Representation of Belief Hierarchies in Games with Incomplete Information

Qingmin Liu*†

Stanford University

This version: October, 2005

Original draft: September 2004

* *S485, GSB, Stanford* e-mail: qiliu@stanford.edu.

†I am most thankful to Yossi Feinberg for numerous invaluable suggestions and conversations and for detailed comments on an early version. I also thank Bob Wilson, Pierpaolo Battigalli, Jeffrey Ely, Françoise Forges, Aviad Heifetz, Jean-Francois Mertens , Marcin Peski, Alfredo Di Tillio and Baozhong Yang for insightful comments, and seminar participants at Stanford GSB, Peking University GSM and Stochastic Methods in Game Theory workshop at Erice for feedbacks. I benefit greatly from communications with Marcin. Of course, all remaining errors are mine.

Abstract

We consider two partition models to be *compatible* if they render the same set of hierarchies of beliefs over a given set of underlying uncertainties. A model is *reduced* if only one partition element corresponds to each hierarchy. We show how to generate all models in a compatible class from the reduced model by introducing a state-dependent correlating mechanism which reveals different information to each player in such a way that no one player gains information about the payoff-relevant states. This provides a solution-concept independent framework for normative analysis of games with incomplete information. We argue that the non-reduced models encapsulate certain payoff-irrelevant yet potentially strategically relevant hidden information entertained by players. Two applications are immediate. We show that the common prior assumption is not an invariant property on a compatible class; we extend Aumann's [2,3] correlated equilibrium to incomplete information games in a straightforward manner.

JEL classification: C72, D80, D82;

Keywords: type spaces, incomplete information, common prior assumption, correlated equilibrium.

1 Introduction

The idea of type spaces, initially proposed by Harsanyi [14] and subsequently formalized by Mertens and Zamir [18,19], Brandenburger and Dekel [6] and Heifetz [15], play an indispensable role in economic modelling utilizing games of incomplete information. Within suitable topologies and under a seemingly vacuous mathematical assumption—the non-redundancy (NR) condition which roughly says that two distinct points in a type space should describe different situations, the universal space exists in the sense that any type space (not necessarily with NR) can be mapped into the space by some belief morphism. The analytic characterization for belief morphisms and hence for all type spaces with redundancy is left open and the problem is clearly important if one notices that the redundancy is actually strategically relevant.

As has been observed (e.g. Battigalli and Siniscalchi [5], Ely and Peski [10], and Dekel et al.[8]), two different models rendering the same set of belief hierarchies over a fixed set of underlying uncertainties, could predict different equilibrium behaviors. The problem, mathematically, lies in the fact that *some* correlations, between redundant types and underlying uncertainties and among redundant types can be embedded into the models in *some* way, while these correlations are prohibited by NR-condition. To visualize the problem, let us look at the following example which simplifies the one in Ely and Peski [10].

FIGURE 1

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

s_1

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

s_2

Coordination Under Uncertainty

EXAMPLE 1: The set of possible underlying uncertainties or states of nature $S = \{s_1, s_2\}$ consists of two possible 2×2 games indexed by s_1 and s_2 as in Figure 1 above. Player 1 is the row player. The additional information the researcher has about the strategic situation is "it is common knowledge[1] that the true game being played is either s_1 or s_2 with equal chance". To formalize this, the researcher comes up with two different partition models.

[1]Knowledge is treated as "belief with probability 1".

The first model $\langle \Omega, (\Pi_i, P_i)_{i=1}^2, \zeta \rangle$ has $\Omega = \{\omega_1, \omega_2\}$ as the state space; $\Pi_i = \{\Omega\}$ is the partition for player $i = 1, 2$; P_i represents the uniform distribution, $i = 1, 2$. The mapping ζ assigns a game to each state of the world, $\zeta(\omega_1) = s_1$ and $\zeta(\omega_2) = s_2$. It is straightforward to see that all the Bayesian-Nash equilibria yield the same expected payoff $(\frac{1}{2}, \frac{1}{2})$.

Now consider the second model $\langle \tilde{\Omega}, (\tilde{\Pi}_i, \tilde{P}_i)_{i=1}^2, \tilde{\zeta} \rangle$: $\tilde{\Omega} = \{\tilde{\omega}_1, \tilde{\omega}_2, \tilde{\omega}_3, \tilde{\omega}_4\}$, $\tilde{P}_i$ represents the uniform distribution, $i = 1, 2$, $\tilde{\zeta}(\tilde{\omega}_1) = \tilde{\zeta}(\tilde{\omega}_4) = s_1$ and $\tilde{\zeta}(\tilde{\omega}_2) = \tilde{\zeta}(\tilde{\omega}_3) = s_2$; the partitions on $\tilde{\Omega}$ are given by Figure 2 below:

$\tilde{\omega}_1$	$\tilde{\omega}_2$
$\tilde{\omega}_3$	$\tilde{\omega}_4$

FIGURE 2

The rows form a partition for player 1 and the columns form a partition for player 2, e.g. $\tilde{\Pi}_1 = \{\{\tilde{\omega}_1, \tilde{\omega}_2\}, \{\tilde{\omega}_3, \tilde{\omega}_4\}\}$. The corresponding partitions of states of nature are given by Figure 3.

s_1	s_2
s_2	s_1

FIGURE 3

In this model, both players believe that s_1 and s_2 happen with equal probability at every state of the world, and hence this fact holds in common knowledge. However, players can now achieve better coordination with the following strategies: player 1 plays U if he believes that the true state of the world lies in the first row and plays D otherwise; player 2 plays L if he believes that the true state of the world lies in the first column and plays R otherwise. The true game is s_1 at states $\tilde{\omega}_1$ or $\tilde{\omega}_4$, and following the strategies, (U, L) and (D, R) are played respectively; at states $\tilde{\omega}_2$ or $\tilde{\omega}_3$, the true game is s_2 and (U, R) and (D, L) are played respectively—(1,1) is obtained.

Although both are consistent with the modeler's information, the two models predict different equilibrium plays. The first model satisfies the NR-condition while the second is not since there are two types rendering the same belief hierarchy for each player. We call the first model *reduced* and the second *non-reduced*.[2] We call the two models *fundamentally compatible* or simply *compatible* to indicate that they represent the same set of belief hierarchies over a given set of

[2]We note that Aumann [4, page 930] uses the term *reduced* for a similar meaning. We did not use the term *redundant* for two reasons: first, the "redundancy" is strategic relevant; secondly, the "redundancy" comes from "hidden information" as we will explain later.

payoff relevant parameters, which are the fundamentals of an incomplete information game without reference to the epistemic structure. Our objectives are (1) structure: we characterize all the compatible models by explicitly formulating the correlations embedded in the information structures in a manageable manner;[3] (2) interpretation: we explain how the players reason about the "redundancy". (3) common prior: we show that the common prior is not an invariant property on a fundamentally compatible class; (4) solution: we characterize the equilibria for a given set of belief hierarchies by considering all possible models.

Structure

Noticing the strategic relevance of redundancy, Ely and Peski [10] show how, given a model, to distill all information that is relevant to the set of rationalizable outcomes. They elegantly construct a "larger" universal space based on infinite hierarchies of beliefs over conditional beliefs or Δ-hierarchies. However, their construction is contingent on the concept of rationalizability; different solution concepts will entail different universal spaces and solution concepts stronger than rationalizability will require even larger universal spaces. So the rationalizability also serves as a "minimal constraint" analogous to the technical NR-condition.[4]

It is desirable to provide a framework in such a way that normative analysis of games with incomplete information (with various solution concepts) can be conducted within the single framework without looking for different ones; this framework should also be linked with the well known Mertens-Zamir framework in order to examine the theories based on Mertens-Zamir framework extensively studied in economics and game theory literature. This can be done only if one can analytically characterize the the structural relation between a reduced model (the well-known Mertens-Zamir space) and its compatible models. This is the main task of the current work.

Let us demonstrate the idea behind our characterization with Example 1. The second model (non-reduced) can be derived from the first one (reduced) by an "unusual" coin tossing procedure: toss a fair coin with a red head (RH) and a green tail (GT) if the state is ω_1 and toss another fair coin with a green head (GH) and a red tail (RT) if the state is ω_2. So the outcomes consist of two dimensions: colors and faces. After the toss, show the color to player 1 and the face

[3]This amounts to characterizing the belief morphism in Mertens-Zamir construction; but we also begin with weaker assumptions.

[4]Dekel et al. [8] essentially regard the reduced space and non-reduced space as the same space and devise a invariant solution concept. These constructions are all solution-concept dependent and none of them tried to characterize the correlations.

to player 2. Then $\{\omega_1\text{RH}, \omega_2\text{RT}, \omega_2\text{GH}, \omega_1\text{GT}\}$ can be identified with $\tilde{\Omega}$; the partitions are given by $\tilde{\Pi}_1 = \{\{\omega_1\text{RH}, \omega_2\text{RT}\}, \{\omega_2\text{GH},\omega_1\text{GT}\}\}$ and $\tilde{\Pi}_2 = \{\{\omega_1\text{RH}, \omega_2\text{GH}\}, \{\omega_1\text{GT},\omega_2\text{RT}\}\}$; the beliefs are uniform by Bayesian updating – this is exactly the second model (see Figure 4 below). We show that such a derivation holds in general: fundamentally compatible models can be fully characterized through such correlating mechanisms.

ω_1RH	ω_2RT
ω_2GH	ω_1GT

FIGURE 4

The characterization, which singles out correlations embedded in a non-reduced model in terms of a correlating mechanism, has several merits due to its analytic nature and solution-concept independency. It allows us to conduct analysis for games of incomplete information within the well-studied Mertens-Zamir framework by adding a correlating mechansim; we can easily examine whether and how the extensively analyzed properties rooted in the reduced models, such as priors and solutions, change over compatible models.

We would like to emphasize that our results clarify the structural relations between more general type spaces rather than simply type spaces with "redundancy." In fact, let $\Omega(S)$ and $\Omega(S \times S')$ be two type spaces built on S and $S \times S'$, respectively; our results characterize all those $\Omega(S \times S')$ such that the "marginal" of the hierarchies in $\Omega(S \times S')$ on S are the same as the hierarchies in $\Omega(S)$.

Interpretation

A non-reduced model is not simply a result of some arbitrary relabelling of belief hierarchies; the "coin" tossed on a reduced model can be interpreted as some specific *payoff-irrelevant* hidden information[5] entertained by players but ignored by the researcher.[6] Thus our result characterizes, in an analytic way, all those hidden information that are compatible with the discovered

[5]In contrast, *payoff-relevant* hidden information, especially for complete information games, is extensively examined in the literature. In the context of equilibrium selction, those hidden information is interpreted as payoff uncertainties and in turn captured by a reduced type space.

[6]In a different context, Brandenburger and Friedenberg [7] consider complete information games where actions serve as the underlying uncertainties. They define "hidden variable" as belief hierarchies and their correlations. It is readily seen that their notion of NR is quite different.

information— Mertens-Zamir belief hierarchies over payoff-relevant parameters.[7] This provides an alternative explanation for the non-reduced models, as well as a workable framework to handle the economic problems arising from payoff-irrelevant hidden information.

In fact, if we regard the non-reduced models as legitimate models to reflect players' reasoning— what they know and what they believe, a player should be able to tell the difference between any two states of the world using his own language. So the seemingly "identical" types must differ in terms of players' beliefs or higher order beliefs over a larger set of primitives than the modeler observes.[8] Actually, Ely and Peksi's [10, section 7.5] interpretation of Δ-hierarchies relies on exogenous "actions" besides the payoff-relevant paramters.

Let us illustrate the point with example 1: when player 1 stands in the first row, he may expect player 2 to learn his information from different sources, either from BBC or CNN. Useful information could be generated if they can pool their information. However, notice that, for player 1 to distinguish the two types of player 2, we have used words "CNN" and "BBC" which are not payoff-relevant parameters on which the original hierarchies are built. We will illustrate this further.

Common Prior

The common prior assumption (CPA), which is closely related with no-trade theorem in economic literature, is made on the partition model rather than directly on the belief hierarchies. We want to examine the implications of non-reduced models on CPA and further on no-trade theorem. In particular, we ask the following question: whether the belief hierarchies over payoff-relevant parameters provide enough information to tell the existence of common priors on a model. This is analogous to the question asked by Ely-Peski: whether the belief hierarchies provide enough information to determine the set of rationalizable outcomes. To visualize our problem, let us look at the following example.

Example 2: Consider the partition model in Figure 5. Player 1's probability assignments

[7] At one extreme, Dekel *et al.* (2005), to explain the non-measurability of players' conjectures in their solution concept, takes nature as an additional player who necessarily knows the true state. This is equivalent to pooling all kinds of outcome relevant hidden information. In contrast, our characterization specifies the whole spectrum of payoff-irrelevant hidden information by the specifications of mechanisms. Important cases include that players have common priors over hidden information.

[8] It is tempting to suggest that we should include statements like "what one will believe if the players could pool their information", or the so called *distributed knowledge*. But note that like common knowledge, distributed knowledge and any other information-relevant operators that players could reason with, should be described in terms of players own language: what they individually know and believe. (see, e.g. Fagin et al. (1995, chapter 2)).

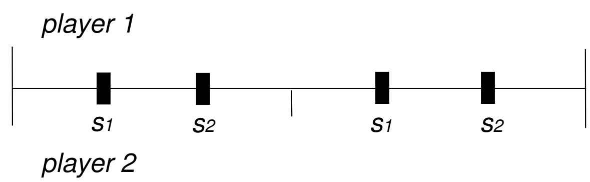


FIGURE 5

over the four states are $(\frac{9}{20}, \frac{1}{20}, \frac{1}{20}, \frac{9}{20})$; player 2's posterior beliefs, conditional on the two partition cells, are $(\frac{1}{2}, \frac{1}{2})$ and $(\frac{1}{2}, \frac{1}{2})$. Clearly there is no common prior on this model, though it contains the same set of belief hierarchies as the two models (with uniform priors) of Example 1: it is commonly known that either s_1 or s_2 prevails with equal probability.

FIGURE 6

	L	S
I	10, 2	0, 1
NI	6, 0	6, 0

s_1

	L	S
I	0, 1	10, 2
NI	6, 0	6, 0

s_2

Investment Under Uncertainty

We now construct a game[9] (see Figure 6 above) to illustrate that the "loss of common prior" is strategic relevant. Player 2, a fund manager, receives \$1 for commission fee whenever player 1, his customer, invests and \$1 bonus whenever he does the right operation. Player 1 could save the money and have \$6 for sure. At s_1, the stock price is going to rise, so the right operation for player 2 to do is "Long"; at s_2, the stock price is going to decline, so the right operation is "Short". It is common knowledge that s_1 and s_2 happen with equal probability, representing the idea that it is commonly known that the fund manager does *not* know more than his costumer.

It is straightforward to see that "Not to Invest"—NI is dominant for player 1 in the reduced model, resulting in the payoff pair (6, 0). But within the model in Figure 5, player 1 invests and expects to receive \$9 and player 2 expects to receive $\$\frac{3}{2}$ in the following equilibrium: Player 1 plays I; Player 2 plays L on the first partition cell and plays S otherwise.

[9]Dekel et al. [8] independently construct a game with similar coordination structure, but their motivation is different.

We now illustrate how the non-reduced model can be generated from the reduced one: a coin is tossed and the outcome (tail or head) is shown only to player 2. Player 2 believes that the same fair coin is tossed at both states, while player 1 thinks that different coins are tossed contingent on which state prevails: he believes that a head will come up with probability $\frac{9}{10}$ at s_1 but only with probability $\frac{1}{10}$ at s_2. Thus conditional a head, player 1 assigns probability $\frac{9}{10}$ on s_1. This leads to the information structure in Figure 5, wherein the previously dominated strategy NI can be played.

Let us further illustrate our hidden-information view of non-reduced models using Example 2. There is other information that the researcher fails to discover, though he finds out it is common knowledge among the players that the stock price is going to rise or decline with equal probability. The fund manager is indoctrinated to "know your customers" in a business school and he happens to hear that the current customer is a superstitious investor: she believes tomorrow's temperature is somehow correlated with the stock price a week later. If it is common knowledge that the fund manager is a layman in finance but instead an expert in weather forecasting, the equilibrium seems reasonable: the customer buys weather information from the fund manager. Were the researcher to confine himself to payoff-relevant parameters, he could find the investment decision irrational.

In Example 2, to obtain from a model with a common prior to a model without common priors, we had chosen a randomizing device without common priors. This is necessarily true as shown in section 4. It is worthwhile to point out the difference between Example 1 and 2 we have given so far. Both examples show that a non-reduced model can be derived from its fundamentally compatible reduced model through a state-dependent correlating mechanism. In Example 1, the correlation comes from the randomizing device itself, e.g. the combinations of colors and faces. This can be called the *objective correlation* in some sense since players agree on the randomizing device. In Example 2, the correlation comes from players' beliefs. It is the disagreement among those players that leads to strategic relevant correlations. This is the *subjective correlation*. This taxonomy is comparable with that in Aumann [2] for complete information games where the state of nature is a singleton. From the hidden-information view of non-reduced models, this taxonomy identified two types of hidden information. Finally, we observe that the taxonomy is strategic relevant as that in Aumann [2]; in particular, for an action which is never a best response in a reduced model to be played by all types with the same belief hierarchy in a compatible non-reduced model we must use some subjective correlation.

We conclude that common prior is not an invariant property within the compatible class. The existence or nonexistence of common priors reflect different kinds of correlations among the underlying uncertainties and economic agents' interactive beliefs. Example 2 shows that a serious researcher would be reluctant to reach the no-trade result if he only knows players' interactive beliefs over payoff-relevant parameters.

Solution

Our characterization, which links compatible models through state-dependent correlating devices, enables us to extend the idea of Aumann [2] in a straightforward way. In Aumann's [2] story, a randomizing device is operated and different players observe different signals before playing a complete information game. The correlated equilibrium is defined as the Nash equilibrium on the resulting information structure. Due to the fact that complete information game contains a single payoff-relevant state, possible correlations arise after enriching the information structures while the (degenerated) beliefs over the original game remains unchanged; it is irrelevant whether the mediator knows the payoffs.

As in Aumann [2], we define correlated equilibrium as the Bayesian-Nash equilibrium on the resulting information structure; but unlike Aumann [2], we have a mediator who knows the true state.[10] The information the mediator reveals to each individual is restricted to be completely uninformative about the fundamentals; however, were the players to aggregate information they receive, they could gain additional information. This is comparable to one player receiving an encrypted message and another player receiving the key. Although actual communication between the players does not take place, the decipherment could emerge in *equilibrium*.

Different solution concepts can be defined depending on whether players are allow to have a common prior on the randomizing device. The subjective version of the solution we characterize turn out to be equivalent to the solution termed "interim rationalizability" independently introduced by Dekel et al. [8]. This will not be surprising if we notice that the two solutions are all based on the same Mertens-Zamir space. Our equilibrium also has some flavor of "agent normal form correlated equilibrium" suggested by Forges [12] where a mediator gives recommendations based on players' types[11]; By applying the characterization in our framework, we obtain new

[10] Forges [12] consider an uninformed mediator who can correlate players' actions only through state-independent mechanisms.

[11] She also provides four additional variations of the correlated equilibrium.

results as well as some properties similar to theirs. We show that we can view the equilibrium concept as an equilibrium for an compatible class, it captures the ignorance of the modeler: the modeler cannot preclude any derived models (or equivalently, compatible payoff-irrelevant hidden information) if he only knows the belief hierarchies over a fixed set of underlying uncertainties and hence considering all possible equilibria in this class of models is equivalent to considering the extended correlated equilibria of the strategic situation known to him. From this perspective, we find the extended equilibrium to be a natural extension of the subjective correlated equilibrium in Aumann [2].

The rest of the paper is organized as follows. We describe the basic model in section 2 and provide the characterization result in section 3; the common prior assumption and the extended equilibrium are discussed in section 4 and section 5, respectively; section 6 concludes. All proofs are in appendices.

2 The Model

Given a set of underlying uncertainties S, which we call *states of nature* parameterizing different games. A *partition model* is a collection $\langle \Omega, (\Pi_i, P_i)_{i \in N}, \zeta \rangle$: Ω is the state space with $\omega \in \Omega$ a state of the world; $N = \{1, ..., n\}$ is the set of players; Π_i is player i's partition on Ω with a typical element π_i and $\Pi_i(\omega) \in \Pi_i$ denotes the element containing ω; $P_i[\cdot|\pi_i]$ denotes player i's belief over π_i; $\zeta : \Omega \to S$ specifies the game being played at a given state of the world. Following the mathematical convention, we simply use Ω to denote the model $\langle \Omega, (\Pi_i, P_i)_{i \in N}, \zeta \rangle$. For well exhibiting the idea and its implications on priors and solutions, we assume S and Ω to be finite in the main text; general cases are treated in Appendix D.

Given a model Ω, player i's (Mertens-zamir) *belief hierarchy* at state ω, $\delta_i(\omega)$, is defined constructively as follows. Player i's first order belief at $\omega \in \Omega$ is given by

$$\delta_i^1(\omega)[s] = P_i[\zeta(\omega') = s|\, \Pi_i(\omega)].$$

Define, for $k \geq 1$,

$$\Phi_i^k = \{\phi_i^k : \phi_i^k = \delta_i^k(\omega) \text{ for some } \omega \in \Omega\};$$

Φ_i^k consists of all the kth order beliefs of player i. Define player i's $k+1$th order belief at ω as

$$\begin{aligned} &\delta_i^{k+1}(\omega)[s \times \phi_{-i}^1 \times \phi_{-i}^2 \times \cdots \times \phi_{-i}^k] \\ &= P_i[\zeta(\omega') = s,\ \delta_{-i}^1(\omega') = \phi_{-i}^1, \cdots, \delta_{-i}^k(\omega') = \phi_{-i}^k |\ \Pi_i(\omega)]. \end{aligned}$$

Player i's (Harsanyi-Mertens-Zamir) belief hierarchy at ω is given by

$$\delta_i(\omega) = (\delta_i^1(\omega), \delta_i^2(\omega), \cdots).$$

Denote $\delta = (\delta_1, \cdots, \delta_n)$. This construction generates coherent belief hierarchies which correspond to types in the universal space. We now formalize the idea of fundamental compatibility.

Definition 1 *Two models, $\langle \Omega, (\Pi_i, P_i)_{i\in N}, \zeta \rangle$ and $\langle \tilde{\Omega}, (\tilde{\Pi}_i, \tilde{P}_i)_{i\in N}, \tilde{\zeta} \rangle$ are fundamentally compatible, written as $\Omega \sim \tilde{\Omega}$, if $\delta(\Omega) = \tilde{\delta}(\tilde{\Omega})$; that is, Ω and $\tilde{\Omega}$ represent the same set of belief hierarchies.*

Remark Fundamental compatibility is obviously an equivalence relation; we define the equivalence relation in terms of the *belief hierarchy* δ instead of the *description* (ζ, δ) for two reasons: first, we want to well exhibit the correlation between states of nature and belief hierarchies; secondly, it is the belief hierarchies that captures players' reasoning; the description is a result of the construction of the mirroring model for such reasoning, as we will show that equivalence of belief hierarchy implies the equivalence of description: this is where the topological assumption (finiteness in this paper and compactness in general) comes in.

Definition 2 *A model $\langle \Omega, (\Pi_i, P_i)_{i\in N}, \zeta \rangle$ is reduced*[12]*. if*

$$\forall i \in N,\ \omega' \notin \Pi_i(\omega) \Rightarrow \delta_i(\omega) \neq \delta_i(\omega').$$

A model is non-reduced if it is not reduced.

Without loss of generality, we add the following assumption on a reduced model: the meet[13]of $\Pi_i,\ i = 1, \cdots, n$, is $\{\Omega\}$.

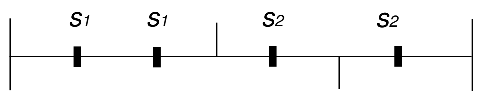
There are two types of "redundancy" in a non-reduced model: first, redundancy related to correlations, which is what we are interested in; secondly, pure redundancy, such as the duplication

[12]Mertens-Zamir [19] formulated "non-redundancy" in terms of description implying the condition here. Again, we have this definition in order to clarifying the correlations between beliefs and states of nature.

[13]Finest common coarsening.

of some existing states illustrated in the following example.

Figure 7



EXAMPLE 3: In Figure 7 above, s_1 is duplicated twice. The pure redundancy is strategic relevant only if one assumes that a player's strategy is, or conjectured by other players to be, non-measurable. Indeed, any normative analysis of games should conform with measurability: it is not pertaining to rationality; it simply says that a player should behave according to his perception.

To focus on the idea of correlation, we introduce the following assumption to rule out the pure redundancy.

Assumption 1 *We assume that* $\omega' \in \Pi_i(\omega) - \{\omega\}$ *for all* $i \in N$ *implies* $\zeta(\omega) \neq \zeta(\omega')$.

In words, if ω and ω' are within the same partition for each player (i.e. within the same element of the join), then they cannot simply duplicate each other.

3 Characterizing Compatibility

As seen in the introduction, compatible models are linked together by the "individually uninformative" state-dependent randomizing devices. The formalization is introduced below and the interpretation is given afterwards.

Definition 3 $\langle (C^\omega)_{\omega\in\Omega}, (q_i^\omega)_{\omega\in\Omega, i\in N}\rangle$ *is a state-dependent correlating mechanism on* Ω, *if*

(a) *there exist finite sets* C_i, $i \in N$, *such that*

$C^\omega \subseteq C_1 \times C_2 \times \cdots \times C_n$, *and*

$C^{\omega'}|_{C_i} = C^\omega|_{C_i}$, *if* $\omega' \in \Pi_i(\omega)$;

(b) q_i^ω *is a probability measure on* C^ω *such that*

if $\omega' \in \Pi_i(\omega)$; *then* $\sum_{\{c\in C^{\omega'}:\, c_i=c_i^\omega\}} q_i^{\omega'}(c) = \sum_{\{c\in C^{\omega}:\, c_i=c_i^\omega\}} q_i^{\omega}(c)$, *for all* $c_i^\omega \in C_i^\omega$.

The interpretation is as follows. At each state ω, a randomizing device with outcomes of n-dimension characteristics (i.e. $C_1^\omega, \cdots, C_n^\omega$) is operated, and each player i is shown only the ith coordinate of the outcome, c_i^ω; the randomizing devices are designed such that player i cannot distinguish the states in $\Pi_i(\omega)$ by simply observing the outcomes – this is condition (a). Note that the superscript ω on c_i^ω only indicates that the observation is obtained when the true state is ω, but player i does not know ω unless $\Pi_i(\omega)$ is a singleton since all states in $\Pi_i(\omega)$ could entail c_i^ω. In condition (b), q_i^ω is interpreted as the prior belief player i holds over the device operated at ω, C^ω; $\sum_{\{c \in C^{\omega'}:\, c_i = c_i^\omega\}} q_i^{\omega'}(c)$ being constant over $\Pi_i(\omega)$ says that the randomizing device cannot change the likelihood ratio between different states; that is, no one player gains payoff-relevant information. Note that if C^ω's and q_i^ω's are independent with ω, the story here is the same as that in Aumann [2]. As argued informally in the introduction, the first component of the mechanism corresponds to "objective correlation" (see Example 1) and the second "subjective correlation" (see Example 2). This taxonomy plays a role in common priors (Proposition 1) and in strategic issues (Proposition 3).

The following theorem fully characterizes the fundamental compatibility. It provides a program to construct all other models from the reduced models. Interpretations of the theorem are provided afterwards. All omitted proofs can be found in appendices.

Theorem 1 *Suppose Ω is reduced and $\tilde{\Omega}$ is non-reduced. Then $\Omega \sim \tilde{\Omega}$ if and only if there exist a state-dependent mechanism $\langle (C^\omega)_{\omega \in \Omega}, (q_i^\omega)_{\omega \in \Omega, i \in N} \rangle$ on Ω, such that*

(i) $\tilde{\Omega} = \{\omega \times c^\omega\}_{\omega \in \Omega,\, c^\omega \in C^\omega}$;

(ii) $\tilde{\Pi}_i(\omega \times c^\omega) = \{\omega' \times c^{\omega'} : \omega' \in \Pi_i(\omega)$ *and* $c^{\omega'} \in C^{\omega'}$ *with* $c_i^{\omega'} = c_i^\omega\}$;

(iii) $\tilde{P}_i(\omega' \times c^{\omega'} | \omega \times c^\omega) = P_i(\omega' | \omega) \cdot q_i^{\omega'}(c^{\omega'} \mid c_i^\omega)$, *where* $q_i^{\omega'}(c^{\omega'} \mid c_i^\omega) = \dfrac{q_i^{\omega'}(c^{\omega'})}{\sum_{\{c \in C^{\omega'}:\, c_i = c_i^\omega\}} q_i^{\omega'}(c)}$;

(iv) $\tilde{\zeta}(\omega \times c^\omega) = \zeta(\omega); \forall \omega \times c^\omega \in \tilde{\Omega}$.

The equalities in the theorem should be read as "can be identified with" in the sense that there is an isomorphism that preserves the properties of a model. Mathematically, the relation between $\tilde{\Omega}$ and Ω is characterized by a transition probability with certain properties.

The theorem can be interpreted as the introduction of a randomizing device when considering a non-reduced model. At each state ω, a randomizing device with n-dimension outcomes is operated; each player i is shown only the ith coordinate, c_i^ω – this is condition (i). After observing c_i^ω, player i then believes that all states in $\Pi_i(\omega)$ and the corresponding outcomes with ith coordinate c_i^ω

are possible – this is condition (ii). In condition (iii), $q_i^{\omega'}(c^{\omega'}|c_i^{\omega})$ is the probability player i assigns to outcome $c^{\omega'}$ when he observes c_i^{ω}; $\tilde{P}_i$ is then the result of Bayesian updating from P_i. After learning c_i^{ω}, player i updates his beliefs as follows:

$$\begin{aligned}\tilde{P}_i(\omega' \times c^{\omega'}|\omega \times c^{\omega}) &= \frac{P_i(\omega'|\,\omega)\cdot q_i^{\omega'}(c^{\omega'})}{\sum_{\omega''\in\Pi_i(\omega)}\sum_{\{c\in C^{\omega''}:\,c_i=c_i^{\omega}\}}[q_i^{\omega''}(c)\cdot P_i(\omega''|\omega)]} &(1)\\ &= \frac{P_i(\omega'|\,\omega)\cdot q_i^{\omega'}(c^{\omega'})}{\sum_{\{c\in C^{\omega''}:\,c_i=c_i^{\omega}\}} q_i^{\omega''}(c)\cdot\sum_{\omega''\in\Pi_i(\omega)}P_i(\omega''|\omega)} &(2)\\ &= P_i(\omega'|\omega)\frac{q_i^{\omega'}(c^{\omega'})}{\sum_{\{c\in C^{\omega''}:\,c_i=c_i^{\omega}\}} q_i^{\omega''}(c)} &(3)\\ &= P_i(\omega'|\omega)\frac{q_i^{\omega'}(c^{\omega'})}{\sum_{\{c\in C^{\omega'}:\,c_i=c_i^{\omega}\}} q_i^{\omega'}(c)}. &(4)\end{aligned}$$

(1) is by Bayes' rule; (2) is by the property of $q_i^{\omega''}$– (b) in Definition 3; (3) is by the definition of P_i; (4) is by the same property of $q_i^{\omega''}$ again.

Note that C_i, the set of signals that player i could observe, can be very large depending on $\tilde{\Omega}$, and there is no uniform bound on $|C_i|$ (due to the non-existence of universal non-reduced space[14]). Nevertheless, given a finite game, there is a "strategic bound" on C_i, depending only on the size of the game, with which all the strategic relevances with (w.r.t. the extended correlated equilibrium) of interest are created.[15] This is the content of Proposition 2 in section 5.

4 On Common Priors

Theorem 1 characterizes the compatible models through correlating mechanisms. In this section, we examine its implication on common priors. As illustrated by Example 2, one need distill more information than just belief hierarchies over the payoff-relevant parameters to tell the existence of a common prior on a non-reduced model. The common prior property is related to the correlations embedded in the model. Our characterization enables us to examine this relationship through the correlating mechanism. Suppose Ω is reduced and $\Omega \sim \tilde{\Omega}$. Let $\langle (C^{\omega})_{\omega\in\Omega}, (q_i^{\omega})_{\omega\in\Omega, i\in N}\rangle$ be the corresponding mechanism as that in Theorem 1. We say that there is a common prior on this mechanism if $q_i^{\omega} = q_j^{\omega}$ for all $i, j \in N$ and $\omega \in \Omega$. We summarized our results here.

[14]In Example 5 of Appendix D2, we show how to construct a belief space without pure redundancy yet with arbitrary cardinality.

[15]I thank Avaid Heifetz for pointing out the question to me.

Proposition 1

(i) *If there is no common prior on* Ω*; then there is no common prior on* $\tilde{\Omega}$;

(ii) *If* $q_i^\omega = q_j^\omega$ *for all* $i, j \in N$ *and* $\omega \in \Omega$*; then there is a common prior on* Ω *if and only if there is a common prior on* $\tilde{\Omega}$.

The converse of (i) does not hold in general as seen from Example 2; (ii) says that to obtain from a model with a common prior to a model without common priors one must have a correlating device without common priors. But the condition in (ii) is not necessary, as shown in Example 4.

Example 4: Let $\Omega = \{\omega\}$ be a singleton set and so the model has a common prior trivially. Toss a coin and show the outcome only to player 2. Suppose $q_1^\omega(\text{"}Head\text{"}) = \frac{1}{2}$ and $q_2^\omega(\text{"}Head\text{"}) = \frac{1}{3}$. Clearly, there is a common prior $(\frac{1}{2}, \frac{1}{2})$ on the resulting model $\tilde{\Omega}$. This shows that a correlating device without common priors operated on a model with a common prior could result in a model with common priors.

Proposition 1 has some implications in economic modelling. If a researcher holds the view that "differences in probabilities express differences in information only" and hence use a model with common prior, then he would, consistently, prefer to model the hidden information with a correlating device with objective correlation only, which guarantees the common prior property on the non-reduced model. Some implications of Proposition 1 on games are discussed in next section.

5 Solutions

5.1 Extended Correlated Equilibrium

Let A_i be a finite set of actions of player i and A_{-i} take its usual meaning. We write $A = A_1 \times \cdots \times A_n$. Player i's (pure) *strategy* is a mapping $\alpha_i : \Omega \to A_i$, which are Π_i-measurable representing the idea that a player knows only his own types. The payoff for player i is given by a bounded real function $g_i : A \times S \to \mathbb{R}$, which is extended multilinearly to mixed strategies. For any model $\langle \Omega, (\Pi_i, P_i)_{i=1}^n, \zeta \rangle$, define

$$\hat{g}_i(a_i, a_{-i}, \omega) := g_i(a_i, a_{-i}, \zeta(\omega)).$$

We still write it as g_i with some abuse of notation. Thus, $\langle \Omega, N, A, (\Pi_i, P_i)_{i\in N}, (g_i)_{i\in N}\rangle$ represents a *game of incomplete information* associated with the partition model Ω.[16]

Definition 4 *A pure strategy profile* $(\alpha_1, \cdots, \alpha_n)$ *is a Bayesian-Nash equilibrium for the game associated with* Ω *if* $E_i[g_i(\alpha_i, \alpha_{-i}, \omega)|\Pi_i] \geq E_i[g_i(a_i, \alpha_{-i}, \omega)|\Pi_i], \ \forall a_i \in A_i, \ i \in N.$

Note that considering mixed strategies entails no real difficulties but cumbersome notations and proofs; it is without loss of generality to consider pure strategies since we allow Ω to be non-reduced to incorporate the randomness in strategies. The equilibrium outcomes are expanded after introducing state-dependent correlating mechanisms. Run $\langle (C^\omega)_{\omega\in\Omega}, (q_i^\omega)_{\omega\in\Omega, i\in N}\rangle$ before playing the game associated with Ω and then show the ith-characteristics of the outcome to player i. In this scenario, player i can obtain two pieces of information when the true state is ω: $\Pi_i(\omega)$ and c_i^ω. But he will not know ω for sure unless $\Pi_i(\omega)$ is a singleton: he would expect to see c_i^ω at any state $\omega' \in \Pi_i(\omega)$ according to the nature of the mechanism. Player i's information set, when c^ω is realized at ω, is $I_i(\omega \times c^\omega) = \{\omega' \times c^{\omega'} : c^{\omega'} = c_i^\omega, c^{\omega'} \in C^{\omega'}, \omega' \in \Pi_i(\omega)\}$. His belief is updated via Bayes' rule.

Definition 5 *A (subjective) correlated strategy for player* i*, relative to the state-dependent correlating mechanism* $\langle (C^\omega)_{\omega\in\Omega}, (q_i^\omega)_{\omega\in\Omega, i\in N}\rangle$*, for the game associated with* Ω *is an* I_i*-measurable mapping* $\alpha_i : \{\omega \times c^\omega : c^\omega \in C^\omega, \omega \in \Omega\} \to A_i$.

We should distinguish the meaning of $I_i(\omega \times c^\omega)$ and the partition element $\tilde{\Pi}_i(\omega \times c^\omega)$ for a non-reduced model $\tilde{\Omega}$; they are mathematically identical due to Theorem 1; this coincidence is exactly the motivation of our solution concept.

Definition 6 *A profile of correlated strategies* $(\alpha_1, \cdots, \alpha_n)$ *is a (subjective) correlated equilibrium, relative to a state-dependent correlating mechanism* $\langle (C^\omega)_{\omega\in\Omega}, (q_i^\omega)_{i\in N,\, \omega\in\Omega}\rangle$*, for the game associated with* Ω *if*

$$E_i[g_i(\alpha_i, \alpha_{-i}, \omega)|I_i] \geq E_i[g_i(a_i, \alpha_{-i}, \omega)|I_i] \quad \forall a_i \in A, \ i \in N.$$

Within the context of complete information games, our definition reduces to the subjective equilibrium in Aumann [2,3]. One can also define an objective version correlated equilibrium by

[16] ζ does not appear in the definition, since it is captured by the payoff g_i's.

assuming common priors on the correlating mechanism and the underlying model. As commented before, there is a "strategic bound" on the correlating mechanism: we can without loss of generality take correlating mechanisms with $C_i = A_i$ for all $i \in N$. This is consistent with the well-known fact for complete information games.

Proposition 2 *Let $\alpha = (\alpha_1, \cdots, \alpha_n)$ be a correlated equilibrium with respect to the mechanism $\langle (C^\omega)_{\omega\in\Omega}, (q_i^\omega)_{\omega\in\Omega, i\in N}\rangle$. One can construct another mechanism $\langle (\tilde{C}^\omega)_{\omega\in\Omega}, (\tilde{q}_i^\omega)_{\omega\in\Omega, i\in N}\rangle$ with $C_i = A_i$ for all $i \in N$, wherein the same equilibrium outcome obtains.*

Proof Define $\tilde{C}^\omega = \{a \in A : a = \alpha(\omega \times c),\ c \in C^\omega\}$ and $\tilde{q}_i^\omega(a) = \sum_{\{c\in C^\omega : \alpha(\omega\times c)=a\}} q_i^\omega(c)$. It is routine to check that $\langle (\tilde{C}^\omega)_{\omega\in\Omega}, (\tilde{q}_i^\omega)_{\omega\in\Omega, i\in N}\rangle$ is a correlating mechanism and $\tilde{\alpha}(\omega \times a) = a$ defines a correlated equilibrium. □

We see from the proof that if $q_i^\omega = q_j^\omega$ then $\tilde{q}_i^\omega = \tilde{q}_j^\omega$; that is, for any correlating mechanism with common priors, we can construct a "strategically equivalent" bounded mechanism inheriting the common prior property. This fact together with Proposition 2 is used to show the strategic relevance of the taxonomy of objective correlation and subjective correlation.

We say an action a_i is *never a best response* for player i at $\omega \in \Omega$ if for any belief over his opponents' strategies, $\sigma_{-i} : \Omega \to \Delta A_{-i}$, which is measurable with respect to Π_{-i}, the join of Π_j's for $j \neq i$, there exists $a_i' \in A_i$ such that $E[g_i(a_i, \sigma_{-i}, \cdot)|\Pi_i(\omega)] < E[g_i(a_i', \sigma_{-i}, \cdot)|\Pi_i(\omega)]$.

Proposition 3 *In a two-player game*[17]*, suppose $\Omega \sim \tilde{\Omega}$ and Ω is reduced. If a_1^* is never a best response for player 1 at $\omega^* \in \Omega$ and if (α_1, α_2) is a Bayesian-Nash equilibrium on $\tilde{\Omega}$ with $\alpha_1(\tilde{\omega}) = a_1^*$ for all $\tilde{\omega} \in \tilde{\Omega}$ such that $P_1^*(\omega^*) = \tilde{P}_1^*(\tilde{\omega})$; then there is no common prior on the correlating mechanism $\langle (C^\omega)_{\omega\in\Omega}, (q_i^\omega)_{\omega\in\Omega, i\in N}\rangle$.*

The main idea of the proposition is as follows. The strategic relevance of the objective correlation comes from the fact that "one player receives an encrypted message and another player receiving the key" and "the encrypted message is deciphered in *equilibrium*" though no actual communication ever takes place. Now, we are forcing one player to take the same action regardless the messages he receives; that is, this player does not bring the encrypted message to the equilibrium and hence the message never have a chance to be deciphered. If it turns out that

[17] As seen from the proof, the result holds for more than two players. We add this assumption for simplicity.

some new fact happens in equilibrium, it must come from the subjective correlation (see Example 2).

Proof By virtue of Theorem 1, the given Bayesian-Nash equilibrium on $\tilde{\Omega}$, α, is identified with a correlated equilibrium on Ω with respect to $\langle (C^\omega)_{\omega\in\Omega}, (q_i^\omega)_{\omega\in\Omega, i\in N}\rangle$. By Proposition 2, we can construct a "strategic equivalent" mechanism $\langle (\tilde{C}^\omega)_{\omega\in\Omega}, (\tilde{q}_i^\omega)_{\omega\in\Omega, i\in N}\rangle$ (as defined in the proof of Proposition 2,) on which the equilibrium $\tilde{\alpha}$, with the same outcome as α, obtains. The new mechanism leads to another information structure, still written as $\tilde{\Omega}$ with some abuse of notations. $\tilde{\alpha}$ can in turn be treated as a Bayesian-Nash equilibrium on $\tilde{\Omega}$. It then suffices to show that there is no common prior on $\langle (\tilde{C}^\omega)_{\omega\in\Omega}, (\tilde{q}_i^\omega)_{\omega\in\Omega, i\in N}\rangle$, by the remark after proposition 2. Notice that $\tilde{C}_1^\omega = \{a_1^*\}$ for $\omega \in \Pi_1(\omega^*)$ by assumption.

Suppose, to the contrary, there is a common prior on the mechanism, i.e. $\tilde{q}_1^\omega = \tilde{q}_2^\omega = \tilde{q}^\omega$ for all $\omega \in \Omega$. Define a joint mixed strategy for players other than 2, $\sigma_2 : \Omega \to \Delta A_2$ as follows:

If $\omega' \in \Pi_1(\omega^*)$, *then* $\sigma_2(\omega')(a_2) = q^{\omega'}(a)$ *where* $a = (a_1^*, a_2) \in \tilde{C}^{\omega'}$.

If $\omega' \notin \Pi_1(\omega^*)$, *then* $\sigma_2(\omega')$ *is defined arbitrarily with only measurability constraint.*

We notice that σ_2 is Π_2-measurable: $q^\omega = q^{\omega'}$ if $\omega' \in \Pi_2(\omega)$ by Definition 3(b). We write $a_2^* = \alpha_2(\omega^*)$ and $a^* = (a_1^*, a_2^*)$. We then claim that a_1^* is a best response to σ_2 at ω^* in the game associated with Ω. To see this,

$$\begin{aligned}
E_1[g_1(a_1^*, \sigma_2, \cdot)|\Pi_1(\omega^*)] &= \sum_{\omega\in\Pi_1(\omega^*)} \sum_{a_2\in A_2} g_1(a_1^*, a_2, \omega) q^\omega((a_1^*, a_2)) P_1[\omega|\Pi_1(\omega^*)] \\
&= \sum_{\omega\times(a_1^*, a_2)\in\tilde{\Pi}_1(\omega\times a^*)} g_1(a_1^*, a_2, \omega\times(a_1^*, a_2))\tilde{P}_1[\omega\times(a_1^*, a_2)|\tilde{\Pi}_1(\omega^*\times a^*)] \\
&= E_1[g_1(a_1^*, \alpha_2, \cdot)|\tilde{\Pi}_1(\omega^*\times a^*)] \\
&\geq E_1[g_1(a_1, \alpha_2, \cdot)|\tilde{\Pi}_1(\omega^*\times a^*)], \forall a_1 \in A_1 \\
&= E_1[g_1(a_1, \sigma_2, \cdot)|\Pi_1(\omega^*)], \forall a_1 \in A_1
\end{aligned}$$

This provides a desired contradiction. □

Propositions 2 and 3 together give the following result.

Corollary 1 *With the assumptions in Proposition 3, if in addition there is a common prior on* Ω*, then there is no common prior on* $\tilde{\Omega}$*.*

Let $\mathcal{N}(\Omega)$ and $\mathcal{C}(\Omega)$ denote the sets of Bayesian-Nash equilibrium outcomes and correlated equilibrium outcomes for the game associated with Ω, respectively. The next result states that the correlated equilibrium outcomes associated with a reduced model are obtained by considering all the Bayesian-Nash equilibrium outcomes associated with fundamentally compatible models. The proof is simply a comparison of definitions due to Theorem 1.

Proposition 4 *Suppose Ω is a reduced model; then*

$$\mathcal{C}(\Omega) = \bigcup_{\tilde{\Omega}\sim\Omega} \mathcal{N}(\tilde{\Omega}).$$ [18]

The following result further strengthens Proposition 4. The correlated equilibrium does not depend on a particular model within a compatibility class.

Proposition 5 *If $\Omega_1 \sim \Omega_2$ then $\mathcal{C}(\Omega_1) = \mathcal{C}(\Omega_2)$.*[19]

The formal proof can be found in Appendix. Here we only sketch the main idea. Let Ω be the reduced model which is compatible with Ω_1 and Ω_2; let $\mathcal{D}_1$ and $\mathcal{D}_2$ be the correlating devices used to obtain Ω_1 and Ω_2 from Ω according to Theorem 1 (see Figure 7). It suffices to show $\mathcal{C}(\Omega_2) = \mathcal{C}(\Omega)$. Note that $\mathcal{C}(\Omega_2) \subset \mathcal{C}(\Omega)$ since any partition models obtained by augmenting Ω_2 via a correlating device is compatible with Ω_2 and hence with Ω. For the other direction, we can construct another model Ω_3 from Ω_2 via a correlating device $\mathcal{D}_3$, such that

$$\mathcal{N}(\Omega_1) \subset \mathcal{N}(\Omega_3) \subset \mathcal{C}(\Omega_2) \tag{5}$$

The intuition for this step is illustrated in Figure 8. Ω_3 is obtained from Ω by sequentially running two correlating device $\mathcal{D}_2$ and $\mathcal{D}_3$. If $\mathcal{D}_3$ is constructed to "mimic" $\mathcal{D}_1$, then any Nash equilibrium in Ω_1 can be achieved in Ω_3 if all players "ignore" the correlating device $\mathcal{D}_2$. Since Ω_1 is arbitrary, $\mathcal{C}(\Omega) \subset \mathcal{C}(\Omega_2)$ follows immediately from Proposition 2 and (5).

In many games, the set of correlated equilibria is strictly bigger than the set of Bayesian-Nash equilibria. For example, by slightly disturbing the payoff in Example 1, we obtain an open set of such games.

[18] See also Dekel et al. [8, Proposition 2] and Battigalli and Siniscalchi [5] for related results in their frameworks.

[19] Similarly, Proposition 1 in Dekel et al. [8] states that two types, representing the same belief hierarchies, have the same rationalizability action set.

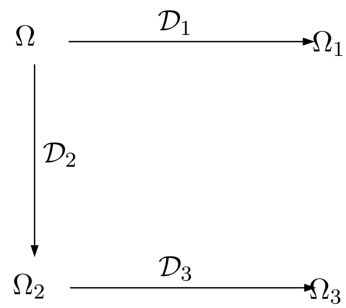


Figure 8

5.2 Rationalizability

Battigalli and Siniscalchi [5] define a concept called Δ-rationalizability by accommodating common knowledge restrictions on first-order beliefs without requiring construction of a type space. Dekel et al. [8] introduces the correlated interim rationalizability concept which in a sense extend Battigalli and Siniscalchi's idea by utilizing information from the complete specification of the infinite hierarchy.

Definition 7 *Let $\langle \Omega, (\Pi_i, P_i)_{i \in N}, \zeta \rangle$ be any model. The interim rationalizable action set for player $i \in N$ at $\omega \in \Omega$, $R_i(\omega)$, is defined inductively as follows:*[20]

$$R_i^0(\omega) = A_i;$$

$$R_i^k(\omega) = \left\{ a_i \in R_i^{k-1}(\omega) : \begin{array}{l} \exists \sigma_{-i} : \omega' \mapsto \mu \in \triangle R_{-i}^{k-1}(\omega'),\ s.t. \\ (1)\sigma_{-i} \text{ is } \Pi_i\text{-measurable}; \\ (2)a_i \in \text{argmax } \mathrm{E}_i[g_i|\omega] \end{array} \right\}$$

$$\text{where } k \geq 1 \text{ and } \quad E_i[g_i|\omega] = \sum_{\Omega} \sum_{A_{-i}} g(a_i, a_{-i}, \omega') \cdot \sigma_{-i}(\omega')(a_{-i}) \cdot P_i(\omega'|\omega);$$

$$R_i(\omega) = \bigcap_{k \geq 0} R_i^k(\omega);$$

[20] Dekel *et al.* iterates over types when defining the concept, but the two approaches are equivalent.

In the definition above, player i's conjecture $\sigma_{-i}(\cdot)$ is Π_i-measurable, i.e. player i knows his conjecture; however, there is no measurability requirement of $\sigma_{-i}(\cdot)$ with respect to $\Pi_j, j \neq i$; i.e. from player i's point of view, other players's strategies can depend on the true states of nature, even though they cannot actually distinguish them.

Dekel et al. [8] also shows that all the rationalizable actions are in the support of some Bayesian Nash equilibrium for some partition model that represents the same set of belief hierarchies. This result together with our Proposition 4 implies the equivalence of the correlated equilibrium and the interim rationalizability.

Proposition 6 *For a given partition model, any correlated equilibrium outcome corresponds to some interim rationalizability outcome, and vice versa.*

6 Conclusion

Our main contribution in this work is to characterize the non-reduced partition models from the standard reduced models through a state-dependent correlating device, where the correlations between (redundant) types and underlying uncertainties and among types are formulated explicitly. Thus, we can capture and analyze all the possible representations of belief hierarchies within the traditional framework initiated by Harsanyi [14]. We argue that the redundancy is caused by payoff-irrelevant hidden information. As applications to the characterization, we show that the common prior assumption is not a property of belief hierarchies over payoff-relevant parameters alone; the correlated equilibrium, which extends Aumann [2,3], can be naturally defined through the state-dependent correlating device. Possible further applications include mechanism design and implementation theory: the social planner knows the payoff-relevant states and all the possible belief hierarchies over those states while agents may entertain further compatible hidden information that are strategic relevant. We leave it for further research.

Appendix A: Proof of Theorem 1

To prove the theorem, we need the following lemmas.

It is straightforward to see that, Ω and $\tilde{\Omega}$ are compatible if and only if there are set valued mappings $h_i : \Omega \rightrightarrows \tilde{\Omega}$, $\forall i \in N$, such that

$$\begin{aligned} &\text{(i)} \quad h_i(\omega) \neq \emptyset, \ \forall \omega \in \Omega, i \in N; \\ &\text{(ii)} \quad h_i(\Omega) = \tilde{\Omega}, \ \forall i \in N; \\ &\text{(iii)} \quad \delta_i(\omega) = \tilde{\delta}_i(\tilde{\omega}), \ \forall \tilde{\omega} \in h_i(\omega), \omega \in \Omega, i \in N. \end{aligned}$$

Conditions (iii) says that h_i's preserve beliefs; conditions (i) and (ii) say that one model contains all the belief hierarchies in the other information system and nothing more. This definition captures the intuitive idea that two compatible models represent the same belief hierarchies.

Lemma 1

(i) $\delta_i^k(\omega) = \delta_i^k(\omega') \Rightarrow \delta_i^l(\omega) = \delta_i^l(\omega')$; $\forall\, k \geq l > 0,\ \omega, \omega' \in \Omega, i \in N$.

(ii) $\zeta(\omega) \times \delta_{-i}^1(\omega) \times \cdots \times \delta_{-i}^k(\omega)$ *is in the support of* $\delta_i^{k+1}(\omega)$, $\forall\, k > 0,\ \omega \in \Omega, i \in N$.

Proof This is straightforward by definitions. □

Lemma 2 *If $\Omega \sim \tilde{\Omega}$; then there exist belief-preserving correspondences $h_i : \Omega \rightrightarrows \tilde{\Omega}$ with the following properties:*

$$\begin{aligned} &\text{(i)} \quad if\ \omega' \notin \Pi_i(\omega),\ h_i(\omega) \cap h_i(\omega') = \emptyset; \\ &\text{(ii)} \quad if\ \omega' \in \Pi_i(\omega),\ h_i(\omega) = h_i(\omega'); \\ &\text{(iii)} \quad if\ \tilde{\delta}_i(\tilde{\omega}) = \delta_i(\omega),\ \tilde{\omega} \in \tilde{\Omega},\ \omega \in \Omega;\ then\ \tilde{\omega} \in h_i(\omega). \end{aligned}$$

Furthermore, h_i's are uniquely determined by the three conditions.

Proof Suppose, to the contrary, (i) is not possible, i.e. $\forall$ belief-preserving correspondence h_i, $\exists\, \omega' \notin \Pi_i(\omega)$, such that $h_i(\omega) \cap h_i(\omega') \neq \emptyset$. We then have

$$\delta_i(\omega) = \tilde{\delta}_i(h_i(\omega)) = \tilde{\delta}_i(h_i(\omega')) = \delta_i(\omega'),$$

which in turn implies that $\omega' \in \Pi_i(\omega)$ by the fact that $\langle \Omega, (\Pi_i, P_i)_{i \in N}, \zeta \rangle$ is reduced – a contradiction.

For (ii), we can take $\hat{h}_i(\omega) = \bigcup_{\omega' \in \Pi_i(\omega)} h_i(\omega')$; $\hat{h}_i$'s satisfy both (i) and (ii).

Given (i) and (ii), suppose, to the contrary, (iii) does not hold; then $\exists \hat{\omega} \in \Omega$ such that $\tilde{\omega} \notin h_i(\omega)$ but $\tilde{\omega} \in h_i(\hat{\omega})$ by the full-range property of h_i. We thus have $\delta_i(\hat{\omega}) = \tilde{\delta}_i(\tilde{\omega}) = \delta_i(\omega)$, implying that $\hat{\omega} \in \Pi_i(\omega)$. But (ii) implies that $h_i(\omega) = h_i(\hat{\omega})$, a contradiction.

Uniqueness is immediate from (iii). $\square$

The three conditions are not independent: (iii) implies (i) and (ii); we list all three for future citations. Property (iii) requires that $h_i(\omega)$ exhausts all the $\tilde{\omega}$'s, in $\tilde{\Omega}$, representing the same belief hierarchy as ω. Thus $h_i(\omega) \subset \tilde{\Pi}_i$. We will assume h_i's satisfy the three conditions in later arguments. The following lemma further strengthens the relationship between compatible models.

Lemma 3 *If $\Omega \sim \tilde{\Omega}$; then for any $\tilde{\omega} \in \tilde{\Omega}$, there exists a unique $\omega \in \Omega$ such that*

$$\text{(i)} \quad \delta_i(\omega) = \tilde{\delta}_i^*(\tilde{\omega}),\ \forall i \in N;$$
$$\text{(ii)} \quad \zeta(\omega) = \tilde{\zeta}(\tilde{\omega}).$$

Proof *Existence.* Fix $i \in N$ and $\tilde{\omega} \in \tilde{\Omega}$. We can find $\tilde{\pi}_i \in \tilde{\Pi}_i$ and $\pi_i \in \Pi_i$ such that $\tilde{\omega} \in \tilde{\pi}_i \subset h_i(\pi_i)$, where h_i's are defined as in Lemma 2. For any $\omega' \in \pi_i$, $\delta_i(\omega') = \tilde{\delta}_i(\tilde{\omega})$, or equivalently,

$$\delta_i^k(\omega') = \tilde{\delta}_i^k(\tilde{\omega}),\ \forall\, \omega' \in \pi_i,\ k \geq 1. \tag{6}$$

Define

$$\begin{aligned} B_0 &= \{\omega \in \pi_i : \zeta(\omega) = \tilde{\zeta}(\tilde{\omega})\}; \\ B_k &= \{\omega \in \pi_i : \zeta(\omega) = \tilde{\zeta}(\tilde{\omega}), \delta_{-i}^l(\omega) = \tilde{\delta}_{-i}^l(\tilde{\omega}), 2 \leq l \leq k\},\ k \geq 1. \end{aligned}$$

Notice that $B_k \neq \emptyset$ for $k \geq 0$, by (6) and Lemma 1; Since $B_{k+1} \subset B_k$ by definition, we conclude that $\bigcap_{k \geq 0} B_k \neq \emptyset$.[21] Take any $\omega \in \bigcap_{k \geq 1} B_k$; we then have $\zeta(\omega) = \tilde{\zeta}(\tilde{\omega})$ and $\delta_{-i}^*(\omega) = \tilde{\delta}_{-i}^*(\tilde{\omega})$. Notice that $\delta_i^*(\omega) = \tilde{\delta}_i^*(\tilde{\omega})$ since $\omega \in \pi_i$. ω satisfies (i) and (ii).

Uniqueness. Suppose to the contrary, for a state $\tilde{\omega} \in \tilde{\Omega}$ there are two states $\omega \neq \omega'$ in Ω, satisfying conditions (i) and (ii); then $\delta_i(\omega) = \tilde{\delta}_i(\tilde{\omega}) = \delta_i(\omega')\ \forall i \in N$, implying that $\omega' \in \Pi_i(\omega)\ \forall i \in N$, by reducedness; furthermore, $(\zeta(\omega), (\delta_i(\omega))_{i \in N}) = (\zeta(\omega'), (\delta_i(\omega'))_{i \in N})$, contradicting Assumption

[21] Note that compactness of sets π_i or regularity of δ_i^k can guarantee this conclusion in the general case.

1. $\square$

By virtue of Lemma 3, there exists a function $f : \tilde{\Omega} \to \Omega$ as follows: $f(\tilde{\omega}) = \omega$ if $(\zeta(\omega), (\delta_i(\omega))_{i \in N}) = (\tilde{\zeta}(\tilde{\omega}), (\tilde{P}_i{}^*(\tilde{\omega}))_{i \in N})$. We call f the *canonical projection* of $\tilde{\Omega}$ to Ω. Lemma 3 strengthens Lemma 2 in the sense that $f(\tilde{\pi}_i) \subset \pi_i$ whenever $\tilde{\pi}_i \subset h_i(\pi_i)$. The next lemma shows the inverse.

Lemma 4 *Fix $i \in N$. Suppose $\tilde{\pi}_i \subset h_i(\pi_i)$, for some $\pi_i \in \Pi_i$ and $\tilde{\pi}_i \in \tilde{\Pi}_i$, where h_i's are defined as in Lemma 2. Then, $\pi_i = f(\tilde{\pi}_i)$ for the canonical projection f.*

Proof As noted above, it suffices to show $\pi_i \subset f(\tilde{\pi}_i)$. Take $\omega \in \pi_i$; then $\delta_i(\omega) = \tilde{\delta}_i(\tilde{\omega})\ \forall \tilde{\omega} \in \tilde{\pi}_i$; that is, $\delta_i^k(\omega) = \tilde{\delta}_i^k(\tilde{\omega}), \forall k \geq 1$. We can find some $\tilde{\omega}' \in \tilde{\pi}_i$, by the same argument as in the proof of Lemma 3, such that $\delta_{-i}^k(\omega) = \tilde{\delta}_{-i}^k(\tilde{\omega}'), \forall k \geq 1$, and $\zeta(\omega) = \tilde{\zeta}(\tilde{\omega}')$. Thus $(\zeta(\omega), (\delta_i(\omega))_{i \in N}) = (\tilde{\zeta}(\tilde{\omega}'), (\tilde{P}_i{}^*(\tilde{\omega}'))_{i \in N})$. We conclude that $\omega = f(\tilde{\omega}') \in f(\tilde{\pi}_i)$ by the definition of f. $\square$

Lemma 4 together with the full range property of h_i, yields the following properties of f.

Lemma 5 $f(\tilde{\Omega}) = \Omega$.

Lemma 6 $f^{-1}(\pi_i) = h_i(\pi_i)$.

In addition, our next result says that f preserves posterior beliefs.

Lemma 7 *Fix $i \in N$, let π_i and $\tilde{\pi}_i$ be arbitrary elements in Π_i and $\tilde{\Pi}_i$, respectively, such that $\tilde{\pi}_i \subset h_i(\pi_i)$; Fix $\omega \in \pi_i$. Then $P_i[\omega|\pi_i] = \tilde{P}_i[f^{-1}(\omega)|\tilde{\pi}_i]$.*

Proof Define

$$\begin{aligned}
B_0 &= \{\tilde{\omega} \in \tilde{\pi}_i : \zeta(\omega) = \tilde{\zeta}(\tilde{\omega})\}; \\
B_k &= \{\tilde{\omega} \in \tilde{\pi}_i : \zeta(\omega) = \tilde{\zeta}(\tilde{\omega}), \delta_{-i}^l(\omega) = \tilde{\delta}_{-i}^l(\tilde{\omega}), 1 \leq l \leq k\},\ k \geq 1. \\
D_0 &= \{\omega' \in \pi_i : \zeta(\omega) = \zeta(\omega')\}; \\
D_k &= \{\omega' \in \pi_i : \zeta(\omega) = \zeta(\omega'), \delta_{-i}^l(\omega) = \delta_{-i}^l(\omega'), 1 \leq l \leq k\},\ k \geq 1.
\end{aligned}$$

B_k and D_k are decreasing and nonempty as shown before. Then,

$$\begin{aligned}
\tilde{P}_i[f^{-1}(\omega)|\tilde{\pi}_i] &= \tilde{P}_i[\{\tilde{\zeta}(\tilde{\omega}) = \zeta(\omega), \tilde{\delta}_j(\tilde{\omega}) = \delta_j(\omega), j \in N\}|\tilde{\pi}_i] \quad (7) \\
&= \tilde{P}_i[\{\tilde{\zeta}(\tilde{\omega}) = \zeta(\omega), \tilde{\delta}_j(\tilde{\omega}) = \delta_j(\omega), j \in N, j \neq i\}|\tilde{\pi}_i] \quad (8)
\end{aligned}$$

$$
\begin{aligned}
&= \tilde{P}_i[\{\tilde{\zeta}(\tilde{\omega}) = \zeta(\omega), \tilde{\delta}^k_{-i}(\tilde{\omega}) = \delta^k_{-i}(\omega), k \geq 1\}|\tilde{\pi}_i] && (9)\\
&= \tilde{P}_i[\lim_{k\to+\infty} B_k|\tilde{\pi}_i] && (10)\\
&= \lim_{k\to+\infty} \tilde{P}_i[B_k|\tilde{\pi}_i] && (11)\\
&= \lim_{k\to+\infty} \tilde{P}_i[\{\tilde{\zeta}(\tilde{\omega}) = \zeta(\omega), \tilde{\delta}^l_{-i}(\tilde{\omega}) = \delta^l_{-i}(\omega), 1 \leq l \leq k\}|\tilde{\pi}_i] && (12)\\
&= \lim_{k\to+\infty} \tilde{\delta}^{k+1}_i(\tilde{\pi}_i)[(\zeta(\omega), \delta^1_{-i}(\omega), \cdots, \delta^k_{-i}(\omega)] && (13)\\
&= \lim_{k\to+\infty} \delta^{k+1}_i(\pi_i)[(\zeta(\omega), \delta^1_{-i}(\omega), \cdots, \delta^k_{-i}(\omega)] && (14)\\
&= \lim_{k\to+\infty} P_i[\{\zeta(\omega') = \zeta(\omega), \delta^l_{-i}(\omega') = \delta^l_{-i}(\omega), 1 \leq l \leq k\}|\pi_i] && (15)\\
&= \lim_{k\to+\infty} P_i[D_k|\pi_i] && (16)\\
&= P_i[\lim_{k\to+\infty} D_k|\pi_i] && (17)\\
&= P_i[\{\zeta(\omega') = \zeta(\omega), \delta^l_{-i}(\omega') = \delta^l_{-i}(\omega), k \geq 1|\pi_i] && (18)\\
&= P_i[\{\zeta(\omega') = \zeta(\omega), \delta_j(\omega') = \delta_j(\omega), j \in N, j \neq i\}|\pi_i] && (19)\\
&= P_i[\{\zeta(\omega') = \zeta(\omega), \delta_j(\omega') = \delta_j(\omega), j \in N\}|\pi_i] && (20)\\
&= P_i[\omega|\pi_i] && (21)
\end{aligned}
$$

In the equations above, (7) is by the definition of f; (8) is because $\tilde{\delta}_i(\tilde{\omega}) = \tilde{\delta}_i(\omega), \forall \tilde{\omega} \in \tilde{\pi}_i$; (9) is by the definition of $\tilde{\delta}_i$; (10) is by the definition of B_k; (11) is by the property of probability measures; (12) is by the definition of B_k; (13) is by the definition of $\tilde{\delta}^{k+1}_i$; (14) is by the equivalence of $\tilde{\delta}_i$ and δ_i; (15) is by the definition of δ^{k+1}_i; (16) is by the definition of D_k; (17) is by the property of probability measures; (18) is by the definition of D_k; (19) is by the definition of δ_j; (20) is because $\delta_i(\omega') = \delta_i(\omega), \forall \omega' \in \pi_i$; (21) is by reducedness. $\square$

We are now ready for the proof of the characterization theorem:

Proof of Theorem 1

Sufficiency. Let $h_i(\omega) = \omega \times C^\omega$ for all $i \in N$; then $h_i(\Omega) = \tilde{\Omega}$. From the remark following Definition 1, we need to show that $\delta_i(\omega) = \tilde{\delta}_i(\tilde{\omega})$ holds for all $\tilde{\omega} \in h_i(\omega)$ and $\omega \in \Omega$. We begin with the first order beliefs: for any $\omega \in \Omega$, $c \in C^\omega$ and $s \in S$,

$$
\begin{aligned}
\tilde{\delta}^1_i(\omega \times c)(s) &= \tilde{P}_i(\tilde{\zeta}(\omega' \times c') = s|\, \tilde{\Pi}_i(\omega \times c)) && (22)\\
&= \sum_{\omega' \in \zeta^{-1}(s)} \sum_{c' \in C^{\omega'}} \tilde{P}_i(\omega' \times c'|\, \tilde{\Pi}_i(\omega \times c)) && (23)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{\omega' \in \zeta^{-1}(s)} \sum_{c' \in C^{\omega'}} P_i(\omega' | \, \omega) \cdot q_i^{\omega'}(c' | \, c_i) & (24)\\
&= \sum_{\omega' \in \zeta^{-1}(s)} P_i(\omega' | \, \omega) \cdot \sum_{c' \in C^{\omega'}} q_i^{\omega'}(c' | \, c_i) & (25)\\
&= \sum_{\omega' \in \zeta^{-1}(s)} P_i(\omega' | \, \omega) & (26)\\
&= P_i(\zeta(\omega') = s | \, \Pi_i(\omega)) & (27)\\
&= \delta_i^1(\omega)(s). & (28)
\end{aligned}
$$

In the equations above, (22) is by the definition of $\tilde{\delta}_i^1$; (23) is by condition (iv) in Theorem 1; (24) is by the definition of $\tilde{P}_i$; (26) is by the definition of $q_i^{\omega'}$. (28) is by the definition of δ_i^1. We then have,

$$
\{\phi_i^1 : \phi_i^1 = \delta_i^1(\omega) \text{ for some } \omega \in \Omega\} = \Phi_i^1 = \tilde{\Phi}_i^1 = \{\tilde{\phi}_i^1 : \tilde{\phi}_i^1 = \tilde{\delta}_i^1(\tilde{\omega}) \text{ for some } \tilde{\omega} \in \tilde{\Omega}\};
$$

Continuing with the second order beliefs, for all $s \in S$ and $\phi_{-i}^1 \in \Phi_{-i}^1 = \tilde{\Phi}_{-i}^1$,

$$
\begin{aligned}
\tilde{\delta}_i^2(\omega \times c)(s \times \phi_{-i}^1) &= \tilde{P}_i(\tilde{\zeta}(\omega' \times c') = s, \tilde{\delta}_{-i}^1(\omega' \times c') = \phi_{-i}^1 | \, \tilde{\Pi}_i(\omega \times c))\\
&= \sum_{\omega' \in \zeta^{-1}(s) \cap (\delta_{-i}^1)^{-1}(\phi_{-i}^1)} \sum_{c' \in C^{\omega'}} \tilde{P}_i(\omega' \times c' | \, \tilde{\Pi}_i(\omega \times c))\\
&= \sum_{\omega' \in \zeta^{-1}(s) \cap (\delta_{-i}^1)^{-1}(\phi_{-i}^1)} \sum_{c' \in C^{\omega'}} P_i(\omega' | \, \omega) \cdot q_i^{\omega'}(c' | \, c_i)\\
&= \sum_{\omega' \in \zeta^{-1}(s) \cap (\delta_{-i}^1)^{-1}(\phi_{-i}^1)} P_i(\omega' | \, \omega) \cdot \sum_{c' \in C^{\omega'}} q_i^{\omega'}(c' | \, c_i)\\
&= \sum_{\omega' \in \zeta^{-1}(s) \cap (\delta_{-i}^1)^{-1}(\phi_{-i}^1)} P_i(\omega' | \, \omega)\\
&= P_i(\zeta(\omega') = s, \ \delta_i^1(\omega' \times c') = \phi_{-i}^1 | \, \Pi_i(\omega))\\
&= \delta_i^2(\omega)(s \times \phi_{-i}^1).
\end{aligned}
$$

Inductively, we can show that $\tilde{\delta}_i^k(\omega \times c) = \delta_i^k(\omega), \ \forall c \in C^{\omega}, k \geq 2$, and hence $\delta_i(\omega) = \tilde{\delta}_i(\tilde{\omega})$ $\forall \tilde{\omega} \in h_i(\omega)$ and $\omega \in \Omega, \ i = 1, \cdots, n$, as required.

Necessity. Given $\Omega \sim \tilde{\Omega}$, we want to construct a correlating mechanism $\langle (C^{\omega})_{\omega}, (q_i^{\omega})_{\omega \in \Omega, i \in N} \rangle$ through which $\tilde{\Omega}$ can be derived from Ω.

Step 1: Construction of C^ω.

Suppose $\tilde{\Pi}_i = \{\tilde{\pi}_i^1, \tilde{\pi}_i^2, \cdots, \tilde{\pi}_i^{n_i}\}$. Let $C_i = \{1, 2, \cdots, n_i\}$. For $\forall i \in N$, define a mapping $k_i : \tilde{\Omega} \to C_i$ by $k_i(\tilde{\omega}) = k$, if $\tilde{\omega} \in \tilde{\pi}_i^k$. Clearly, k_i is a well defined $\tilde{\Pi}_i$-measurable mapping with $\tilde{\Pi}_i(\tilde{\omega}) = \tilde{\pi}_i^{k_i(\tilde{\omega})}$. Define, for any $\omega \in \Omega$,

$$C^\omega = \{(k_1(\tilde{\omega}), k_2(\tilde{\omega}), \cdots, k_n(\tilde{\omega})) : \tilde{\omega} \in f^{-1}(\omega)\}$$

where f is the canonical projection from $\tilde{\Omega}$ onto Ω. C^ω, with a typical element c^ω, is well-defined and non-empty by Lemma 5. Notice that $C^\omega|_{C_i} = C^{\omega'}|_{C_i}$ if $\omega' \in \Pi_i(\omega)$ by Lemma 4 and Lemma 6. So C^ω such defined satisfies (a) in Definition 3.

Step 2: Construction of q_i^ω.

Define $\rho : \tilde{\omega} \mapsto (f(\tilde{\omega}), k_1(\tilde{\omega}), k_2(\tilde{\omega}), \cdots, k_n(\tilde{\omega}))$. We claim that $\rho : \tilde{\Omega} \to \rho(\tilde{\Omega})$ is one-to-one. To see this, suppose there exist $\tilde{\omega}, \tilde{\omega}' \in \tilde{\Omega}, \tilde{\omega} \neq \tilde{\omega}'$, such that $\rho(\tilde{\omega}) = \rho(\tilde{\omega}')$; then $f(\tilde{\omega}) = f(\tilde{\omega}')$ and $k_i(\tilde{\omega}) = k_i(\tilde{\omega}'), \forall i \in N$. The former implies that $\zeta(\tilde{\omega}) = \zeta(\tilde{\omega}')$; the latter implies that $\tilde{\omega} \in \tilde{\Pi}_i(\tilde{\omega}'), \forall i \in N$. This violates Assumption 1.

Define, for $\omega \in \Omega$ and $c^\omega \in C^\omega$,

$$q_i^\omega(c^\omega) = \frac{\sum_{c_i \in C_i^\omega} \tilde{P}_i[\rho^{-1}(\omega \times c^\omega) | \tilde{\pi}_i^{c_i}]}{|C_i^\omega| P_i[\omega | \Pi_i(\omega)]}.$$

We claim that q_i^ω satisfies (b) in Definition 3.

$$\begin{aligned} q_i^\omega(C^\omega) &= \frac{\sum_{c_i \in C_i^\omega} \tilde{P}_i[\rho^{-1}(\omega \times C^\omega) | \tilde{\pi}_i^{c_i}]}{|C_i^\omega| P_i[\omega | \Pi_i(\omega)]} \\ &= \frac{\sum_{c_i \in C_i^\omega} \tilde{P}_i[f^{-1}(\omega) | \tilde{\pi}_i^{c_i}]}{|C_i^\omega| P_i[\omega | \Pi_i(\omega)]} \\ &= 1 \end{aligned}$$

In the last line, we used Lemma 7 and the fact $\tilde{\pi}_i^{c_i} \subset h_i(\Pi_i(\omega))$ by the definition of C^ω. The additivity of q_i^ω is straightforward by noting that ρ is one-to-one. Thus q_i^ω is a probability measure on C^ω.

For a fixed $c_i^\circ \in C_i^\omega$, consider

$$
\begin{aligned}
\sum_{\{c^\omega \in C^\omega:\ c_i^\omega = c_i^\circ\}} q_i^\omega(c) &= \sum_{\{c^\omega \in C^\omega:\ c_i^\omega = c_i^\circ\}} \frac{\sum_{c_i \in C_i^\omega} \tilde{P}_i[\rho^{-1}(\omega \times c^\omega)|\tilde{\pi}_i^{c_i}]}{|C_i^\omega| P_i[\omega|\ \Pi_i(\omega)]} \\
&= \frac{\sum_{\{c^\omega \in C^\omega:\ c_i^\omega = c_i^\circ\}} \tilde{P}_i[\rho^{-1}(\omega \times c^\omega)|\tilde{\pi}_i^{c_i^\circ}]}{|C_i^\omega| P_i[\omega|\ \Pi_i(\omega)]} \\
&= \frac{\tilde{P}_i[f^{-1}(\omega) \cap \tilde{\pi}_i^{c_i^\circ}]|\tilde{\pi}_i^{c_i^\circ}]}{|C_i^\omega| P_i[\omega|\ \Pi_i(\omega)]} \\
&= \frac{1}{|C_i^\omega|}.
\end{aligned}
$$

As shown in step 1, $C^\omega|_{C_i} = C^{\omega'}|_{C_i}$ if $\omega' \in \Pi_i(\omega)$. We conclude that condition (b) in Definition 3 is satisfied.

We have constructed a state-dependent correlating mechanism $\langle (C^\omega)_{\omega \in \Omega}, (q_i^\omega)_{\omega \in \Omega, i \in N} \rangle$ in the two steps. We now show that $\tilde{\Omega}$ can be derived from Ω through the correlating mechanism as stated in Theorem 1.

Step 3 (i)

$$
\begin{aligned}
C^\omega &= \{(k_1(\tilde{\omega}), k_2(\tilde{\omega}), \cdots, k_n(\tilde{\omega})) : \tilde{\omega} \in f^{-1}(\omega)\} \\
&= \{(c_1, c_2, \cdots, c_n) : \exists \tilde{\omega} \in \tilde{\Omega},\ \text{s.t. } \rho(\tilde{\omega}) = (\omega, c_1, c_2, \cdots, c_n)\}
\end{aligned}
$$

Consequently, $\rho(\tilde{\Omega}) = \{\omega \times c^\omega\}_{c^\omega \in C^\omega, \omega \in \Omega}$ by Lemma 5; that is, $\tilde{\Omega}$ is isomorphic to $\{\omega \times c^\omega\}_{c^\omega \in C^\omega, \omega \in \Omega}$. ρ, as an imbedding of $\tilde{\Omega}$ in the space of outcomes of the randomizing device, is the desired isomorphism in Theorem 1.

Step 3 (ii)

For an arbitrary $\tilde{\omega} \in \tilde{\Omega}$,

$$
\begin{aligned}
\rho\left(\tilde{\Pi}_i(\tilde{\omega})\right) &= \rho\left(\tilde{\pi}_i^{k_i(\tilde{\omega})}\right) && (29) \\
&= \left\{\rho(\tilde{\omega}') : \tilde{\omega}' \in \tilde{\pi}_i^{k_i(\tilde{\omega})}\right\} && (30) \\
&= \left\{\left(f(\tilde{\omega}'), k_1(\tilde{\omega}'), \cdots, k_i(\tilde{\omega}'), \cdots, k_n(\tilde{\omega}')\right) : \tilde{\omega}' \in \tilde{\pi}_i^{k_i(\tilde{\omega})}\right\} && (31) \\
&= \left\{\left(f(\tilde{\omega}'), k_1(\tilde{\omega}'), \cdots, k_i(\tilde{\omega}), \cdots, k_n(\tilde{\omega}')\right) : \tilde{\omega}' \in \tilde{\pi}_i^{k_i(\tilde{\omega})}\right\} && (32) \\
&= \left\{\omega' \times c^{\omega'} : \omega' \in \Pi_i(f(\tilde{\omega})), c^{\omega'} \in C^{\omega'}, c_i^{\omega'} = k_i(\tilde{\omega})\right\} && (33)
\end{aligned}
$$

(31) is by the definition of ρ; (32) is due to the measurability of k_i; (33) is by Lemma 4 and definition of $C^{\omega'}$.

We write $\omega = f(\tilde{\omega})$ and $c^{\omega} = (k_1(\tilde{\omega}), \cdots, k_i(\tilde{\omega}), \cdots, k_n(\tilde{\omega}))$. Then $\rho(\tilde{\omega}) = \omega \times c^{\omega}$. Identifying $\rho\left(\tilde{\Pi}_i(\tilde{\omega})\right)$ with $\tilde{\Pi}_i(\rho(\tilde{\omega}))$, we have $\tilde{\Pi}_i(\omega \times c^{\omega}) = \rho\left(\tilde{\Pi}_i(\tilde{\omega})\right) = \{\omega' \times c^{\omega'} : \omega' \in \Pi_i(\omega) \text{ and } c^{\omega'} \in C^{\omega'} \text{ with } c_i^{\omega'} = c_i^{\omega}\}$, which is the desired form in the theorem.

Step 3 (iii)

With the identification, for $\omega' \in \Pi_i(\omega)$ we have, $q_i^{\omega'}(c^{\omega'}|\ c_i^{\omega}) = \frac{q_i^{\omega'}(c^{\omega'})}{\sum_{\{c \in C^{\omega'}:\ c_i = c_i^{\omega}\}} q_i^{\omega'}(c)}$ and $q_i^{\omega}(c^{\omega}) = \frac{\sum_{c_i \in C_i^{\omega}} \tilde{P}_i[\rho^{-1}(\omega \times c^{\omega})|\tilde{\pi}_i^{c_i}]}{|C_i^{\omega}| P_i[\omega|\ \Pi_i(\omega)]}$. Substitution yields the desired result.

Step 3 (iv)

With the identification, we have $\tilde{\zeta}(\omega \times c^{\omega}) = \zeta(\omega), \forall \omega \times c^{\omega} \in \tilde{\Omega}$ immediately by the definition of ρ. $\square$

Appendix B: Proof of Proposition 1

Proof (i) We apply a well-known result[22]: there is no common prior on Ω if and only if there exist real functions $\lambda_i : \Omega \to \mathbb{R}$, $i = 1, 2, \cdots, n$, such that $\sum_{i=1}^{n} \lambda_i \equiv 0$ and $E_i[\lambda_i|\Pi_i] > 0, \forall i \in N$. Define functions $\tilde{\lambda}_i : \tilde{\Omega} \to \mathbb{R}$ as $\tilde{\lambda}_i(\omega \times c) = \lambda_i(\omega)$ for all $\omega \times c \in \tilde{\Omega}$ according to Theorem 1; then $\sum_{i=1}^{n} \tilde{\lambda}_i \equiv 0$ by definition. We have, by Theorem 1 and the definition of $\tilde{\lambda}_i$,

$$\begin{aligned}
E_i[\tilde{\lambda}_i|\tilde{\Pi}_i(\omega \times c)] &= \sum_{\omega' \times c' \in \tilde{\Pi}_i(\omega \times c)} \tilde{\lambda}_i(\omega' \times c')\tilde{P}_i[\omega' \times c'|\tilde{\Pi}_i(\omega \times c')] \\
&= \sum_{\omega' \in \Pi_i(\omega)} \sum_{\{c' \in C^{\omega'}:\ c'_i = c_i\}} \lambda_i(\omega')P_i(\omega'|\omega)q_i^{\omega'}(c'|c_i) \\
&= \sum_{\omega' \in \Pi_i(\omega)} \lambda_i(\omega')P_i(\omega'|\omega) \sum_{\{c' \in C^{\omega'}:\ c'_i = c_i\}} q_i^{\omega'}(c'|c_i) \\
&= \sum_{\omega' \in \Pi_i(\omega)} \lambda_i(\omega')P_i(\omega'|\omega) \\
&> 0.
\end{aligned}$$

So there is no common prior on $\tilde{\Omega}$, as claimed in (i).

(ii) Suppose there is a common prior $\mathbf{P}$ on Ω. Define $\tilde{\mathbf{P}}(\omega \times c) = \mathbf{P}(\omega)q^{\omega}(c)$ for $\omega \times c \in \tilde{\Omega}$.

[22] e.g. Feinberg [11].

Notice that the subscript of q_i^ω has been suppressed by assumption. It is routine to check that $\tilde{\mathbf{P}}$ is a probability measure on $\tilde{\Omega}$. We will show that $\tilde{\mathbf{P}}$ is a common prior on $\tilde{\Omega}$. By definition,

$$\begin{aligned}
\tilde{\mathbf{P}}(\tilde{\Pi}_i(\omega \times c)) &= \sum_{\omega' \times c' \in \tilde{\Pi}_i(\omega \times c)} \tilde{\mathbf{P}}(\omega' \times c') \\
&= \sum_{\omega' \in \Pi_i(\omega)} \sum_{\{c' \in C^{\omega'}:\, c'_i = c_i\}} \mathbf{P}(\omega') q^{\omega'}(c') \\
&= \sum_{\omega' \in \Pi_i(\omega)} \mathbf{P}(\omega') \sum_{\{c' \in C^{\omega'}:\, c'_i = c_i\}} q^{\omega'}(c') \\
&= \mathbf{P}(\Pi_i(\omega)) \sum_{\{c' \in C^{\omega'}:\, c'_i = c_i\}} q^{\omega'}(c')
\end{aligned}$$

We then have

$$\begin{aligned}
\tilde{\mathbf{P}}(\omega' \times c' | \tilde{\Pi}_i(\omega \times c)) &= \frac{\tilde{\mathbf{P}}(\omega' \times c')}{\tilde{\mathbf{P}}(\tilde{\Pi}_i(\omega \times c))} \\
&= \frac{\mathbf{P}(\omega') q^{\omega'}(c')}{\mathbf{P}(\Pi_i(\omega)) \sum_{\{c' \in C^{\omega'}:\, c'_i = c_i\}} q^{\omega'}(c')} \\
&= \mathbf{P}(\omega' | \Pi_i(\omega)) q^{\omega'}(c' | c_i) \\
&= P(\omega' | \Pi_i(\omega)) q^{\omega'}(c' | c_i) \\
&= \tilde{P}_i(\omega' \times c' | \tilde{\Pi}_i(\omega \times c))
\end{aligned}$$

as required by (ii). □

Appendix C: Proof of Proposition 5

We first show the following simple lemmas; they are straightforward in a complete information setting.

Lemma 8 one-shot mechanism vs. sequential mechanism

Let Ω be an arbitrary partition model (possibly non-reduced) and a state-dependent correlating mechanism $\langle (C^\omega)_{\omega \in \Omega}, (q_i^\omega)_{\omega \in \Omega, i \in N} \rangle$ on it; then the partition model generated by this mechanism, $\tilde{\Omega}$ is fundamentally compatible with Ω, where $\tilde{\Omega} = \{\omega \times c^\omega\}_{\omega \in \Omega,\, c^\omega \in C^\omega}$, $\tilde{\Pi}_i(\omega \times c^\omega) = I(\omega \times c^\omega)$ and $\tilde{P}_i(\omega' \times c^{\omega'} | \, \omega \times c^\omega) = P_i(\omega' | \, \omega) \cdot q_i^{\omega'}(c^{\omega'} | \, c_i^\omega)$.

Remark This lemma was proved in the sufficiency part of Theorem 1, where we did not use the reducedness assumption. Let us interpret the lemma in the following way. Let Ω and Ω_2 be the models stated in the Proposition (see Figure 7 in the text.) Theorem 1 said Ω_2 is generated from Ω by a state-dependent correlating mechanism $\mathcal{D}_2$. Let Ω_3 be a model generated from Ω_2 by a mechanism $\mathcal{D}_3$, then $\Omega_3 \sim \Omega$ by this lemma and hence can be generated from Ω by another mechanism $\mathcal{D}$ by Theorem 1 again. So, running $\mathcal{D}_2$ and $\mathcal{D}_3$ sequentially is equivalent to running $\mathcal{D}$ alone.

Corollary 2 *Let Ω and Ω_1 be partition models in the proposition. Then $\mathcal{C}(\Omega_1) \subset \mathcal{C}(\Omega)$.*

To make our further life easier, we suffer a little bit now by introducing more notations. We denote

$$\Omega := \langle \Omega, (\Pi_i, P_i)_{i\in N}, \zeta \rangle;$$

$$\Omega_1 := \langle \hat{\Omega}, (\hat{\Pi}_i, \hat{P}_i)_{i\in N}, \hat{\zeta} \rangle \text{ and } \mathcal{D}_1 = \langle (\hat{C}^{\omega})_{\omega\in\Omega}, (\hat{q}_i^{\omega})_{\omega\in\Omega, i\in N} \rangle;$$

$$\Omega_2 = \langle \tilde{\Omega}, (\tilde{\Pi}_i, \tilde{P}_i)_{i\in N}, \tilde{\zeta} \rangle \text{ and } \mathcal{D}_2 = \langle (\tilde{C}^{\omega})_{\omega\in\Omega}, (\tilde{q}_i^{\tilde{\omega}})_{\tilde{\omega}\in\tilde{\Omega}, i\in N} \rangle;$$

Define a mechanism $\mathcal{D}_3 = \langle (\hat{\tilde{C}}^{\tilde{\omega}})_{\tilde{\omega}\in\tilde{\Omega}}, (\hat{\tilde{q}}_i^{\tilde{\omega}})_{\tilde{\omega}\in\tilde{\Omega}, i\in N} \rangle$ on Ω_2 by the following:

$$\hat{\tilde{C}}^{\tilde{\omega}} = \hat{C}^{\omega}, \ \forall \tilde{\omega} = \omega \times c \in \tilde{\Omega}; \tag{34}$$

$$\hat{\tilde{q}}_i^{\tilde{\omega}} = \hat{q}_i^{\omega}, \ \forall \tilde{\omega} = \omega \times c \in \tilde{\Omega}. \tag{35}$$

$\tilde{\omega} = \omega \times c$ is due to the characterization of Theorem 1. It is straightforward to check that $\mathcal{D}_3$ is well defined. Denote $\Omega_3 := \langle \hat{\tilde{\Omega}}, (\hat{\tilde{\Pi}}_i, \hat{\tilde{P}}_i)_{i\in N}, \hat{\tilde{\zeta}} \rangle$ be the model generated from Ω_2 by the mechanism $\mathcal{D}_3$.

Lemma 9 common ignorance: $\mathcal{N}(\Omega_1) \subset \mathcal{N}(\Omega_3) \subset \mathcal{C}(\Omega_2)$.

The intuition behind the lemma is the following: Ω_3 is obtained by operating two consecutive devices $\mathcal{D}_2$ and $\mathcal{D}_3$ (see Figure 6 in the text); but when coordinating their actions, players can ignore the correlating device $\mathcal{D}_2$ if all others do so. We can design $\mathcal{D}_3$ in a way such that it "mimics" $\mathcal{D}_1$; So, the equilibria obtained by using coordinating device $\mathcal{D}_1$ are still obtained. We term this property *common ignorance*.

Proof From Theorem 1, a generic point in $\hat{\Omega}$ can be written as $\omega \times \hat{c}$ and

$$\hat{\Pi}_i(\omega \times \hat{c}) = \{\omega' \times \hat{c}'\}_{\hat{c}'\in\hat{C}^{\omega'},\, \hat{c}'_i=\hat{c}_i,\, \omega'\in\Pi_i(\omega)}. \tag{36}$$

Let $(\hat{\alpha}_1, \cdots, \hat{\alpha}_n)$ be a Bayesian-Nash equilibrium for the game associated with Ω_1; then by definition,

$$\text{(i)} \quad \hat{\alpha}_i : \hat{\Omega} \to A_i \text{ is } \hat{\Pi}_i\text{-measurable}, \forall i \in N. \tag{37}$$

$$\text{(ii)} \quad E_i[g_i(\hat{\alpha}_i, \hat{\alpha}_{-i}, \omega \times \hat{c})|\hat{\Pi}_i] \geq E_i[g_i(a_i, \hat{\alpha}_{-i}, \omega \times \hat{c})|\hat{\Pi}_i] \quad \forall a_i \in A_i,\ i \in N. \tag{38}$$

A generic point in $\hat{\tilde{\Omega}}$ can be written as $\tilde{\omega} \times \hat{\tilde{c}}$, and further as $\omega \times \tilde{c} \times \hat{\tilde{c}}$; Similarly to (34),

$$\hat{\tilde{\Pi}}_i(\omega \times \tilde{c} \times \hat{\tilde{c}}) = \{(\omega' \times \tilde{c}' \times \hat{\tilde{c}}')\}_{\hat{\tilde{c}}' \in \hat{\tilde{C}}^{\omega' \times \tilde{c}'},\ \hat{\tilde{c}}_i' = \hat{\tilde{c}}_i;\ \omega' \times \tilde{c}' \in \tilde{\Pi}(\omega \times \tilde{c})} \tag{39}$$

$$= \{(\omega' \times \tilde{c}' \times \hat{\tilde{c}}')\}_{\hat{\tilde{c}}' \in \hat{\tilde{C}}^{\omega' \times \tilde{c}'},\ \hat{\tilde{c}}_i' = \hat{\tilde{c}}_i;\ \tilde{c}' \in \tilde{C}^{\omega'},\ \tilde{c}_i' = \tilde{c}_i;\ \omega' \in \Pi(\omega)} \tag{40}$$

$$= \{(\omega' \times \tilde{c}' \times \hat{\tilde{c}}')\}_{\hat{\tilde{c}}' \in \hat{C}^{\omega'},\ \hat{\tilde{c}}_i' = \hat{\tilde{c}}_i;\ \tilde{c}' \in \tilde{C}^{\omega'},\ \tilde{c}_i' = \tilde{c}_i;\ \omega' \in \Pi(\omega)} \tag{41}$$

(39) and (40) are by Theorem 1; (41) is by (35). Clearly, $\hat{\tilde{\Pi}}_i(\omega \times \tilde{c} \times \hat{\tilde{c}}) \subset \tilde{\Pi}_i(\omega \times \tilde{c} \times \hat{\tilde{c}})$. In the game associated with Ω_3, define a strategy profile $(\hat{\tilde{\alpha}}_1, \cdots, \hat{\tilde{\alpha}}_n)$, where $\hat{\tilde{\alpha}}_i : \hat{\tilde{\Omega}} \to A_i$ for all $i \in N$, as follows:

$$\hat{\tilde{\alpha}}_i(\omega \times \tilde{c} \times \hat{\tilde{c}}) := \hat{\alpha}_i(\omega \times \hat{\tilde{c}}). \tag{42}$$

$\hat{\alpha}_i$ is $\hat{\Pi}_i$-measurable, and hence $\hat{\tilde{\alpha}}_i$ is $\hat{\tilde{\Pi}}_i$-measurable by comparing (36) and (41). On the other hand,

$$E_i \quad \left[g_i(\hat{\tilde{\alpha}}_i, \hat{\tilde{\alpha}}_{-i}, \cdot)|\hat{\tilde{\Pi}}_i(\omega \times \tilde{c} \times \hat{\tilde{c}})\right] \tag{43}$$

$$= \sum_{\omega' \times \tilde{c}' \times \hat{\tilde{c}}' \in \hat{\tilde{\Pi}}_i(\omega \times \tilde{c} \times \hat{\tilde{c}})} g_i(\hat{\tilde{\alpha}}_i, \hat{\tilde{\alpha}}_{-i}, \omega' \times \tilde{c}' \times \hat{\tilde{c}}') \cdot \hat{\tilde{P}}_i[\omega' \times \tilde{c}' \times \hat{\tilde{c}}'|\omega \times \tilde{c} \times \hat{\tilde{c}}] \tag{44}$$

$$= \sum_{\omega' \times \tilde{c}' \times \hat{\tilde{c}}' \in \hat{\tilde{\Pi}}_i(\omega \times \tilde{c} \times \hat{\tilde{c}})} g_i(\hat{\tilde{\alpha}}_i, \hat{\tilde{\alpha}}_{-i}, \omega' \times \tilde{c}' \times \hat{\tilde{c}}') \cdot \tilde{P}_i[\omega' \times \tilde{c}'|\omega \times \tilde{c}] \cdot \hat{\tilde{q}}_i^{\omega' \times \tilde{c}'}(\hat{\tilde{c}}'|\hat{\tilde{c}}_i) \tag{45}$$

$$= \sum_{\omega' \times \tilde{c}' \times \hat{\tilde{c}}' \in \hat{\tilde{\Pi}}_i(\omega \times \tilde{c} \times \hat{\tilde{c}})} g_i(\hat{\tilde{\alpha}}_i, \hat{\tilde{\alpha}}_{-i}, \omega' \times \tilde{c}' \times \hat{\tilde{c}}') \cdot P_i[\omega'|\omega] \cdot \tilde{q}_i^{\omega'}(\tilde{c}'|\tilde{c}_i) \cdot \hat{\tilde{q}}_i^{\omega' \times \tilde{c}'}(\hat{\tilde{c}}'|\hat{\tilde{c}}_i) \tag{46}$$

$$= \sum_{\omega' \times \tilde{c}' \times \hat{\tilde{c}}' \in \hat{\tilde{\Pi}}_i(\omega \times \tilde{c} \times \hat{\tilde{c}})} g_i(\hat{\alpha}_i, \hat{\alpha}_{-i}, \omega' \times \hat{\tilde{c}}') \cdot P_i[\omega'|\omega] \cdot \tilde{q}_i^{\omega'}(\tilde{c}'|\tilde{c}_i) \cdot \hat{\tilde{q}}_i^{\omega' \times \tilde{c}'}(\hat{\tilde{c}}'|\hat{\tilde{c}}_i) \tag{47}$$

$$= \sum_{\omega' \times \tilde{c}' \times \hat{\tilde{c}}' \in \hat{\tilde{\Pi}}_i(\omega \times \tilde{c} \times \hat{\tilde{c}})} g_i(\hat{\alpha}_i, \hat{\alpha}_{-i}, \omega' \times \hat{\tilde{c}}') \cdot P_i[\omega'|\omega] \cdot \tilde{q}_i^{\omega'}(\tilde{c}'|\tilde{c}_i) \cdot \hat{q}_i^{\omega'}(\hat{\tilde{c}}'|\hat{\tilde{c}}_i) \tag{48}$$

$$= \sum_{\omega' \times \hat{\tilde{c}}' \in \hat{\Pi}_i(\omega \times \hat{\tilde{c}})} \sum_{\{\tilde{c}' \in \tilde{C}^{\omega'} : \tilde{c}_i' = \tilde{c}_i\}} \tilde{q}_i^{\omega'}(\tilde{c}'|\tilde{c}_i) \cdot g_i(\hat{\alpha}_i, \hat{\alpha}_{-i}, \omega' \times \hat{\tilde{c}}') \cdot P_i[\omega'|\omega] \cdot \hat{q}_i^{\omega'}(\hat{\tilde{c}}'|\hat{\tilde{c}}_i) \tag{49}$$

$$= \sum_{\omega' \times \hat{\tilde{c}}' \in \hat{\Pi}_i(\omega \times \hat{c})} g_i(\hat{\alpha}_i, \hat{\alpha}_{-i}, \omega' \times \hat{\tilde{c}}') \cdot P_i[\omega'|\omega] \cdot \hat{q}_i^{\omega'}(\hat{\tilde{c}}'|\hat{\tilde{c}}_i) \quad (50)$$

$$= \sum_{\omega' \times \hat{\tilde{c}}' \in \hat{\Pi}_i(\omega \times \hat{c})} g_i(\hat{\alpha}_i, \hat{\alpha}_{-i}, \omega' \times \hat{\tilde{c}}') \cdot \hat{P}_i[\omega' \times \hat{\tilde{c}}'|\omega \times \hat{\tilde{c}}] \quad (51)$$

$$= E_i \left[g_i(\hat{\alpha}_i, \hat{\alpha}_{-i}, \omega \times \hat{\tilde{c}}) | \hat{\Pi}_i(\omega \times \hat{\tilde{c}}) \right] \quad (52)$$

In the equations above, (45),(46) and (51) are by Theorem 1; (47) is by (42); (48) is by (35); (49) is obtained by comparing (36) and (41). Similarly, we have

$$E_i \left[g_i(a_i, \hat{\tilde{\alpha}}_{-i}, \cdot) | \hat{\tilde{\Pi}}_i(\omega \times \tilde{c} \times \hat{\tilde{c}}) \right] = E_i \left[g(a_i, \hat{\alpha}_{-i}, \cdot) | \hat{\Pi}_i(\omega \times \hat{\tilde{c}}) \right], \ \forall a_i \in A_i. \quad (53)$$

(38), (52) and (56) together yield

$$E_i \left[g_i(\hat{\tilde{\alpha}}_i, \hat{\tilde{\alpha}}_{-i}, \cdot) | \hat{\tilde{\Pi}}_i \right] \geq E_i \left[g(a_i, \hat{\tilde{\alpha}}_{-i}, \cdot) | \hat{\Pi}_i \right], \ \forall a_i \in A_i.$$

Thus the profile $(\hat{\tilde{\alpha}}_1, \cdots, \hat{\tilde{\alpha}}_n)$ forms a Bayesian-Nash equilibrium for the game associated with Ω_3. This proves the first half of the lemma; the second half follows immediately from the definition of correlated equilibrium.□

Proof of Proposition 5 By Corollary 5, $\mathcal{C}(\Omega_2) \subset \mathcal{C}(\Omega)$. From Proposition 2, $\mathcal{C}(\Omega) = \bigcup_{\Omega_1 \sim \Omega} \mathcal{N}(\Omega_1)$; by Lemma 9, $\mathcal{N}(\Omega_1) \subset \mathcal{C}(\Omega_2), \forall \Omega_1 \sim \Omega$. Thus, $\mathcal{C}(\Omega) \subset \mathcal{C}(\Omega_2)$. □

Appendix D: General Topological Case

D1 Basic Definitions and Notation

D1.1 Topology

We use compact and metrizable topology in this note; more general topology will also work (see Liu [17]). Let (X, τ) be a compact metrizable space. ΔX, the space of Borel probability measures on X, which is endowed with the weak* topology generated by sets of the form $\{\mu : \mu(A) \geq r\}$ for any Borel set A and real number r. ΔX is weak* compact and metrizable. The handy properties of metrizability greatly shorten the proofs:

Suppose X and Y are compact and metrizable. Denote $\sigma(X \times Y)$ as the Borel σ-algebra on $X \times Y$ and $\sigma(X) \otimes \sigma(Y)$ as the product σ-algebra on $X \times Y$. It follows that X and Y are complete, separable (Polish) and second countable, and that $\sigma(X \times Y) = \sigma(X) \otimes \sigma(Y)$. If μ is a Borel measure on X, then μ is regular and a net of Borel measure $\{\mu_\alpha\}$ weak* converges to μ if and only if $\int f d\mu_\alpha \to \int f d\mu$ for any bounded continuous real function f on X.

D1.2 Belief Space

$N = \{1, \cdots, n\}$ is a finite set of players; S is a compact metrizable space of payoff-relevant parameters; we call elements of S *states of nature*; T_i is the universal type space for player i, which is also compact and metrizable by construction; $\langle S \times \prod_{i=1}^n T_i, (t_i)_{i \in N} \rangle$ is the universal belief space, where $S \times \prod_{i=1}^n T_i$ is endowed with the product topology and is mapped under a continuous function t_i to $\Delta(S \times \prod_{i=1}^n T_i)$, the weak* compact space of Borel probability measures on $S \times \prod_{i=1}^n T_i$.[23] By Mertens-Zamir construction, $\omega' \in \text{supp} t_i(\omega)$ implies $t_i(\omega) = t_i(\omega')$, and $\pi_i(\omega) = \pi_i(\omega') = t_i$ implies $t_i(\omega) = t_i(\omega')$, i.e. a player "knows" his own type.

Let $\langle \Omega, (t_i)_{i \in N} \rangle$ be a belief-closed subspace (BL-subspace): Ω is a subset of $S \times \prod_{i=1}^n T_i$ from which it inherits the product topology and t_i is restricted on Ω with $t_i(\omega)(\Omega) = 1$ for any $\omega \in \Omega$. We assume Ω is closed and hence it is compact. We further require that $\omega \in \text{supp} t_i(\omega)$ meaning that a player cannot exclude the truth.[24] Following the mathematical convention, we simply call Ω a *belief space* without mentioning those t_i's.

Denote π_0 and π_i as the projection from Ω into S and T_i, respectively. They are all continuous by the definition of product topology. Hierarchies of beliefs are retraced from the belief space inductively as follows.

The first order belief mapping for player i, $\delta_i^1 : \Omega \to \Delta S$, is given by $\delta_i^1(\omega) := t_i(\omega) \circ \pi_0^{-1}$; unlike the usual composition, for any $\omega \in \Omega$ and any Borel set $B \subset S$, $\delta_i^1(\omega)(B) = t_i(\omega)(\{\omega' : \pi_0(\omega') \in B\})$.

Denote, for $k \geq 1$, $\delta^k := \prod_{i=1}^n \delta_i^k$, the kth order belief mapping profile for all players. The $k+1$th order belief mapping for player i, $\delta_i^{k+1} : \Omega \to \Delta\left(S \times \prod_{l=1}^k \delta^l(\Omega)\right)$, is given by $\delta_i^{k+1}(\omega) := t_i(\omega) \circ (\pi_0 \times \prod_{l=1}^k \delta^l)^{-1}$, i.e. for any $\omega \in \Omega$ and any Borel set $B \subset S \times \prod_{l=1}^k \delta^l(\Omega)$, $\delta_i^{k+1}(\omega)(B) = t_i(\omega)\left(\{\omega' : \pi_0(\omega') \times \prod_{l=1}^k \delta^l(\omega') \in B\}\right)$. The definition is valid only if δ^l's are measurable; since

[23] t_i is a stochastic kernel in the terminology of stochastic processes. We adopt the following definitions: For a topological space (X, τ), a Borel measure is simply a measure on $\sigma(\tau)$, the σ-algebra generated by all open sets. A Borel measure μ is regular if it is outer regular, $\mu(A) = \inf\{\mu(G) : G \text{ open}, A \subset G\}$, $\forall A \in \sigma(\tau)$, and tight, $\mu(A) = \sup\{\mu(K) : K \text{ compact}, K \subset A\}$, $\forall A \in \sigma(\tau)$.

[24] $t_i(\omega)$ has a unique compact support for any $\omega \in \Omega$ and $i \in N$. See, e.g. Theorem 10.13 in Aliprantis and Border [1]. In particular, $\text{supp} t_i(\omega) = t_i^{-1}(t_i(\omega))$. To see this, recall that $\omega' \in \text{supp} t_i(\omega)$ implies $t_i(\omega) = t_i(\omega')$ and conversely, if $t_i(\omega) = t_i(\omega')$, then $\omega' \in \text{supp} t_i(\omega') = \text{supp} t_i(\omega)$.

then $(\pi_0 \times \prod_{l=1}^{k} \delta^l)^{-1}(B)$ is Borel whenever B is. This is shown in Lemma 10 below.

Denote $\delta_i = (\delta_i^1, \delta_i^2, \cdots)$, which is the hierarchy of beliefs for player i generated by the belief space. For convenience, denote $\delta = (\delta_i)_{i \in N}$, $\delta_i^* = (\delta_i^0, \delta_i)$ and $\delta^* = (\delta_0, \delta)$ with $\delta_i^0 \equiv \delta_0 \equiv \pi_0$. $\delta^*(\omega)$ is the *description* of the state of world ω, including both the state of nature and the profile of belief hierarchies. $\delta_i^*(\omega)$ is the description from player i's perspective. Note that $\delta^* \neq (\delta_i^*)_{i \in N}$.

Lemma 10 *δ_i^k is continuous with respect to the weak* topology for all $i \in N$ and $k \geq 0$; δ_i, δ and δ^* are continuous with respect to the product topology. Hence, all those mappings are Borel measurable.*

Proof The continuity of $\delta_i^0 = \pi_0$ is immediate since the projection is continuous. Suppose a net $\{\omega_\alpha\}$ converges to ω in Ω and hence $t_i(\omega_\alpha) \to t_i(\omega)$ with respect to the weak* topology. By definition, $\delta_i^1(\omega) = t_i(\omega) \circ {\pi_0}^{-1}$. For $f \in C(\Omega)$,[25] we have, by transformation of measures,

$$\int_S f d\delta_i^1(\omega_\alpha) = \int_S f \; dt_i(\omega_\alpha) \circ \pi_0^{-1} = \int_\Omega f \circ \pi_0 \; dt_i(\omega_\alpha)$$

which converges to $\int_\Omega f \circ \pi_0 \; dt_i(\omega)$ by the continuity of t_i; this in turn equals $\int_S f \; dt_i(\omega) \circ \pi_0^{-1} = \int_S f \; d\delta_i^1(\omega)$, showing that δ_i^1 is continuous with respect to the weak* topology. Since $\delta_i^{k+1}(\omega) = t_i(\omega) \circ (\pi_0 \times \prod_{l=1}^{k} \delta^l)^{-1}$, the continuity of δ_i^{k+1} is obtained by the usual induction. Thus δ_i, δ and δ^* are all continuous with respect to the product topology. Borel measurability is implied by continuity. $\square$

According to Mertens-Zamir (1985), Ω is "non-redundant" (NR) if $\omega \neq \omega'$ implies $\delta^*(\omega) \neq \delta^*(\omega')$.[26] This implies that for each i, if $\omega' \notin \text{supp} t_i(\omega)$ then $\delta_i(\omega) \neq \delta_i(\omega')$, conforming with the usual intuition that non-redundancy means no two types of a player represent the same belief hierarchy. This is the content of Lemma 11. The "redundancy" is an unfortunate term as the payoff-irrelevant redundancy is nevertheless strategic relevant. Instead, we call a "non-redundant" belief space *reduced* to capture its mathematical content and *non-reduced* otherwise.

Let $\langle \tilde{\Omega}, (\tilde{t}_i)_{i \in N} \rangle$ be a belief-closed subspace of a non-reduced belief space $\langle S \times \prod_{i=1}^{n} \tilde{T}_i, \tilde{t}_i \rangle$. Again, we assume that $\tilde{\Omega}$, S and $\tilde{T}_i$'s are compact and metrizable and $\tilde{t}_i$ is weak* continuous. Note that $\langle S \times \prod_{i=1}^{n} \tilde{T}_i, \tilde{t}_i \rangle$ does not represent a universal space, but it is without loss of generality

[25] Ω is completely regular since it is compact and Hausdorff; hence, we can simply take $f \in C(\Omega)$ instead of u.s.c. function $f \in \Omega^{\mathbb{R}}$.

[26] Mertens and Zamir (1985, definition 2.4 and proposition 2.5) formulate the non-redundancy condition (NR) in terms of a separation condition which implies the property here.

to introduce it because any belief-closed space can be regarded as a subspace of a product space; we introduce it for easy description of the topological and measurable structure on $\tilde{\Omega}$. The hierarchies of beliefs can be retraced in a similar manner as we do for the reduced space and a continuity result similar to Lemma 1 holds.

D2 Results

Assumption 1 now takes the following form:

ASSUMPTION 1 For a non-reduced space $\tilde{\Omega}$, if $\tilde{\omega}' \in \text{supp}(t_i(\tilde{\omega}))$ for all $i \in N$ and $\tilde{\omega} \neq \tilde{\omega}'$ then $\tilde{\delta}^*(\tilde{\omega}) \neq \tilde{\delta}^*(\tilde{\omega}')$.

The universal space without NR condition does not exist even with the above restrictions and other stringent topological conditions. Further more, each player could have as many times as possible.

Example 5: The universal space does not exist without imposing NR-condition. We can construct a belief space with arbitrary cardinality even under Assumption 1: no pure redundancy is involved. Let $S = \{s\}$ be a singleton. Then the Mertens-Zamir universal space is also a singleton. Let C_1 and C_2 be any metrizable spaces endowed with the Borel σ-algebras. Let ν_i be an arbitrary probability measure on C_i such that ν_i. Define a mapping $\mu_i : C_i \to \Delta C_i$ as $\mu_i(c_i)(U) = 1$ if $c_i \in U$ and 0 otherwise for any Borel $U \subset C$. μ_i is continuous with respect to the weak* topology.[27] Let $C = C_1 \times C_2$ with the product topology. Define a mapping t_i on C as $t_i(c) = \mu_i(c_i) \otimes \nu_{-i}$. We claim that $t_i(c)$ is weak* continuous. To see this, for any bounded continuous function f on C, we have by Fubini's theorem (see, e.g., 1.d.3.i Mertens-Sorin-Zamir [18] for a topological version)

$$\int_C f \; dt_i(c) = \int_{C_{-i}} \int_{C_i} f(\cdot, \cdot) \; d\mu_i(c_i) d\nu_{-i} = \int_{C_{-i}} f(c_i, \cdot) \; d\nu_{-i}.$$

Thus t_i is weak* continuous. $\langle C, (t_i)_{i=1,2} \rangle$ defines a belief space, on which Assumption 1 is satisfied. If there is a universal space then the space should have a fixed cardinality, this contradicts the choice of C.

The non-existence result does not prevent us from understanding the redundant belief space, which is the main point in this paper.

[27] $\int f \; d\mu_i(c_i) = f(c_i)$ for any bounded continuous function f on C_i. The weak* continuity of μ_i follows immediately.

Definition 8 *Two belief spaces Ω and $\tilde{\Omega}$ are fundamentally compatible or simply compatible, written as $\Omega \sim \tilde{\Omega}$, if $\delta(\Omega) = \tilde{\delta}(\tilde{\Omega})$ i.e. they represents the same set of belief hierarchies.*

Let C be a compact metrizable space. Define a semiring on $\Omega \times C$ as $\mathcal{S} = \{W \times U : W \in \sigma(\Omega), U \in \sigma(C)\}$.

Let $q_i : \Omega \times C \to \Delta C$ be a function such that $q_i(\omega, c)(U)$ is measurable in ω for any $c \in C$ and Borel set $U \subset C$. $\{q_i(\omega, \cdot)\}_{\omega \in \Omega}$ is a family of transition[28] on C. $\langle C, (q_i)_{i \in N} \rangle$ is called a transition indexed by Ω.

Proposition 7

$$\tilde{t}_i^*(\omega \times c)(W \times U) = \int_W q_i(\omega', c)(U)\ t_i(\omega)(d\omega')$$

as a function $\Omega \times C \to \mathbb{R}^{\mathcal{S}}$ can be extended to a unique function $\tilde{t}_i : \Omega \times C \to \Delta(\Omega \times C)$.

The proof can be found in Section 4.

Definition 9 *We call $\langle C, (q_i)_{i \in N} \rangle$ adopted to Ω if the extended function is jointly continuous in ω and c.*

There is in general no sufficient and necessary conditions to go from separate continuity to joint continuity. Thus, if we were to impose more conditions (e.g. the property obviously holds in discrete case) on t_i and q_i which implies this property, the belief space we are considering is necessarily restricted (through t_i). Our full characterization below also illustrate that this definition is appropriate.

Theorem 2 *Suppose Ω is a reduced belief space. Then $\langle \Omega, (t_i)_{i \in N} \rangle \sim \langle \tilde{\Omega}, (\tilde{t}_i)_{i \in N} \rangle$ if and only if there exists a transition $\langle C, (q_i)_{i \in N} \rangle$ adapted to Ω, such that*

(1) $\tilde{\Omega}$ can be imbedded in $\Omega \times C$ such that $\tilde{\Omega}|_\Omega = \Omega$;[29]

(2) $\tilde{t}_i(\omega \times c)$ is the unique extension of the following measure on the semiring of measurable rectangles $\{W \times U : W \in \sigma(\Omega), U \in \sigma(C)\}$:

$$\tilde{t}_i^*(\omega \times c)(W \times U) = \int_W q_i(\omega', c)(U)\ t_i(\omega)(d\omega').$$

[28] Note that we do not require the measurability in c, see Liu [17] for related discussions.

[29] In this expression, we have identified $\tilde{\Omega}$ as a subspace in $\Omega \times C$.

In particular, we can take $\langle C, (q_i)_{i\in N}\rangle$ *with the following additional properties:*

(i) $C = C_1 \times \cdots \times C_n$, C_i *is compact and metrizable;*

(ii) $q_i(\omega \times c) = q_i(\omega \times c')$ *if* $c_i = c'_i$;

(iii) $c \in supp(q_i(\omega \times c))$;

(iv) *If* $\omega \times c \in \tilde{\Omega}$ *and* $\omega' \in supp(t_i(\omega))$ *then there exists a point* $\omega' \times c' \in \tilde{\Omega}$ *such that* $c_i = {c'}_i$.

The additional properties define an "individually uninformative" state-dependent correlating mechanism in terms of conditional probabilities[30]: (i) says the random device has n dimensions; (ii) says that player i only observes the ith coordinate; (iii) says that player i cannot exclude the true outcome; (iv) says the same type of player i expects to make the same observations.

D3 Proofs

D3.1 Two Alternative Definitions of Compatibility

The compatibility can be equivalently defined in terms of description δ^* or belief morphism φ. The results are of interest in their own right since they connect the intuitive notion of belief hierarchy with the abstract objects such as description and belief morphism. This is where the topology comes in.

Definition 10 *A continuous function* $\varphi : \tilde{\Omega} \to \Omega$ *is a belief morphism if,*

(1) $\pi_0(\varphi(\tilde{\omega})) = \pi_0(\varphi(\omega))$, *for each* $\tilde{\omega} \in \tilde{\Omega}$;

(2) $\tilde{t}_i(\tilde{\omega})(\varphi^{-1}(E)) = t_i(\varphi(\tilde{\omega}))(E)$, *for any Borel set* $E \subset \Omega$ *and* $\tilde{\omega} \in \tilde{\Omega}$.

We first prove several results which are easy to see in the discrete case.

Lemma 11 *In a reduced space* Ω, *if* $\omega' \notin supp(t_i(\omega))$ *then* $\delta_i(\omega) \neq \delta_i(\omega')$.

Proof Suppose to the contrary $\omega' \notin \mathrm{supp}t_i(\omega)$ and $\delta_i(\omega) = \delta_i(\omega')$. Define, for $k \geq 0$, $A_k = \{\omega'' \in \mathrm{supp}t_i(\omega') : \delta^l(\omega'') = \delta^l(\omega), 0 \leq l \leq k\}$.Since $\delta_i(\omega') = \delta_i(\omega)$ and in particular $\delta_i^1(\omega') = \delta_i^1(\omega)$ i.e. $t_i(\omega') \circ \pi_0^{-1} = t_i(\omega) \circ \pi_0^{-1}$, we have $\pi_0(\omega) \in \mathrm{supp}(t_i(\omega) \circ \pi_0^{-1}) = \mathrm{supp}(t_i(\omega') \circ \pi_0^{-1})$. So there is some $\omega'' \in \mathrm{supp}t_i(\omega')$ with $\pi_0(\omega) = \pi_0(\omega'')$, showing that $A_0 \neq \emptyset$. $A_k \neq \emptyset$ by the usual induction. Note also that $\{A_k\}$ is a decreasing sequence and that A_k is compact since $\mathrm{supp}t_i(\omega)$ is compact and $(\delta^0, \cdots, \delta^k)$ is continuous (Lemma 1); So $\bigcap_{k=1}^{\infty} A_k \neq \emptyset$. $\square$

[30]The mechanism in terms of unconditional probability can be easily obtained, e.g. for any $i \in N$ and $\omega \in \Omega$, define $q_i(\omega)$ as a probability measure that put unit mass on a point $c^* \notin C$; then $\langle C \cup \{c^*\}, (q_i)_{i\in N}\rangle$ is a mechanism according to Definition 3 in the main text and the properties of those conditional probabilities are trivially obtained.

Lemma 12 *If $\Omega \sim \tilde{\Omega}$ then there exists a unique function $\varphi : \tilde{\Omega} \to \Omega$ such that $\tilde{\delta}^*(\tilde{\omega}) = \delta^*(\varphi(\tilde{\omega}))$; that is, $\varphi = \delta^{*-1} \circ \tilde{\delta}^*$ is well defined. Furthermore, φ is surjective and continuous.*

Proof We will show that for any $\tilde{\omega} \in \tilde{\Omega}$ there exists a unique $\omega \in \Omega$ such that $\tilde{\delta}^*(\tilde{\omega}) = \delta^*(\omega)$. By compatibility, there exists $\omega' \in \Omega$ such that $\delta_i(\omega') = \tilde{\delta}_i(\tilde{\omega})$, i.e. $\delta_i^k(\omega') = \tilde{\delta}_i^k(\tilde{\omega}), k \geq 1$. Define $B_k = \{\omega \in \Omega : \delta^l(\omega) = \tilde{\delta}^l(\tilde{\omega}), 0 \leq l \leq k\},\ k \geq 0$. Notice that $B_k \neq \emptyset$ since $\delta_i^{k+1}(\omega') = \tilde{\delta}_i^{k+1}(\tilde{\omega})$ implies they have the same support; B_k is closed since δ^k is continuous. Thus $\bigcap_{k\geq 0} B_k \neq \emptyset$ by the compactness of Ω. Take any $\omega \in \bigcap_{k\geq 1} B_k$; we then have $\delta^*(\omega) = \tilde{\delta}^*(\tilde{\omega})$. The existence of ω is established. The uniqueness of such ω follows from the reducedness of Ω. Thus $\tilde{\delta}^*(\tilde{\omega}) = \delta^*(\omega)$ uniquely determines a function $\varphi : \tilde{\omega} \mapsto \omega$. $\varphi = \delta^{*-1} \circ \tilde{\delta}^*$ is well defined.

We claim that φ is surjective, i.e. $\varphi^{-1}(\omega) \neq \emptyset$ for any $\omega \in \Omega$. To see this, we simply define $B_k = \{\tilde{\omega} \in \tilde{\Omega} : \delta^l(\omega) = \tilde{\delta}^l(\tilde{\omega}), 0 \leq l \leq k\},\ k \geq 0$, for any fixed ω. Exactly the same argument as above shows that the existence of an $\tilde{\omega}$ such that $\tilde{\delta}^*(\tilde{\omega}) = \delta^*(\omega)$. (But we do not have uniqueness since $\tilde{\Omega}$ is non-reduced.) By the definition of φ, $\varphi(\tilde{\omega}) = \omega$.

We now show that φ is continuous. Take a closed subset (and hence compact) $F \subset \Omega$. $\delta^*(F)$ is compact and hence closed by continuity of δ^* (Recall that the space of belief hierarchies is Hausdorff.) Thus $\varphi^{-1}(F) = \left(\tilde{\delta}^*\right)^{-1}(\delta^*(F))$ is closed by continuity of $\tilde{\delta}^*$, showing the continuity of φ. $\square$

Proposition 8 *For a reduced space Ω, $\Omega \sim \tilde{\Omega}$ if and only if $\delta^*(\Omega) = \tilde{\delta}^*(\tilde{\Omega})$.*

Proof Suppose $\Omega \sim \tilde{\Omega}$. By Lemma 12, $\delta^{*-1} \circ \tilde{\delta}^*(\tilde{\Omega}) = \varphi(\tilde{\Omega}) = \Omega$; thus, $\tilde{\delta}^*(\tilde{\Omega}) = \delta^*(\Omega)$. On the other hand, $\tilde{\delta}^*(\tilde{\Omega}) = \delta^*(\Omega)$ implies $\tilde{\delta}(\tilde{\Omega}) = \delta(\Omega)$ by definition, showing that $\Omega \sim \tilde{\Omega}$. $\square$

Proposition 9 *For a reduced belief space Ω, $\Omega \sim \tilde{\Omega}$ if and only if there exists a surjective belief morphism $\varphi : \tilde{\Omega} \to \Omega$.*

Proof The sufficiency follows immediately from the definition.[31] For the necessity part, we show that $\varphi = \delta^{*-1} \circ \tilde{\delta}^*$ defined in Lemma 12 is a surjective belief morphism. For a Borel set $W \subset \Omega$ define, for $k \geq 0$,

$$B_k \ := \ \{\tilde{\omega}' \in \tilde{\Omega} : \tilde{\delta}^l(\tilde{\omega}') = \delta^l(\omega'),\ 0 \leq l \leq k, \omega' \in W\}$$

[31] Heifetz and Samet (1998, Proposition 5.2) also shows that belief morphism preserves descriptions in a purely measure-theoretic setup. The necessity part is not topology free.

$$
\begin{aligned}
D_k &:= \{(\tilde{\delta}^l(\tilde{\omega}'))_{0\le l\le k} : \tilde{\omega}' \in B_k\} = \{(\delta^l(\omega'))_{0\le l\le k} : \omega' \in W\} \\
E_k &:= \{\omega' \in \Omega : (\delta^l(\omega'))_{0\le l\le k} \in D_k\}
\end{aligned}
$$

It is easy to show the following properties. Their proofs are left to the readers.

Claim 1: $\{B_k\}_{k=0,1,\cdots}$ is a decreasing sequence and

$$
B := \bigcap_{k\ge 0} B_k = \{\tilde{\omega}' \in \tilde{\Omega} : \tilde{\delta}(\tilde{\omega}') = \delta(\omega'),\ \omega' \in W\} = \varphi^{-1}(W).
$$

Claim 2: The duality between B_k and D_k.

$$
B_k := \{\tilde{\omega}' \in \tilde{\Omega} : (\tilde{\delta}_l(\tilde{\omega}'))_{0\le l\le k} \in D_k\}.
$$

Claim 3: $\{E_k\}_{k=0,1,\cdots}$ is a decreasing sequence with $W \subset E_k$ and $W = \bigcap_{k\ge 0} E_k$.[32]

Thus for any $\tilde{\omega} \in \tilde{\Omega}$,

$$
\begin{aligned}
\tilde{t}_i(\tilde{\omega})(\varphi^{-1}(W)) &= \tilde{t}_i(\tilde{\omega})(B) \\
&= \tilde{t}_i(\tilde{\omega})(\lim_{k\to+\infty} B_k) \quad \text{(by Claim 1)} \\
&= \lim_{k\to+\infty} \tilde{t}_i(\tilde{\omega})(B_k) \\
&= \lim_{k\to+\infty} \tilde{\delta}_i^{k+1}(\tilde{\omega})(D_k) \quad \text{(by Claim 2 and the definition of } \tilde{\delta}_i^{k+1}\text{)} \\
&= \lim_{k\to+\infty} \delta_i^{k+1}(\varphi(\tilde{\omega}))(D_k) \quad \text{(by Lemma 12)} \\
&= \lim_{k\to+\infty} t_i(\varphi(\tilde{\omega}))(E_k) \\
&= t_i(\varphi(\tilde{\omega}))(\lim_{k\to+\infty} E_k) \\
&= t_i(\varphi(\tilde{\omega}))(W) \quad \text{(by Claim 3)}.
\end{aligned}
$$

Also notice that $\pi_0(\varphi(\tilde{\omega})) = \pi_0(\tilde{\omega})$ is implied by $\tilde{\delta}^*(\tilde{\omega}) = \delta^*(\varphi(\tilde{\omega}))$.

D3.2 Proof of Proposition 1

The following result is of interest in his own right.

[32]The equality follows from the non-reducedness of Ω.

Proposition 10

$$\tilde{t}_i^*(\omega \times c)(W \times U) = \int_W q_i(\omega', c)(U)\ t_i(\omega)(d\omega')$$

defines a measure on the semiring of measurable rectangles $\mathcal{S} = \{W \times U : W \in \sigma(\Omega), U \in \sigma(C)\}$

Proof We only need to show countable additivity. Suppose $\{W_n \times U_n\}$ is a pairwise disjoint sequence in $\mathcal{S}$ and $\bigcup_{n=1}^{\infty} W_n \times U_n \in \mathcal{S}$, i.e. $\bigcup_{n=1}^{\infty} W_n \times U_n = W \times U$ for some $W \in \sigma(\Omega)$ and $U \in \sigma(C)$. Thus, for any $v \in \Omega$ and $u \in C$,

$$\begin{aligned}
\chi_W(v) \times \chi_U(u) &= \chi_{W \times U}(v \times u) \\
&= \chi_{\bigcup_{n=1}^{\infty} W_n \times U_n}(v \times u) \\
&= \sum_{n=1}^{\infty} \chi_{W_n \times U_n}(v \times u) \\
&= \sum_{n=1}^{\infty} \chi_{W_n}(v) \times \chi_{U_n}(u)
\end{aligned}$$

For some $\omega' \in \Omega$ and $c \in C$, integration over U with respect to $q_i(\omega', c)$ yields

$$\chi_W(v) q_i(\omega', c)(U) = \sum_{n=1}^{\infty} \chi_{W_n}(v) q_i(\omega', c)(U_n).$$

Then integration over W with respect to $t_i(\omega)$, we have

$$\int_W q_i(\omega', c)(U)\ t_i(\omega)(d\omega') = \sum_{n=1}^{\infty} \int_{W_n} q_i(\omega', c)(U_n) t_i(\omega)(d\omega')$$

This proves the σ-additivity of $\tilde{t}_i^*(\omega \times c)$ on $\mathcal{S}$. $\square$

By a version of Carathéodory Extension Theorem[33] $t_i^*(\omega \times c)$ extend to a unique measure $t_i(\omega \times c)$ on $\sigma(\mathcal{S})$, the Borel σ-algebra on $\Omega \times C$. In particular, for any $\tilde{E} \subset \tilde{\Omega}$,

$$\tilde{t}_i(\omega \times c)(\tilde{E}) = \inf \left\{ \sum_{n=1}^{\infty} \tilde{t}_i^*(\omega \times c)(W_n \times U_n) : W_n \times U_n \in \mathcal{S} \text{ and } \tilde{E} \subset \bigcup_{n=1}^{\infty} W_n \times U_n \right\}.$$

Clearly $t_i(\omega \times c)(\Omega \times C) = 1$.

[33] See Aliprantis and Border (1999, 9.22).

D3.3 Proof of Theorem 2: Sufficiency

Given two belief spaces $\Omega \sim \tilde{\Omega}$. Define $\varphi : \tilde{\Omega} \to \Omega$ as $\varphi(\omega \times c) = \omega$. Clearly, $\varphi(\tilde{\Omega}) = \Omega$ and $\pi_0(\tilde{\omega}) = \pi_0(\varphi(\tilde{\omega}))$. For any measurable set $E \subset \Omega$ and $\omega \times c \in \tilde{\Omega}$, $\varphi^{-1}(E) = (E \times C) \cap \tilde{\Omega}$ and hence $\tilde{t}_i(\omega \times c)(\varphi^{-1}(E)) = \tilde{t}_i(\omega \times c)(E \times C)$. To conclude that φ is a belief morphism, we need $\tilde{t}_i(\omega \times c)(E \times C) = t_i(\omega)(E)$. To see this, note that $\tilde{t}_i(\omega \times c)$ is extended from $\tilde{t}_i^*(\omega \times c)$, they agree on $\mathcal{S}$. Thus $\tilde{t}_i(\omega \times c)(E \times C) = \tilde{t}_i^*(\omega \times c)(E \times C) = \int_E q_i(\omega', c)(C) t_i(\omega)(d\omega') = t_i(\omega)(E)$.

D3.4 Proof of Theorem 2: Necessity

The proof contains several steps: (1) define $C = \prod_{i=1}^n T_i$ and imbed $\tilde{\Omega}$ into $\Omega \times C$; (2) define $\tilde{t}_i^*$ as the restriction of $\tilde{t}_i$ on the semiring from which $\tilde{t}_i$ is the unique extension; (3) $\tilde{t}_i^*(\omega \times c)(\cdot \times U)$ as a measure on Ω for a fixed U is absolutely continuous with respect to $t_i(\omega)$ and thus Radon-Nikodym Theorem gives a derivative; (4) the Radon-Nikodym derivative has the desired properties.

Let φ be the surjective belief morphism from $\tilde{\Omega}$ to Ω. $\tilde{\delta}(\varphi(\tilde{\omega})) = \delta(\omega)$. Define a mapping $\rho : \tilde{\Omega} \to \Omega \times \prod_{i=1}^n \tilde{T}_i$ as $\rho : \tilde{\omega} \mapsto (\varphi(\tilde{\omega}), \pi_1(\tilde{\omega}), \cdots, \pi_n(\tilde{\omega}))$. Recall that π_i is the projection mapping.

Lemma 13 *ρ is an imbedding; that is, ρ maps $\tilde{\Omega}$ to $\rho(\tilde{\Omega})$ homeomorphically.*

Proof φ and π_i's are all continuous, so ρ is continuous. Suppose $\tilde{\omega}, \tilde{\omega}' \in \tilde{\Omega}$ and $\tilde{\omega} \neq \tilde{\omega}'$. If $\pi_i(\tilde{\omega}) = \pi_i(\tilde{\omega}')$ for all $i \in N$, it must be that $\varphi(\tilde{\omega}) \neq \varphi(\tilde{\omega}')$; otherwise, we have $\tilde{\delta}^*(\tilde{\omega}) = \tilde{\delta}^*(\tilde{\omega}')$—violating Assumption 1. Thus ρ is one-to-one. Since ρ maps a compact metric space into a compact metric space, ρ is a homeomorphism between $\tilde{\Omega}$ and $\rho(\tilde{\Omega})$ and hence an imbedding. □

For $\omega \times c \in \Omega \times C$ such that $\omega \times c = \rho(\tilde{\omega})$, define a measure on the semiring $\mathcal{S}$ via $\tilde{t}_i^*(\omega \times c)(W \times C) = \tilde{t}_i(\tilde{\omega})(\rho^{-1}(W \times U))$. The Carathéodory Extension Theorem in turns implies that the unique extension of $\tilde{t}_i^*(\omega \times c)$ on $\sigma(\mathcal{S})$, $\tilde{t}_i(\omega \times c)$, is exactly $t_i(\tilde{\omega})\rho^{-1}$. So we can identify the two belief spaces $\langle \tilde{\Omega}, \tilde{t}_i \rangle$ and $\langle \rho(\tilde{\Omega}), \tilde{t}_i \rangle$.

Lemma 14 *For each tuple (ω, c, U) with $\omega \times c = \rho(\tilde{\omega}) \in \Omega$, there exists a nonnegative measurable function $q_i(\cdot, \omega, c)(U)$ on Ω such that*

$$\tilde{t}_i^*(\omega \times c)(W \times U) = \int_W q_i(\omega', \omega, c)(U)\ t_i(\omega)(d\omega').$$

Proof For a given tuple (ω, c, U), $\tilde{t}_i^*(\omega \times c)(\cdot \times U)$ is a finite Borel measure over Ω (but not necessarily a probability measure.) We show $\tilde{t}_i^*(\omega \times c)(\cdot \times U)$ is absolutely continuous with respect to $t_i(\omega)$; then the lemma is immediate by Radon-Nikodym Theorem. To see this, notice that $\tilde{t}_i^*(\omega \times c)(W \times U) = \tilde{t}_i(\tilde{\omega})(\rho^{-1}(W \times U)) \leq \tilde{t}_i(\tilde{\omega})(\rho^{-1}(W \times C)) = \tilde{t}_i(\tilde{\omega})(\varphi^{-1}(W)) = t_i(\omega)(W)$. The last equality is by the definition of belief morphism. It follows immediately that $t_i(\omega)(W) = 0$ implies $\tilde{t}_i^*(\omega \times c)(W \times U) = 0$. $\square$

Lemma 15 *With the assumptions in Lemma 14,*

(1) *The Radon-Nikodym derive $q_i(\omega', \omega, c)$ can be rewritten as $q_i(\omega', c)$. .*

(2) $q_i(\omega', c) = q_i(\omega', c')$ *if* $c_i = {c'}_i$.

Proof Notice that since $\varphi(\tilde{\omega}) = \omega$, $t_i(\omega) = \tilde{t}_i(\tilde{\omega}) \circ \varphi^{-1}$ by the definition of belief morphism. Since $\pi_i(\tilde{\omega}) = c_i$, we have that $\tilde{t}_i(\tilde{\omega}) = \tilde{t}_i(\tilde{\omega}')$ for any $\tilde{\omega}'$ such that $\pi_i(\tilde{\omega}') = c_i$. With some abuse of notation, we can write $c_i = \tilde{t}_i(\tilde{\omega})$ if $\pi_i(\tilde{\omega}) = c_i$. Thus, if $\omega \times c = \rho(\tilde{\omega}) \in \Omega$, then $\tilde{t}_i^*(\omega \times c) = c_i \circ \rho^{-1}$ and $t_i(\omega) = c_i \circ \varphi^{-1}$. Hence the Radon-Nikodym derive can be written as $q_i(\omega', c)$ independently with ω, as seen from the following version of Lemma 14:

For each tuple (ω, c, U) with $\omega \times c = \rho(\tilde{\omega}) \in \Omega$, there exists a nonnegative measurable function $q_i(\cdot, c)(U)$ on Ω such that

$$(c_i \circ \rho^{-1})(W \times U) = \int_W q_i(\omega', c)(U) \ (c_i \circ \varphi^{-1})(d\omega').$$

(2) is obvious from the above expression.

Lemma 16

For any $(\omega, c) \in \rho(\tilde{\Omega})$, $q_i(\omega', c)$ is a probability measure on $\sigma(C)$, $c_i \circ \varphi^{-1}$-a.s.

(i) *$q_i(\omega', c)(U)$ is a countably additive set function on $\sigma(C)$, $c_i \circ \varphi^{-1}$-a.s.*

(ii) *$q_i(\omega', c)(C) = 1$, $c_i \circ \varphi^{-1}$-a.s.*

Proof

(i) Let $\{U_{k=1,2,\cdots}\}$ be a sequence of pairwise disjoint Borel subsets of C. Then

$$(c_i \circ \rho^{-1})(W \times \bigcup_{k \geq 1} U_k) = \sum_{k \geq 1} (c_i \circ \rho^{-1})(W \times U_k)$$

Thus,

$$\begin{aligned}\int_W q_i(\omega', c)(\bigcup_{k\geq 1} U_k)(c_i \circ \varphi^{-1})(d\omega') &= \sum_{k\geq 1}\int_W q_i(\omega', c)(U_k)(c_i \circ \varphi^{-1})(d\omega') \\ &= \int_W \sum_{k\geq 1} q_i(\omega', c)(U_k)(c_i \circ \varphi^{-1})(d\omega') \text{ (Monotone Convergence Theorem)}\end{aligned}$$

Since the equality holds for any Borel subset $W \subset \Omega$, we conclude that, for any $\omega \times c \in \rho(\tilde{\Omega})$,

$$q_i(\omega', c)(\bigcup_{k\geq 1} U_k) = \sum_{k\geq 1} q_i(\omega', c)(U_k), \qquad c_i \circ \varphi^{-1}\text{-a.s.}$$

(ii) By the relation of ρ and φ,

$$(c_i \circ \rho^{-1})(W \times C) = c_i(\rho^{-1}(W \times C)) = c_i(\varphi^{-1}(W)) = (c_i \circ \varphi^{-1})(W).$$

This together with lemma 14 gives,

$$\int_W 1\ (c_i \circ \varphi^{-1})(d\omega') = \int_W q_i(\omega', c)(C)\ (c_i \circ \varphi^{-1})(d\omega') \text{ for any } W.$$

Thus $q_i(\omega', c)(C) = 1$, $c_i \circ \varphi^{-1}$-a.s. We can take a version of the q_i such that it is a transition.

□

Now we have constructed $\langle C, (q_i)_{i\in N}\rangle$ with the desired properties.

References

[1] C. D. Aliprantis, K. C. Border, Infinite Dimensional Analysis. Springer-Verlag, Berlin (1999)

[2] R. J. Aumann, Subjectivity and Correlation in Randomized Strategies, J. Math. Econ. 1 (1974), 67–96.

[3] R. J. Aumann, Correlated Equilibrium as an Expression of Bayesian Rationality, Econometrica 55(1987), 1–18.

[4] R. J. Aumann, Common Priors: A Reply to Gul, Econometrica 66(1998), 928–938.

[5] B. Battigalli, M. Siniscalchi, Rationalization and Incomplete Information, Advances in Theoretical Economics, Berkeley Electronic Press, 3(1)(2003), 1073–1073.

[6] A. Brandenburger, E. Dekel, Hierarchies of Beliefs and Common Knowledge, J. Econ. Theory 59(1987), 189–198.

[7] A. Brandenburger, A. Friedenberg, Can Hidden Variables Explain Correlation? miemo (2005) New York University.

[8] E. Dekel, D. Fudenberg, S. Morris, Interim Rationalizability, mimeo(2005) Princeton University.

[9] E. B. Dynkin, On extreme X-harmonic functions, forthcoming Mathematical Research Letters **12** 2006 10001–10011.

[10] J. Ely, M. Peski, Hierarchies of Belief and Interim Rationalizability, mimeo (2005) Northwestern University.

[11] Y. Feinberg, Characterizing Common Priors in the Form of Posteriors, J. Econ. Theory 91 (2002), 127–179.

[12] F. Forges, Five Legitimate Definitions of Correlated Equilibrium in Games with Incomplete Information, Theory Dec. 35 (1993), 277–310.

[13] F. Gul, A Comment on Aumann's Bayesian View, Econometrica 66(1998), 923–928.

[14] J. C. Harsanyi, Games with Incomplete Information Played by "Bayesian" Players, Parts I, II, and III, Manage. Sci. 14(1967-68), 159–182, 320–334, 486–502.

[15] A. Heifetz, The Bayesian Formulation of Incomplete Information - the Non-Compact Case, Int. J. Game Theory 21(1993) 329-338.

[16] B. L. Lipman, Finite Order Implications of Common Priors, Econometrica 71(2003), 1255-1267.

[17] Q. Liu, Structural Relations Between Type Spaces, mimeo (2005) Stanford University.

[18] J. F. Mertens, S. Sorin, S. Zamir, Repeated Games: Part A: Background Material," CORE Discussion Paper No. 9420, Université Catholique de Louvain (1994)

[19] J.F. Mertens, S. Zamir, Formulation of Bayesian Analysis for Games with Incomplete Information, Int. J. Game Theory 14(1985), 1–29.

[20] D. Novikov, Hahn Decomposition and Radon-Nikodym Theorem with A Parameter. mimeo (2005), Purdue University.