^j \in \sigma\left(f^j\right)$. Note, $\bigcap_j E^j \in \mathcal{B}(\Omega)$. Then, by definition of conditional probability

$$\mu(\textstyle\bigcap_j E^j) = \int_\Omega \mu(\bigcap_j E^j \| \sigma(g))_\omega \, \mathrm{d}\mu(\omega).$$

Using conditional independence (equation B1) and sufficiency (equation B2),

$$\mu(\textstyle\bigcap_j E^j) = \int_\Omega \prod_j \mu(E^j \| \sigma(g^j))_\omega \, \mathrm{d}\mu(\omega). \tag{B3}$$

For each $j = 1, \ldots, m$, define maps $h^j : Y^j \to \mathbb{R}$ so that $h^j \circ g^j = \mu(E^j \| \sigma(g^j))$. Note, by Lemma B1, this is well defined. Also define $h : \prod_j Y^j \to \mathbb{R}$ by $h\left(y^1, \ldots, y^m\right) = \prod_j h^j\left(y^j\right)$. Then $h \circ g = \prod_j \mu(E^j \| \sigma(g^j))$. Using these properties,

$$\begin{aligned} \mu(\textstyle\bigcap_j E^j) &= \int_\Omega \prod_j \mu(E^j \| \sigma(g^j))_\omega \, \mathrm{d}\mu(\omega) \\ &= \int_{Y^1 \times \cdots \times Y^m} \prod_j h^j\left(y^j\right) \, \mathrm{d} \operatorname{marg}_{Y^1 \times \cdots \times Y^m} \mu(\omega), \end{aligned}$$

where the first line is equation B3 and the second line uses Billingsley [7, 1995, Theorem 16.13]. Now use the fact that $\mathrm{marg}_{Y^1\times\cdots\times Y^m}\,\mu$ is a product measure, and Fubini's Theorem, to get

$$\mu(\textstyle\bigcap_j E^j) = \int_{Y^1\times\cdots\times Y^m} \prod_j h^j\left(y^j\right)\,\mathrm{d}\,\mathrm{marg}_{Y^1\times\cdots\times Y^m}\,\mu(\omega) = \prod_j \int_{Y^j} h^j\left(y^j\right)\,\mathrm{d}\,\mathrm{marg}_{Y^j}\,\mu(\omega). \qquad \text{(B4)}$$

Again using Billingsley [7, 1995, Theorem 16.13], note that, for each $j = 1, \ldots, m$,

$$\int_\Omega \mu(E^j \| \sigma(g^j))_\omega \,\mathrm{d}\mu(\omega) = \int_{Y^j} h^j\left(y^j\right)\,\mathrm{d}\,\mathrm{marg}_{Y^j}\,\mu(\omega).$$

So, by equation B4 and the definition of conditional probability,

$$\mu(\textstyle\bigcap_j E^j) = \prod_j \int_\Omega \mu(E^j \| \sigma(g^j))_\omega \,\mathrm{d}\mu(\omega) = \prod_j \mu(E^j),$$

as required. ■

Proposition 5.1 is an immediate corollary of Proposition B1. The next two examples demonstrate that Proposition 5.1 is tight. Neither CI nor SUFF alone is enough to conclude that independence over hierarchies implies independence over strategies.

Example B1 *Here, type t^c satisfies CI and assesses Ann's and Bob's hierarchies as independent, but does not assess their strategies as independent. Let $T^a = \{t^a\}$, $T^b = \{t^b, u^b\}$, $T^c = \{t^c\}$. Let $\lambda^b\left(t^b\right) \neq \lambda^b\left(u^b\right)$. Figure B1 depicts the measure $\lambda^c\left(t^c\right)$.*[16]

	L	R
U	$\frac{16}{192}$	$\frac{64}{192}$
D	$\frac{12}{192}$	$\frac{48}{192}$

(t^a, t^b)

	L	R
U	$\frac{13}{192}$	$\frac{13}{192}$
D	$\frac{13}{192}$	$\frac{13}{192}$

(t^a, u^b)

Figure B1

Note that $\left[t^b\right] \neq \left[u^b\right]$, since $\lambda^b\left(t^b\right) \neq \lambda^b\left(u^b\right)$. We have

$$\lambda^c\left(t^c\right)\left([U]\cap[L] \mid [t^a]\cap[t^b]\right) = \frac{4}{35} = \frac{4}{7}\times\frac{1}{5} = \lambda^c\left(t^c\right)\left([U] \mid [t^a]\cap[t^b]\right)\times\lambda^c\left(t^c\right)\left([L] \mid [t^a]\cap[t^b]\right),$$

and the corresponding equalities hold for $\lambda^c\left(t^c\right)\left([D]\cap[L] \mid [t^a]\cap[t^b]\right)$, $\lambda^c\left(t^c\right)\left([U]\cap[R] \mid [t^a]\cap[t^b]\right)$, and $\lambda^c\left(t^c\right)\left([D]\cap[R] \mid [t^a]\cap[t^b]\right)$. Also, we clearly have

$$\lambda^c\left(t^c\right)\left([s^a]\cap[s^b] \mid [t^a]\cap[u^b]\right) = \frac{1}{4} = \lambda^c\left(t^c\right)\left([s^a] \mid [t^a]\cap[u^b]\right)\times\lambda^c\left(t^c\right)\left([s^b] \mid [t^a]\cap[u^b]\right)$$

---

[16] This example is due to Ethan Bueno de Mesquita.

for $s^1 = U$ or D and $s^2 = L$ or R, so that CI is established.

Independence over hierarchies is immediate since $\lambda^c(t^c)([t^a]) = 1$, so that

$$\begin{aligned}\lambda^c(t^c)([t^a] \cap [t^b]) &= \lambda^c(t^c)([t^a]) \times \lambda^c(t^c)([t^b]),\\ \lambda^c(t^c)([t^a] \cap [u^b]) &= \lambda^c(t^c)([t^a]) \times \lambda^c(t^c)([u^b]).\end{aligned}$$

Yet we also have

$$\lambda^c(t^c)([U] \cap [L]) = \frac{29}{192} \neq \frac{106}{192} \times \frac{54}{192} = \lambda^c(t^c)([U]) \times \lambda^c(t^c)([L]),$$

so that independence over strategies is violated. (It is easy to check that SUFF is also violated, as it must be.)

Example B2 *Here, type t^c satisfies SUFF and again assesses Ann's and Bob's hierarchies as independent, but again does not assess their strategies as independent. Let $T^a = \{t^a\}$, $T^b = \{t^b, u^b\}$, $T^c = \{t^c\}$. Let $\lambda^b(t^b) \neq \lambda^b(u^b)$. Figure B2 depicts the measure $\lambda^c(t^c)$.*

	L	R
U	$\frac{1}{32}$	$\frac{1}{32}$
D	$\frac{1}{32}$	$\frac{5}{32}$

(t^a, t^b)

	L	R
U	$\frac{2}{32}$	$\frac{4}{32}$
D	$\frac{2}{32}$	$\frac{16}{32}$

(t^a, u^b)

Figure B2

First, notice that

$$\begin{aligned}\lambda^c(t^c)\left([U] \mid [t^a] \cap [t^b]\right) &= \lambda^c(t^c)\left([U] \mid [t^a] \cap [u^b]\right) = \frac{1}{4} = \lambda^c(t^c)([U] \mid [t^a]),\\ \lambda^c(t^c)\left([D] \mid [t^a] \cap [t^b]\right) &= \lambda^c(t^c)\left([D] \mid [t^a] \cap [u^b]\right) = \frac{3}{4} = \lambda^c(t^c)([D] \mid [t^a]).\end{aligned}$$

The corresponding equalities with respect to Bob are immediate, since $\lambda^c(t^c)([t^a]) = 1$. This establishes SUFF.

Independence over hierarchies is immediate, as in Example B1, since $\lambda^c(t^c)([t^a]) = 1$ again. Yet we also have

$$\lambda^c(t^c)([U] \cap [L]) = \frac{3}{32} \neq \frac{8}{32} \times \frac{6}{32} = \lambda^c(t^c)([U]) \times \lambda^c(t^c)(L),$$

so that independence over strategies is violated. (It is readily checked that CI is also violated.)

We conclude by giving a characterization of CI in terms of a sufficiency-like condition (Lemma B2), and of SUFF in terms of a conditional independence-like condition (Lemma B3). Both results follow immediately from Dellacherie-Meyer [10, 1978, pp.9, 36], plus an induction, and so we omit the argument.

Lemma B2 *The random variables* $\overrightarrow{s}^{1}_{i}, \ldots, \overrightarrow{s}^{i-1}_{i}, \overrightarrow{s}^{i+1}_{i}, \ldots, \overrightarrow{s}^{n}_{i}$ *are* $\lambda^i(t^i)$*-conditionally independent given* η^{-i} *if and only if, for each* $j \neq i$ *and each* $E^j \in \sigma(\overrightarrow{s}^{j}_{i})$*,*

$$\lambda^i\left(t^i\right)\left(E^j||\sigma(\eta^{-i})\right) = \textstyle\prod_{j\neq i} \lambda^i\left(t^i\right)\left(E^j||\sigma(\eta^{-i}; \overrightarrow{s}^{k}_{i} : k \neq i, j)\right) \quad a.s.$$

Lemma B3 *The random variable* η^j_i *is* $\lambda^i(t^i)$*-sufficient for* $\overrightarrow{s}^{j}_{i}$ *if and only if, for each* $E^j \in \sigma(\overrightarrow{s}^{j}_{i})$*, and each* $F^k \in \sigma(\eta^k_i)$ *for* $k \neq i, j$*,*

$$\lambda^i\left(t^i\right)\left(E^j \cap \textstyle\bigcap_{k\neq i,j} F^k||\sigma(\eta^j_i)\right) = \lambda^i\left(t^i\right)(E^j||\sigma(\eta^j_i)) \times \textstyle\prod_{k\neq i,j} \lambda^i\left(t^i\right)(F^k||\sigma(\eta^j_i)) \quad a.s.$$

Appendix C Proofs for Section 6

Lemma C1 *Let* E *be a closed subset of a Polish space* X *and* $\mathcal{M}(X;E)$ *be the set of* $\mu \in \mathcal{M}(X)$ *with* $\mu(E) = 1$*. Then* $\mathcal{M}(X;E)$ *is closed.*

Proof. Take a sequence μ_n of measures in $\mathcal{M}(X;E)$, with $\mu_n \to \mu$. It follows from the Portmanteau Theorem that $\limsup \mu_n(E) \leq \mu(E)$. Since $\limsup \mu_n(E) = 1$ for all n, $\mu(E) = 1$ and so $\mu \in \mathcal{M}(X;E)$ as desired. ■

Lemma C2 *The set* R^i_m *is closed for each* i *and* m*.*

Proof. By induction on m.

$m = 1$: Let $E\left(s^i\right)$ be the set of $\mu \in \mathcal{M}\left(S^{-i} \times T^{-i}\right)$ such that s^i is optimal under μ. It suffices to show the sets $E\left(s^i\right)$ are closed. If so, since λ^i is continuous, $(\lambda^i)^{-1}(E(s^i))$ is closed. The set R^i_1 is simply the (finite) union over all sets $\left\{s^i\right\} \times (\lambda^i)^{-1}(E(s^i))$ with $(\lambda^i)^{-1}(E(s^i)) \neq \emptyset$; so, R^i_1 is closed.

First, notice that for each $s^{-i} \in S^{-i}$, the set $\left\{s^{-i}\right\} \times T^{-i}$ is clopen. It follows that

$$\mathrm{cl}\left(\left\{s^{-i}\right\} \times T^{-i}\right) \setminus \mathrm{int}\left(\left\{s^{-i}\right\} \times T^{-i}\right) = \left\{s^{-i}\right\} \times T^{-i} \setminus \left\{s^{-i}\right\} \times T^{-i} = \emptyset,$$

and so, for each $\mu \in \mathcal{M}\left(S^{-i} \times T^{-i}\right)$, $\mathrm{cl}\left(\left\{s^{-i}\right\} \times T^{-i}\right) \setminus \mathrm{int}\left(\left\{s^{-i}\right\} \times T^{-i}\right)$ is μ-null.

Now, take a sequence μ_n of measures in $E\left(s^i\right)$, with $\mu_n \to \mu$. The Portmanteau Theorem, together with the fact that each $\mathrm{cl}\left(\left\{s^{-i}\right\} \times T^{-i}\right) \setminus \mathrm{int}\left(\left\{s^{-i}\right\} \times T^{-i}\right)$ is μ-null, implies that $\mu_n\left(\left\{s^{-i}\right\} \times T^{-i}\right) \to \mu\left(\left\{s^{-i}\right\} \times T^{-i}\right)$.

For each $r^i \in S^i$ and integer n, define

$$x_n(r^i) = \textstyle\sum_{s^{-i}\in S^{-i}} \left[\pi^i\left(s^i, s^{-i}\right) - \pi^i\left(r^i, s^{-i}\right)\right] \mathrm{marg}_{S^{-i}}\, \mu_n\left(s^{-i}\right).$$

Note that $x_n(r^i) \geq 0$, and $x_n(r^i) \to x(r^i)$ where

$$x(r^i) = \sum_{s^{-i} \in S^{-i}} \left[\pi^i \left(s^i, s^{-i} \right) - \pi^i \left(r^i, s^{-i} \right) \right] \operatorname{marg}_{S^{-i}} \mu \left(s^{-i} \right).$$

Thus $x(r^i) \geq 0$ (if $x(r^i) < 0$, then eventually $x_n(r^i) < 0$). With this, $\mu \in E\left(s^i\right)$ as desired.

$m \geq 2$: Assume the lemma holds for m. Then, using the induction hypothesis, it suffices to show that $S^i \times B^i\left(R_m^{-i}\right)$ is closed, i.e., that $B^i\left(R_m^{-i}\right)$ is closed. The induction hypothesis gives that R_m^{-i} is closed. So, by Lemma C1, $\mathcal{M}\left(S^{-i} \times T^{-i}; R_m^{-i}\right)$ is closed in $\mathcal{M}\left(S^{-i} \times T^{-i}\right)$. Since λ^i is continuous, $B^i\left(R_m^{-i}\right)$ is closed. ■

We note the following:

Corollary C1 *If* $t^i \in B^i\left(R_m^{-i}\right)$ *then* $\operatorname{Supp} \lambda^i\left(t^i\right) \subseteq R_m^{-i}$. *Similarly, if* $t^i \in B^i\left(\bigcap_m R_m^{-i}\right)$ *then* $\operatorname{Supp} \lambda^i\left(t^i\right) \subseteq \bigcap_m R_m^{-i}$.

Proof of Proposition 6.1. Begin with part (i), and fix an interactive structure. We will show that the set $\operatorname{proj}_S \bigcap_m R_m$ is a BRS. From this it follows that, for each $\left(s^1, t^1, \ldots, s^n, t^n\right) \in \bigcap_m R_m$, $\left(s^1, \ldots, s^n\right)$ is IU. To see that $\operatorname{proj}_S \bigcap_m R_m$ is a BRS, fix $\left(s^i, t^i\right) \in \bigcap_m R_m^i$. Certainly s^i is optimal under $\operatorname{marg}_{S^{-i}} \lambda^i\left(t^i\right)$, since $\left(s^i, t^i\right) \in R_1^i$. Also, for all m, $\lambda^i\left(t^i\right)\left(R_m^{-i}\right) = 1$, and so $\lambda^i\left(t^i\right)\left(\bigcap_m R_m^{-i}\right) = 1$. From this, $\lambda^i\left(t^i\right)(\operatorname{proj}_{S^{-i}}(\bigcap_m R_m^{-i}) \times T^{-i}) = 1$, or

$$\operatorname{marg}_{S^{-i}} \lambda^i\left(t^i\right)(\operatorname{proj}_{S^{-i}} \textstyle\bigcap_m R_m^{-i}) = 1,$$

as required.

Now part (ii). Construct a structure Φ as follows. For each i, set $T^i = S_M^i$. For each i and $s^i \in S_M^i$, there is a measure $\mu\left(s^i\right) \in \mathcal{M}\left(S^{-i}\right)$, with $\mu\left(s^i\right)\left(S_M^{-i}\right) = 1$, under which s^i is optimal. For each $s^i \in S_M^i$, fix such a measure $\mu\left(s^i\right)$. Construct the measure $\lambda^i(s^i)$ on $S^{-i} \times T^{-i}$ as follows:

$$\lambda^i(s^i)(s^{-i}, t^{-i}) = \begin{cases} \mu\left(s^i\right)\left(s^{-i}\right) & \text{if } s^{-i} = t^{-i}, \\ 0 & \text{otherwise.} \end{cases}$$

Note that the construction is exactly as depicted in Figure 9.2.

We will show that $S_M \subseteq \operatorname{proj}_S \bigcap_m R_m$. Note, by construction, $\operatorname{marg}_{S^{-i}} \lambda\left(s^i\right) = \mu\left(s^i\right)$ and $\lambda\left(t^i\right)(\{\left(s^{-i}, s^{-i}\right) : s^{-i} \in S_M^{-i}\}) = 1$. Certainly, if $s^i \in S_M^i$ then $\left(s^i, s^i\right) \in R_1^i$. Assume inductively that, for all j, $\left(s^j, s^j\right) \in R_m^j$. Then certainly $\lambda^i(t^i)(R_m^{-i}) = 1$, so that $\left(s^i, s^i\right) \in R_{m+1}^i$. Thus $\left(s^i, s^i\right) \in \bigcap_m R_m^i$, establishing the result. ■

Appendix D A Finite-Levels Result

Section 7 showed that IU does not characterize the conditions of CI and RCBR (resp. SUFF and RCBR). Here, we consider an immediate implication of this result, corresponding to the case where players only reason up to some finite number of levels.

Suppose it is given that player i only reasons to m levels. In this case, the hidden variable associated with player i is his hierarchy of beliefs up to m levels.

To formalize this, begin by noticing that, if $\delta_m^i\left(t^i\right) = \delta_m^i\left(u^i\right)$ then $\delta_n^i\left(t^i\right) = \delta_n^i\left(u^i\right)$ for all $n \leq m$. Define composite maps $\eta_{i,m}^j = \delta_m^j \circ \overrightarrow{t}_i^j$ and $\eta_m^{-i} = \delta_m^{-i} \circ \overrightarrow{t}_i$.

Definition D1 *The random variables* $\overrightarrow{s}_i^1, \ldots, \overrightarrow{s}_i^{i-1}, \overrightarrow{s}_i^{i+1}, \ldots, \overrightarrow{s}_i^n$ *are* $\lambda^i(t^i)$***-conditionally independent given the random variable*** η_m^{-i} *if, for all* $j \neq i$ *and* $E^j \in \sigma(\overrightarrow{s}_i^j)$,

$$\lambda^i\left(t^i\right)\left(\bigcap_{j\neq i} E^j || \sigma(\eta_m^{-i})\right) = \prod_{j\neq i} \lambda^i\left(t^i\right)\left(E^j || \sigma(\eta_m^{-i})\right) \quad a.s.$$

Say the type t^i *satisfies* m***-conditional independence (****m****-CI)*** *if* $\overrightarrow{s}_i^1, \ldots, \overrightarrow{s}_i^{i-1}, \overrightarrow{s}_i^{i+1}, \ldots, \overrightarrow{s}_i^n$ *are* $\lambda^i(t^i)$*-conditionally independent given* η_m^{-i}.

Definition D2 *The random variable* $\eta_{i,m}^j$ *is* $\lambda^i(t^i)$***-sufficient for the random variable*** $\overrightarrow{s}_i^j$ *if, for each* $j \neq i$ *and* $E^j \in \sigma(\overrightarrow{s}_i^j)$,

$$\lambda^i\left(t^i\right)\left(E^j || \sigma(\eta_m^{-i})\right) = \lambda^i\left(t^i\right)\left(E^j || \sigma(\eta_{i,m}^j)\right) \quad a.s.$$

Say the type t^i *satisfies* m***-sufficiency (****m****-SUFF)*** *if, for each* $j \neq i$, $\eta_{i,m}^j$ *is* $\lambda^i(t^i)$*-sufficient for* $\overrightarrow{s}_i^j$.

We then get the following corollary to Theorem 7.1:

Corollary D1 *Fix* $m \geq 1$.

(i) There is a game G*, a player* i *in* G*, and an IU strategy* s^i *of* i*, such that the following holds: For any structure* Φ*, there is no* $t^i \in T^i$ *such that* $(s^i, t^i) \in R_{m+1}^i$ *and* t^i *satisfies* m*-CI.*

(ii) There is a game G*, a player* i *in* G*, and an IU strategy* s^i *of* i*, such that the following holds: For any structure* Φ*, there is no* $t^i \in T^i$ *such that* $(s^i, t^i) \in R_{m+1}^i$ *and* t^i *satisfies* m*-SUFF.*

This follows from taking player i to be Charlie in the game of Figure 7.1, the strategy s^i to be his choice Y, and repeating the steps of the proof of Theorem 7.1.

Appendix E Proofs for Section 8

First a definition: A set $\times_{i=1}^n Q^i \subseteq S$ is an **independent best-response set** (**IBRS**) (Pearce [18, 1984]) if, for each i and every $s^i \in Q^i$, there is a $\mu(s^i) \in \times_{j\neq i}\mathcal{M}\left(S^j\right)$ with $\mu(Q^{-i}) = 1$, under which s^i is optimal. It is well known that the set of rationalizable profiles is an IBRS, and every IBRS is contained in the rationalizable set.

To prove Proposition 8.1, we follow exactly the proof of Proposition 6.1(ii), in Appendix C. Throughout, simply replace the set of player i's IU strategies with the set of i's rationalizable strategies. We have to show, in addition, that the structure Φ constructed there satisfies CI and SUFF.

Using the IBRS property, for each rationalizable strategy s^i, there is a product measure $\mu(s^i)$ on S^{-i}, which assigns probability 1 to the rationalizable strategies of players $j \neq i$ and under which s^i is optimal. Construct the measure $\lambda^i(s^i)$ as before.

We now give the intuition for why CI and SUFF hold, and then the formal proof of the proposition. For each $j \neq i$, fix a rationalizable strategy s^j of player j. Consider the hierarchy of beliefs of j induced by the measure $\lambda^j(s^j)$. CI requires that the conditional of $\lambda^i(s^i)$, conditioned on the event that each player $j \neq i$ has the hierarchy induced by $\lambda^j(s^j)$, be a product measure. But this conditional comes from $\mu(s^i)$, conditioned on a certain rectangular subset of strategies for players $j \neq i$. (For each $j \neq i$, consider the other strategies r^j of player j with measures $\lambda^j(r^j)$ that map to the same hierarchy as that induced by $\lambda^j(s^j)$. Take the product of these subsets.) Since $\mu(s^i)$ is a product measure, so is its conditional on any rectangular subset. The same argument establishes SUFF.

Proof of Proposition 8.1. Repeat the proof of Proposition 6.1(ii), replacing S^i_M with the set of i's rationalizable strategies.

Fix $s^i \in T^i$ and $\left(s^{-i}, t^{-i}\right) \in S^{-i} \times T^{-i}$. First assume that $s^{-i} = t^{-i}$. Since $\mu\left(s^i\right)$ is a product measure, we have

$$\lambda^i\left(s^i\right)(\textstyle\bigcap_{j\neq i}\left[s^j\right]\cap\bigcap_{j\neq i}[t^j]) = \mu\left(s^i\right)\left(s^{-i}\right) = \prod_{j\neq i}\mu\left(s^i\right)\left(\{s^j\}\times S^{-i-j}\right) = \prod_{j\neq i}\lambda^i(s^i)([s^j]\cap[t^j]). \tag{E1}$$

Again using the fact that $\mu\left(s^i\right)$ is a product measure, we have

$$\begin{aligned}
\lambda^i\left(s^i\right)(\bigcap_{j\neq i}\left[t^j\right]) &= \lambda^i\left(s^i\right)\left(\left\{\left(r^{-i}, r^{-i}\right) : \delta^{-i}\left(r^{-i}\right) = \delta^{-i}\left(t^{-i}\right)\right\}\right) \\
&= \sum_{\{r^{-i}:\delta^{-i}(r^{-i})=\delta^{-i}(t^{-i})\}} \mu\left(s^i\right)\left(r^{-i}\right) \\
&= \prod_{j\neq i}\sum_{\{r^j:\delta^j(r^j)=\delta^j(t^j)\}} \mu\left(s^i\right)\left(\{r^j\}\times S^{-i-j}\right) \\
&= \prod_{j\neq i}\lambda^i\left(s^i\right)\left(S^{-i-j}\times T^{-i-j}\times\left\{\left(r^j, r^j\right) : \delta^j\left(r^j\right) = \delta^j\left(t^j\right)\right\}\right) \\
&= \prod_{j\neq i}\lambda^i\left(s^i\right)\left(\left[t^j\right]\right).
\end{aligned} \tag{E2}$$

Thus, if $\lambda^i\left(s^i\right)(\bigcap_{j\neq i}\left[t^j\right]) > 0$, E1 and E2 yield

$$\lambda^i\left(s^i\right)\left(\bigcap\nolimits_{j\neq i}\left[s^j\right] \mid \bigcap\nolimits_{j\neq i}\left[t^j\right]\right)=\frac{\prod_{j\neq i}\lambda^i(s^i)([s^j]\cap[t^j])}{\prod_{j\neq i}\lambda^i(s^i)([t^j])}=\prod\nolimits_{j\neq i}\lambda^i(s^i)([s^j]\mid[t^j]). \tag{E3}$$

Now, fix some $k \neq i$, and successively sum the left and right hand sides of E3 over all $s^j = t^j$, for each $j \neq i, k$, to get

$$\lambda^i\left(s^i\right)\left(\left[s^k\right] \mid \bigcap\nolimits_{j\neq i}\left[t^j\right]\right)=\lambda^i(s^i)([s^k]\mid[t^k]). \tag{E4}$$

Equation E4 shows that SUFF is satisfied. (The case that $s^k \neq t^k$ is immediate, since the left and right hand sides are then zero.) Replacing each term on the right side of E3 with the left side of E4 shows that CI is satisfied. (The case that $s^j \neq t^j$ for some $j \neq i$ is again immediate, since the left and right hand sides of E3 are then zero.) ■

Proposition E1 *Fix a game G and a structure* $\left\langle S^1,\ldots,S^n;T^1,\ldots,T^n;\lambda^1,\ldots,\lambda^n\right\rangle$*, where for every* i*, each* $\mathrm{marg}_{S^{-i}}\lambda^i\left(t^i\right)$ *is a product measure. If there is RCBR at the state* $\left(s^1,t^1,\ldots,s^n,t^n\right)$*, then the strategy profile* $\left(s^1,\ldots,s^n\right)$ *is rationalizable in G.*

Proof. Follow the proof of Proposition 6.1(i), in Appendix C. Simply note that since each $\mathrm{marg}_{S^{-i}}\lambda^i\left(t^i\right)$ is a product measure, we get that $\mathrm{proj}_S\bigcap_m R_m$ is an IBRS. ■

We conclude with an example of a structure where: (i) associated with each type is an independent measure on the strategies of other players; but (ii) there is a type that does not satisfy CI.

Example E1 *Consider a three-player game, with strategy sets* $S^a = \{U, D\}$*,* $S^b = \{L, R\}$*, and* $S^c = \{Y\}$*. The type sets are* $T^a = \{t^a, v^a\}$*,* $T^b = \{t^b, u^b\}$*, and* $T^c = \{t^c\}$*, where:*

$\lambda^a(t^a)$ *assigns probability 1 to* (L, t^b, Y, t^c)*;*
$\lambda^a(u^a)$ *assigns probability 1 to* (R, u^b, Y, t^c)*;*
$\lambda^b(t^b)$ *assigns probability 1 to* (U, t^a, Y, t^c)*;*
$\lambda^b(u^b)$ *assigns probability 1 to* (D, u^a, Y, t^c)*;*
$\lambda^c(t^c)$ *assigns probability* $\frac{1}{4}$ *to each of* $(U, t^a, L, t^b), (D, t^a, R, t^b), (D, u^a, L, u^b), (U, u^a, R, u^b)$*.*

Note, $\delta^a\left(t^a\right) \neq \delta^a\left(u^a\right)$ *and* $\delta^b\left(t^b\right) \neq \delta^b\left(u^b\right)$*.*

	L	*R*
U	¼	0
D	0	¼

(t^a, t^b)

	L	*R*
U	0	¼
D	¼	0

(u^a, u^b)

Figure E1

Figure E1 depicts the measure $\lambda^c(t^c)$. Clearly, the marginal on the strategy sets of each type's measure is independent. But CI is violated. For example:

$$\frac{1}{2} = \lambda^c(t^c)([U] \cap [L] \mid [t^a] \cap [t^b]) \neq \lambda^c(t^c)([U] \mid [t^a] \cap [t^b]) \times \lambda^c(t^c)([L] \mid [t^a] \cap [t^b]) = \frac{1}{2} \times \frac{1}{2}.$$

(Note that SUFF is satisfied. As for RCBR, we can easily add payoffs for the players–just makes them all 0's–so that RCBR holds at every state.)

Appendix F An Injectivity Condition

Here we show a method of constructing measures that satisfy conditional independence and sufficiency. We start with (finite) sets $X^1, \ldots, X^m$ of 'observables,' and a measure on $\times_{i=1}^m X^i$. Suppose we can find associated (finite) sets $Y^1, \ldots, Y^m$, to be interpreted as sets of hidden variables, and for each i, an injection from X^i to Y^i. Then there is a natural way to construct a measure on $\times_{i=1}^m \left(X^i \times Y^i\right)$ that agrees the original measure on $\times_{i=1}^m X^i$, and which satisfies conditional independence and sufficiency defined with respect to the hidden variables.

Some notation: Let $\left[x^i\right] = \left\{x^i\right\} \times X^{-i} \times Y$, and define $\left[y^i\right]$ similarly.

Proposition F1 *Let $X^1, \ldots, X^m, Y^1, \ldots, Y^m$ be finite sets and, for each $i = 1, \ldots, m$, let $f^i : X^i \to Y^i$ be an injection. Then, given a measure $\mu \in \mathcal{M}(\times_{i=1}^m X^i)$, there is a measure $\nu \in \mathcal{M}(\times_{i=1}^m \left(X^i \times Y^i\right))$ with:*

(i) $\operatorname{marg}_{\times_{i=1}^m X^i} \nu = \mu$*;*

(ii) $\nu\left(\bigcap_{i=1}^m \left[x^i\right] \mid \bigcap_{i=1}^m \left[y^i\right]\right) = \prod_{i=1}^m \nu\left(\left[x^i\right] \mid \bigcap_{i=1}^m \left[y^i\right]\right)$ *whenever* $\nu\left(\bigcap_{i=1}^m \left[y^i\right]\right) > 0$*;*

(iii) for each $i = 1, \ldots, m$, $\nu(\left[x^i\right] \mid \bigcap_{j=1}^m \left[y^j\right]) = \nu\left(\left[x^i\right] \mid \left[y^i\right]\right)$ *whenever* $\nu(\bigcap_{j=1}^m \left[y^j\right]) > 0$.

Figure F1 depicts the case $m = 2$. Since f^1 and f^2 are injective, the measure ν will assign positive probability to at most one point in each (y^1, y^2)-plane. (We'll give it the probability $\mu((f^1)^{-1}(y^1), (f^2)^{-1}(y^2))$.) Conditions (i), (ii), and (iii) are then clear.

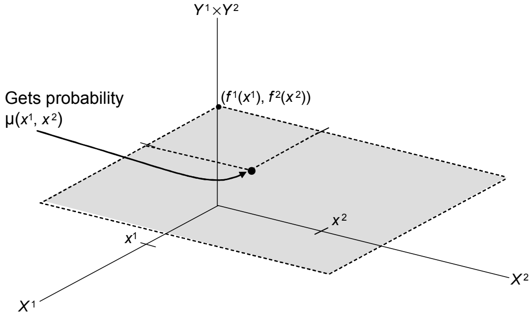


Figure F1

Proof. Define $\nu \in \mathcal{M}(\times_{i=1}^m \left(X^i \times Y^i\right))$ by

$$\nu\left(x^1, y^1, \ldots, x^m, y^m\right) = \begin{cases} \mu\left(x^1, \ldots, x^m\right) & \text{if, for each } i = 1, \ldots, m,\ f^i\left(x^i\right) = y^i, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly $\operatorname{marg}_{\times_{i=1}^m X^i} \nu = \mu$, establishing condition (i).

To establish (ii), first assume $y^i = f^i\left(x^i\right)$ for all i. Then

$$\nu\left(\bigcap_{i=1}^m \left[x^i\right] \mid \bigcap_{i=1}^m \left[y^i\right]\right) = \frac{\nu\left(x^1, f^1\left(x^1\right), \ldots, x^m, f^m\left(x^m\right)\right)}{\nu\left(\bigcap_{i=1}^m \left[f^i\left(x^i\right)\right]\right)} = 1.$$

Also, for each i,

$$\nu\left(\left[x^i\right] \mid \bigcap_{j=1}^m \left[y^j\right]\right) = \frac{\nu\left(\left[x^i\right] \cap \bigcap_{j=1}^m \left[f^j\left(x^j\right)\right]\right)}{\nu\left(\bigcap_{j=1}^m \left[f^j\left(x^j\right)\right]\right)} = 1,$$

so (ii) holds. Next notice that if $y^i \neq f^i\left(x^i\right)$ for some i, then

$$\nu\left(\bigcap_{j=1}^m \left[x^j\right] \mid \bigcap_{j=1}^m \left[y^j\right]\right) = \nu\left(\left[x^i\right] \mid \bigcap_{j=1}^m \left[y^j\right]\right) = 0,$$

and (ii) again holds.

Turning to (iii), if $y^i = f^i\left(x^i\right)$, then

$$\nu\left(\left[x^i\right] \mid \left[y^i\right]\right) = \frac{\nu\left(\left[x^i\right] \cap \left[f^i\left(x^i\right)\right]\right)}{\nu\left(\left[f^i\left(x^i\right)\right]\right)} = 1,$$

and

$$\nu\left(\left[x^i\right] \mid \bigcap_{j=1}^m \left[y^j\right]\right) = \frac{\nu\left(\left[x^i\right] \cap \bigcap_{j=1}^m \left[y^j\right]\right)}{\nu\left(\bigcap_{j=1}^m \left[y^j\right]\right)} = 1,$$

so (iii) holds. If $y^i \neq f^i\left(x^i\right)$, then

$$\nu\left(\left[x^i\right] \mid \bigcap_{j=1}^m \left[y^j\right]\right) = \nu\left(\left[x^i\right] \mid \left[y^i\right]\right) = 0.$$

and (iii) again holds. ■

We now give some implications of Proposition F1. Recall (Definition 6.5) that if $\times_{i=1}^n Q^i$ is a BRS, then for every i and each $s^i \in Q^i$, there is a $\mu\left(s^i\right) \in \mathcal{M}\left(S^{-i}\right)$ with $\mu\left(s^i\right)\left(Q^{-i}\right) = 1$, under which s^i is optimal.

Proposition F2 *Fix a game* $G = \left\langle S^1, \ldots, S^n; \pi^1, \ldots, \pi^n\right\rangle$ *and a BRS* $\times_{i=1}^n Q^i$ *of* G. *Suppose the measures* $\mu\left(s^i\right)$ *can be chosen so that* $\mu\left(r^i\right) \neq \mu\left(s^i\right)$ *if* $r^i \neq s^i$. *Then there is a structure* $\left\langle S^1, \ldots, S^n; T^1, \ldots, T^n; \lambda^1, \ldots, \lambda^n\right\rangle$ *such that, for each strategy profile* $\left(s^1, \ldots, s^n\right) \in \times_{i=1}^n Q^i$, *there is a state* $\left(s^1, t^1, \ldots, s^n, t^n\right)$ *at which there is RCBR, and each type satisfies CI and SUFF.*

Proof. For each i, let T^i be a copy of the set Q^i. We now apply Proposition F1. Fix a player i. For each $j \neq i$, set $X^j = Q^j$ and $Y^j = T^j$. The identity map gives the injection f^j from X^j to Y^j.

For $t^i = s^i \in Q^i$, we construct $\lambda^i\left(t^i\right) \in \mathcal{M}\left(S^{-i} \times T^{-i}\right)$ from $\mu\left(s^i\right)$, the same way that ν is constructed from μ in Proposition F1. (For this, identify a measure on $\mathcal{M}\left(S^{-i} \times T^{-i}\right)$ with support contained in $Q^{-i} \times T^{-i}$, with a measure on $\mathcal{M}\left(Q^{-i} \times T^{-i}\right)$.)

Notice that if $t^j \neq u^j$ then $\operatorname{marg}_{S^{-j}} \lambda^j\left(t^j\right) \neq \operatorname{marg}_{S^{-j}} \lambda^j\left(u^j\right)$. From this it follows that, for any $t^j \neq u^j$, $\delta^j\left(t^j\right) \neq \delta^j\left(u^j\right)$. That is,

$$\left[t^j\right] = S^{-i} \times T^{-i-j} \times \left\{u^j \in T^j : \delta^j\left(u^j\right) = \delta^j\left(t^j\right)\right\} = S^{-i} \times T^{-i-j} \times \left\{t^j\right\}.$$

So, by Proposition F1, t^i satisfies CI and SUFF.

It remains to show that $Q \subseteq \operatorname{proj}_S \bigcap_m R_m$. By construction, if $s^i \in Q^i$ then $\left(s^i, s^i\right) \in R_1^i$. Again by construction, $\lambda\left(s^i\right)\left(\left\{\left(s^{-i}, s^{-i}\right) : s^{-i} \in Q^{-i}\right\}\right) = 1$. Assume inductively that, for all j, $\left\{\left(s^j, s^j\right) : s^j \in Q^j\right\} \subseteq R_m^j$. Then certainly $\lambda^i(s^i)(R_m^{-i}) = 1$, so that $\left(s^i, s^i\right) \in R_{m+1}^i$. Thus $\left(s^i, s^i\right) \in \bigcap_m R_m^i$, which establishes the result. ■

Recall that the IU set is a BRS. So, Proposition F2 tells us that if the game G satisfies the injectivity condition then there is a structure Φ such that, for each IU strategy profile $\left(s^1, \ldots, s^n\right)$, there is a state $\left(s^1, t^1, \ldots, s^n, t^n\right)$ at which there is RCBR, and each type satisfies CI and SUFF.

Now fix an n-player strategic game form $\left\langle S^1, \ldots, S^n\right\rangle$. A particular game can then be identified with a point $\left(\pi^1, \ldots, \pi^n\right) \in \mathbb{R}^{n \times |S|}$. Following Battigalli-Siniscalchi [5, 2003], say the game $\left(\pi^1, \ldots, \pi^n\right)$ satisfies the **strict best-response property** if for each $s^i \in S_M^i$, there exists $\mu \in \mathcal{M}\left(S^{-i}\right)$ with $\mu\left(S_M^{-i}\right) = 1$ such that s^i is the unique strategy optimal under μ. Note, if a game $\left(\pi^1, \ldots, \pi^n\right)$ satisfies the strict best-response property, it also satisfies the injectivity condition (but not vice versa).

Proposition F3 *Let Γ be the set of games for which the strategies consistent with RCBR, CI, and SUFF are strictly contained in the IU strategies. The set Γ is nowhere dense in $\mathbb{R}^{n \times |S|}$.*

Proof. By Proposition F2 and the above remarks, Γ is contained in the sets of games that fail the strict best-response property. Proposition 4.4 in Battigalli-Siniscalchi [5, 2003] shows that for $n = 2$, the set of games that fail the strict best-response property is nowhere dense. Their argument readily extends to $n > 2$, giving our result. ■

Appendix G Relationship to IU Contd.

Here we give a formal treatment of the construction in Section 9.2. Begin by defining modified CI and modified SUFF.

Definition G1 *The random variables $\overrightarrow{s}_i^1, \ldots, \overrightarrow{s}_i^{i-1}, \overrightarrow{s}_i^{i+1}, \ldots, \overrightarrow{s}_i^n$ are $\lambda^i(t^i)$-**conditionally independent given the random variable** $\overrightarrow{t}_i$ if, for all $j \neq i$ and $E^j \in \sigma(\overrightarrow{s}_i^j)$,*

$$\lambda^i\left(t^i\right)\left(\bigcap_{j \neq i} E^j || \sigma(\overrightarrow{t}_i)\right) = \prod_{j \neq i} \lambda^i\left(t^i\right)\left(E^j || \sigma(\overrightarrow{t}_i)\right) \quad a.s.$$

Say the type t^i *satisfies* ***modified CI*** *if* $\overrightarrow{s}_i^1, \ldots, \overrightarrow{s}_i^{i-1}, \overrightarrow{s}_i^{i+1}, \ldots, \overrightarrow{s}_i^n$ *are* $\lambda^i(t^i)$*-conditionally independent given* $\overrightarrow{t}_i$*.*

Definition G2 *The random variable* $\overrightarrow{t}_i^j$ *is* $\lambda^i(t^i)$***-sufficient for the random variable*** $\overrightarrow{s}_i^j$ *if, for each* $j \neq i$ *and* $E^j \in \sigma(\overrightarrow{s}_i^j)$*,*

$$\lambda^i\left(t^i\right)\left(E^j||\sigma(\overrightarrow{t}_i)\right) = \lambda^i\left(t^i\right)\left(E^j||\sigma(\overrightarrow{t}_i^j)\right) \quad a.s.$$

Say the type t^i *satisfies* ***modified SUFF*** *if, for each* $j \neq i$*,* $\overrightarrow{t}_i^j$ *is* $\lambda^i(t^i)$*-sufficient for* $\overrightarrow{s}_i^j$*.*

Proposition G1 *Fix a game* $G = \left\langle S^1, \ldots, S^n; \pi^1, \ldots, \pi^n \right\rangle$*. There is an associated structure* $\left\langle S^1, \ldots, S^n; T^1, \ldots, T^n; \lambda^1, \ldots, \lambda^n \right\rangle$ *such that, for each IU strategy profile* $\left(s^1, \ldots, s^n\right)$*, there is a state* $\left(s^1, t^1, \ldots, s^n, t^n\right)$ *at which RCBR holds, and each type satisfies modified CI and modified SUFF.*

Proof. Construct the structure $\left\langle S^1, \ldots, S^n; T^1, \ldots, T^n; \lambda^1, \ldots, \lambda^n \right\rangle$ as in the proof of Proposition 6.1. As shown there, $S_M \subseteq \mathrm{proj}_S \bigcap_m R_m$. Take each $X^j = S_M^j$ and $Y^j = T^j$. The identity map gives an injection f^j from X^j to Y^j. Then, the construction of each $\lambda^i\left(s^i\right)$ is exactly the construction of each measure $\nu \in \mathcal{M}(\times_{i=1}^m \left(X^i \times Y^i\right))$ in the proof of Proposition F1. So, Proposition F1 establishes that each type satisfies modified CI and modified SUFF. ■

Appendix H Extensions

Here we comment briefly on two possible extensions to our analysis: (i) payoff uncertainty, and (ii) dummy players. We think both extensions are interesting. But do note that both extensions involve analyzing a given game G, by first changing G to a new game and then analyzing this new game.[17] So in the end we don't view this route as explaining correlations in the original game.

i. Payoff uncertainty Payoff functions as well as hierarchies of beliefs are part of the structure of the game. So, uncertainty over payoffs is another source of hidden variables that could be used to explain correlation. In the text, the game was given–i.e., there was no uncertainty over the payoff functions $\pi^1, \ldots, \pi^n$. Now introduce a little uncertainty about payoff functions. Specifically, assume the payoff functions are common $(1-\varepsilon)$-belief (Monderer-Samet [16, 1989], Stinchcombe [20, 1988]), for some (small) $\varepsilon > 0$. (For short, say that a game G itself is common $(1-\varepsilon)$-belief.)

Go back to the game of Figure 7.1 and the associated structure in Figure 9.1. The idea is that now Ann's types t_U^a and t_M^a will both give probability ε to Bob's having a different payoff function from the given one, and the types will differ in what Bob's alternative payoff function is. So t_U^a and t_M^a will induce different hierarchies of beliefs (now defined over strategies and payoff functions).

---

[17] Of course, similar techniques have been used in the equilibrium refinements literature. The key difference is that there the idea is that the analyst doesn't know the true game. Here, as will be seen below, the players themselves know they are in a different game–which we find a less appealing assumption.

We do a similar construction for Bob, and this way (unmodified) CI and SUFF hold in the new structure. We give a general construction in an online appendix.[18]

Yet, this is not a route to understanding the IU strategies in the original game. In introducing payoff uncertainty, we've changed the game from the original one, in which the payoff functions were given.

Even if we allow this change to the game, there is another difficulty. If we introduce payoff uncertainty, we lose a complete characterization of IU. In the online appendix, we show that given $\varepsilon > 0$, we can find a game G where the conditions of RCBR, CI, and SUFF (all redefined for the case of payoff uncertainty), and common $(1-\varepsilon)$-belief of G, allow a non-IU strategy to be played in G.[19] So, while payoff uncertainty can–arguably–rescue the converse direction (part (ii)) of Proposition 6.1, we then lose the forward direction (part (i)).

ii. Dummy players Here the idea is to add another player to the game. In the game of Figure 7.1, we add a fourth player ("Dummy") with a singleton strategy set. Ann's types t^a_U and t^a_M will differ in what they think Dummy thinks about the strategies chosen. Likewise with Bob, and we again get (unmodified) CI and SUFF. The online appendix again gives a general construction.

The same objection applies here. Adding a dummy player is changing the game. The basic question remains: What correlations can be understood in the original game?

---

[18] "Can Hidden Variables Explain Correlation? Online Appendix" available on our webpages.

[19] Note that we first fix ε, and common $(1-\varepsilon)$-belief relative to this ε. Then we find a game where our conditions allow a non-IU strategy. This order is important. Epistemic conditions should be stated independent of a particular game. If the conditions are allowed to depend on the game in question, then the condition could simply be that a strategy profile we're interested in is chosen. This wouldn't be a useful epistemic analysis.

References

[1] Aumann, R., "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics,* 1, 1974, 76-96.

[2] Aumann, R., "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55, 1987, 1-18.

[3] Battigalli, P., and M. Siniscalchi, "Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games," *Journal of Economic Theory*, 88, 1999, 188-230.

[4] Battigalli, P. and M. Siniscalchi, "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory*, 106, 2002, 356-391.

[5] Battigalli, P., and M. Siniscalchi, "Rationalization and Incomplete Information," *Advances in Theoretical Economics*, 3, 2003.

[6] Bernheim, D., "Rationalizable Strategic Behavior," *Econometrica*, 52, 1984, 1007-1028.

[7] Billingsley, P., *Probability and Measure*, 3rd edition, Wiley Series in Probability and Statistics, 1995.

[8] Brandenburger, A., and E. Dekel, "Rationalizability and Correlated Equilibria," *Econometrica*, 55, 1987, 1391-1402.

[9] Diaconis, P., "Finite Forms of de Finetti's Theorem on Exchangeability," *Synthese*, 36, 1977, 271-281.

[10] Dellacherie, C. and P.-A. Meyer, *Probabilities and Potential*, North-Holland Mathematics Studies, 1978.

[11] Kohlberg, E., and J.-F. Mertens, "On the Strategic Stability of Equilibria," *Econometrica*, 54, 1986, 1003-1038.

[12] Marx, L., and J. Swinkels, "Order Independence for Iterated Weak Dominance," *Games and Economic Behavior*, 18, 1997, 219-245.

[13] Mertens, J.-F., "Stable Equilibria–A Reformulation. Part 1. Definition and Basic Properties," *Mathematics of Operations Research*, 14, 1989, 575-624.

[14] Mertens, J.-F., "Ordinality in Non-Cooperative Games," *International Journal of Game Theory*, 32, 2003, 387-430.

[15] Mertens, J.-F., and S. Zamir, "Formulation of Bayesian Analysis for Games with Incomplete Information," *International Journal of Game Theory*, 14, 1985, 1-29.

[16] Monderer, D., and D. Samet, "Approximating Common Knowledge with Common Beliefs," *Games and Economic Behavior*, 1, 1989, 170-190.

[17] Munkres, J., *Topology: A First Course*, Prentice-Hall, 1975.

[18] Pearce, D., "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52, 1984, 1029-1050.

[19] Savage, L., *The Foundations of Statistics*, Wiley, 1954.

[20] Stinchcombe, M. "Approximate Common Knowledge," unpublished, University of California at San Diego, 1988.

[21] Tan, T., and S. Werlang, "The Bayesian Foundations of Solution Concepts of Games," *Journal of Economic Theory*, 45, 1988, 370-391.