Interactive epistemology and solution concepts for games with asymmetric information

Pierpaolo Battigalli* Alfredo Di Tillio† Edoardo Grillo‡ Antonio Penta§

December 10, 2010

Abstract

We use an interactive epistemology framework to provide a systematic analysis of some solution concepts for games with asymmetric information. We characterize solution concepts using expressible epistemic assumptions, represented as events in the canonical space generated by primitive uncertainty about the payoff relevant state, payoff irrelevant information, and actions. In most of the paper we adopt an interim perspective, which is appropriate to analyze genuine incomplete information. We relate Delta-rationalizability (Battigalli and Siniscalchi, 2003) to interim correlated rationalizability (Dekel, Fudenberg, and Morris, 2007) and to rationalizability in the interim strategic form. We also consider the ex ante perspective, which is appropriate to analyze asymmetric information about an initial chance move. We prove the equivalence between interim correlated rationalizability and an ex ante notion of correlated rationalizability.

1 Introduction

In the last few years, ideas related to rationalizability have been increasingly applied to the analysis of games with asymmetric information, interpreted either as games with genuine incomplete information or games with imperfect information about an initial chance move.[1] Yet there seems to be no canonical definition of rationalizability for this class of games. Some authors put forward and apply notions that avoid the specification of a type space à la Harsanyi—Battigalli (2003), Battigalli and Siniscalchi (2003, 2007), Bergemann and Morris (2005, 2007). Others instead deal with

*IGIER and Department of Economics, Università Luigi Bocconi, pierpaolo.battigalli@unibocconi.it

†IGIER and Department of Economics, Università Luigi Bocconi, alfredo.ditillio@unibocconi.it

‡Department of Economics, Princeton University, egrillo@princeton.edu

§Department of Economics, University of Wisconsin-Madison, apenta@ssc.wisc.edu

[1] See Battigalli (2003, section 5), and Battigalli and Siniscalchi (2003, section 6) for references to applications of rationalizability to models of reputation, auctions and signaling. Bergemann and Morris (2005) apply a notion of iterated dominance to robust implementation. Carlsson and van Damme (1993) show that global games can be solved by iterated dominance—see also Morris and Shin (2007) for a recent evaluation of this result and its applications.

solution concepts for the Bayesian game obtained by appending a type space to the basic economic environment—Ely and Pęski (2006), Dekel, Fudenberg, and Morris (2007). However, while the adoption of a type space (and Bayesian Nash equilibrium) is common practice in economics, applying notions of rationalizability to Bayesian games requires some care. It is well known that some modeling details, which arguably should not matter, do instead affect the conclusions of the analysis, as the following two examples illustrate.

1.1 Ex ante vs interim perspective

Although it seems natural to transform a Bayesian game into a strategic form game and apply standard rationalizability, there is more than one way to do this, and the results vary accordingly. Indeed, unlike with Bayesian Nash equilibrium, rationalizability in the *ex ante strategic form*, where each player chooses a mapping from types into actions, is a *refinement* of rationalizability in the *interim strategic form*, where each type of each player chooses an action.[2] To see this, consider the following game, where the payoff state $\theta \in \Theta = \{\theta', \theta''\}$ is known only to player 2:

θ':

	L	M	R
T	$0,3$	$0,2$	$3,0$
B	$2,0$	$2,2$	$2,3$

θ'':

	L	M	R
T	$3,3$	$0,2$	$0,0$
B	$2,0$	$2,2$	$2,3$

Assume that player 1 ascribes equal probabilities to θ' and θ'' and that there is common belief in this fact. Thus we obtain a Bayesian game where player 1 has only one (uninformed) type, who assigns equal probabilities to the two (informed) types of player 2, which we can identify with θ' and θ''. Given any conjecture μ about player 1's action, L is a best reply only if $\mu[T] \geq 2/3$, while R is a best reply only if $\mu[T] \leq 1/3$. Thus the two strategies of player 2 that specify L (resp. R) under θ' and R (resp. L) under θ'', cannot be ex ante best replies to any conjecture μ. If player 1 assigns zero probability to these strategies, the expected payoff of T is at most $3/2$, hence T is not ex ante rationalizable. On the other hand, interim rationalizability regards the two types of player 2 as different players: under θ' player 2 may believe $\mu[T] \geq 2/3$, while under θ'' she may believe $\mu[T] \leq 1/3$ (or vice versa). Thus, in the second iteration of the interim rationalizability procedure player 1 may assign probability close to 1 to L under θ' and R under θ'', and hence choose T as a best response. This implies that every action is interim rationalizable.

The difference between ex ante and interim rationalizability, as illustrated in this example, has been accepted as a natural consequence of the fact that the latter allows different types of the same player to hold different conjectures. However, we maintain that it is disturbing: ex ante expected payoff maximization is equivalent to interim expected payoff maximization,[3] and rationalizability is

[2]This holds under weak conditions on players' (subjective) priors: either (a) priors have a common support, or (b) each player assigns positive prior probability to each one of his types.

[3]Interim maximization implies ex ante maximization; under the assumptions in footnote 2, also the converse is true.

supposed to capture just the behavioral consequences of the assumption that players are expected payoff maximizers and have common belief in this fact. Given the above, how can ex ante and interim rationalizability deliver different results? There must be additional assumptions (i.e. besides rationality and common belief in rationality) determining the discrepancy. As we prove in section 4, these have little to do with the fact that types are treated as distinct players in the interim strategic form; instead, the reasons are to be found in the different *conditional independence* assumptions underlying the two solution concepts; once these assumptions are removed, and thus correlation is allowed, we obtain equivalent ex ante and interim solution concepts.

1.2 Redundant types

Rationalizability in the (ex ante or interim) strategic form is not invariant to the addition of *redundant types*, that is, multiple types that encode the same information and hierarchy of beliefs. Indeed, Ely and Pęski (2006) and Dekel, Fudenberg, and Morris (2007) noticed that adding redundant types may enlarge the set of rationalizable outcomes.[4] In particular, they illustrate this for *interim independent rationalizability*, which, as we remark later on in the paper, is equivalent to rationalizability in the interim strategic form. Dekel, Fudenberg, and Morris (2007) also introduce *interim correlated rationalizability*, a weaker notion that is invariant to the addition of redundant types. To illustrate, consider the following game, borrowed from Dekel, Fudenberg, and Morris (2007), where the payoff state $\theta \in \Theta = \{\theta', \theta''\}$ is unknown to both players:

	B	N
B	$2, -4$	$-1, 0$
N	$0, -1$	$0, 0$

θ'

	B	N
B	$-4, 2$	$-1, 0$
N	$0, -1$	$0, 0$

θ''

Assume that it is common belief that each player attaches equal probabilities to the two states. The simplest Bayesian game representing this situation has only one type for each player. In this case the ex ante and interim strategic forms coincide, and B is dominated, hence not rationalizable:

	B	N
B	$-1, -1$	$-1, 0$
N	$0, -1$	$0, 0$

But we can think of another Bayesian game representing the same situation, where each player $i = 1, 2$ has two types, t_i and t_i', and beliefs are generated by the common prior below:

	t_2	t_2'
t_1	$1/4$	0
t_1'	0	$1/4$

θ'

	t_2	t_2'
t_1	0	$1/4$
t_1'	$1/4$	0

θ''

[4] Liu (2009) and Sadzik (2009) analyze related issues of invariance of solution concepts to redundancies.

As before, it is common belief that θ' and θ'' are considered equally likely, therefore we are just adding redundant types. But in the induced Bayesian game, B is rationalizable for both types of both players. Since ex ante rationalizability implies interim rationalizability, to see this it suffices to show that there are ex ante rationalizable strategies where either type chooses B. Let XY denote the strategy where t_i' chooses X and t_i'' chooses Y. The ex ante strategic form (with every payoff multiplied by 4 for convenience) is as follows:

	BB	BN	NB	NN
BB	$-4,-4$	$-4,-2$	$-4,-2$	$-4,0$
BN	$-2,-4$	$\underline{1},-5$	$-5,\underline{1}$	$-2,0$
NB	$-2,-4$	$-5,\underline{1}$	$\underline{1},-5$	$-2,0$
NN	$0,-4$	$0,-2$	$0,-2$	$\underline{0},\underline{0}$

Note that BB is dominated, but the set of strategy profiles $\{BN,NB,NN\} \times \{BN,NB,NN\}$ has the best response property (Pearce, 1984): as the underlined payoffs indicate, each strategy in the set of player i is a best response to some strategy in (and hence to some belief on) the set of player $-i$. Thus, BN and NB are ex ante rationalizable, and B is interim rationalizable for every type, hence interim correlated rationalizable for every type.

Adding redundant types can expand the rationalizable set of the strategic form. As we have already argued, (ex ante or interim) rationalizability must capture more than just common belief of expected payoff maximization in a situation of incomplete information. How are these additional assumptions related to the presence of redundant types? The reason is that a player may have *payoff-irrelevant information* that the other player believes to be correlated with the payoff state. The example shows that this is possible even if this payoff-irrelevant information does not affect the players' hierarchies of beliefs about the payoff state. Since actions may depend on this information, it is possible that a player's beliefs satisfy conditional independence when considering all the information of the other player, and yet when they are conditioned only on the payoff-relevant information (and hierarchy of beliefs over the payoff state) of the other player, they exhibit correlation between the payoff state and the other player's action. Thus, whenever redundancy can indeed be expressed in terms of payoff-irrelevant information, and such information is taken into account, conditional independence has less bite, and the set of rationalizable actions accordingly expands.

1.3 Expressible epistemic characterizations

The two issues illustrated above should make us suspicious about solution concepts mechanically obtained by applying a known solution algorithm (rationalizability) to the strategic forms of Bayesian games. In order to understand better the various solution concepts and their different predictions, a *formal* analysis of their underlying assumptions is needed, and this is precisely what we propose in this paper. Indeed, the problem with these notions is that they are not completely transparent because, unlike rationalizability in games of complete information, they have not been characterized using *expressible assumptions* about rationality and beliefs. To see what we mean, recall that for games with complete information, Tan and Werlang (1988) show that an action is rationalizable if and only if it is consistent with rationality, i.e. expected payoff maximization, and common belief in rationality.[5] These assumptions are expressible in a language describing primitive terms (actions) and terms derived from the primitives (beliefs about actions, beliefs about actions and beliefs of others, etc.). As explained in Heifetz and Samet (1998), the expressions in such a language can be represented as (and indeed identified with) measurable subsets of the canonical state space where each state specifies the players' actions and hierarchies of beliefs about actions.[6]

Our aim is to characterize rationalizability in games of incomplete information in the same manner, and hence achieve a deeper understanding of the issues illustrated above. Expressing assumptions such as rationality, in the context of games with incomplete information, requires of course a richer language. Thus, our primitives include not only the actions A_i available to each player $i \in I = \{1, 2\}$,[7] but also the private information x_i possessed by this player, as well as the payoff state θ. In order to get a general formulation, we explicitly disentangle the aspects of i's information that are payoff-relevant, denoted by θ_i, from those that are not, denoted by y_i. Thus, i's information is $x_i = (\theta_i, y_i) \in \Theta_i \times Y_i = X_i$, and no player's payoff depends on y_i. We write $\theta = (\theta_0, \theta_1, \theta_2) \in \Theta_0 \times \Theta_1 \times \Theta_2 = \Theta$ and we let $g_i : \Theta \times A_1 \times A_2 \to \mathbb{R}$ denote i's payoff function, so that we permit payoff uncertainty to persist (via θ_0) even after pooling all players' information. Note that y_i can be *strategically* relevant because i's action can depend on it, and the other player can believe that it is correlated with θ_0, thus inducing a potential correlation between θ_0 and i's action.[8] Indeed, our formulation allows us to state characterization results that otherwise could not be stated—we discuss the role of θ_0 and y_i in more detail, when we preview our results below.

In the language described above, an *expressible assumption* about player i is a measurable subset of the space $\Theta_i \times Y_i \times A_i \times H_i$, where H_i is the space of hierarchies of beliefs based on the state of nature and each player's information and actions. More precisely, an element of H_i is a sequence $(\mu_i^1, \mu_i^2, \ldots)$ where $\mu_i^1 \in H_i^1 = \Delta(\Theta_0 \times X_{-i} \times A_{-i})$ and then, recursively, $\mu_i^k \in \Delta(\Theta_0 \times X_{-i} \times A_{-i} \times H_{-i}^{k-1})$.[9] Note that there is no redundancy in the construction, in the sense that every two points in the state space $\Theta_0 \times (X_1 \times A_1 \times H_1) \times (X_2 \times A_2 \times H_2)$ must differ in terms of states of nature, information, actions, or beliefs thereof. Thus, our notion of expressibility is precisely that of Heifetz and Samet (1998).[10]

[5] Brandenburger and Dekel (1987) prove a related result.

[6] This canonical state space is what Mertens and Zamir (1985) call the *universal belief space*, when we take their "parameter space" to be the set of action profiles. More details on Heifetz and Samet (1998) are provided below.

[7] We assume two players for convenience; we comment more on this assumption later on in the paper.

[8] Economic examples abound: geological information and satellite photographs of a tract of land on sale are thought to be correlated with the value of the recoverable resources, expert reports on an object are thought to be correlated with the value of this object, personality traits and propensities may be thought to be correlated with ability, etc. The applied theorist who models a particular situation typically specifies these payoff-irrelevant variables.

[9] As usual, we impose a *coherency* requirement on the sequences defining H_i.

[10] In the formalism of Heifetz and Samet (1998), every subset $S \subseteq \Theta_0 \times X \times A$ is an *expression*, and if e and f are expressions, then $\neg e$, $e \cap f$ and $B_i^p(e)$ are also expressions for each $i \in I$ and $p \in [0, 1]$, which we read as "not e", "e and f" and "player i attaches probability at least p to e," respectively. Heifetz and Samet (1998) show that

We insist on solution concepts being characterized using only expressible assumptions, because the primitive terms of the language (states of nature, information, actions) and hence the derived terms (beliefs and beliefs about beliefs over the primitive terms) are suggested by the problem at hand; our basic tenet is that once *all* potentially relevant parameters of the problem are specified, we are bound to use the corresponding (primitive and derived) terms and nothing more.

To see how we characterize solution concepts, think again of complete information. In that case, rationalizability gives for each player i a subset $R_i \subseteq A_i$. The cited result of Tan and Werlang (1988) can then be stated as follows: if the sets Θ_0, X_1, X_2 are all singletons—and hence can be omitted from notation—then an action a_i belongs to R_i if and only if $(a_i, h_i) \in A_i \times H_i$ for some $h_i = (\mu_i^1, \mu_i^2, \ldots)$ such that i is rational, that is, a_i is a best reply to μ_i^1, and there is common belief of rationality at h_i, so that μ_i^2 gives probability zero to pairs (a_{-i}, μ_{-i}^1) where $-i$ is not rational, μ_i^3 gives probability zero to triplets $(a_{-i}, \mu_{-i}^1, \mu_{-i}^2)$ where $-i$ is not rational, or does not give probability one to i being rational, and so on ad infinitum. Similarly, rationalizability with incomplete information specifies a subset of A_i as a function of i's information and beliefs. More precisely, an interim notion specifies a correspondence into A_i, whose domain is either X_i or some abstract set T_i of types à la Harsanyi, whereas an ex ante notion specifies a subset of *strategies*, which are functions from X_i to A_i.[11] The exercise we perform is then entirely analogous to the one sketched above for complete information. Given a correspondence $S_i : X_i \rightrightarrows A_i$, we look for expressible assumptions in $X_i \times A_i \times H_i$ which restricted to each $x_i \in X_i$, give exactly $S_i(x_i)$. Similarly, given a correspondence $S_i : T_i \rightrightarrows A_i$, we look for expressible assumptions which, restricted to features of $t_i \in T_i$, give exactly $S_i(t_i)$.

1.4 Preview of results

Our exploration begins with *belief-free rationalizability* and Δ-*rationalizability*. The first solution concept specifies a correspondence $R_i : X_i \rightrightarrows A_i$ obtained by iterated elimination, for each $x_i = (\theta_i, y_i)$, of actions that are non-best replies, given θ_i, to some conjecture $\mu_i \in \Delta(\Theta_0 \times X_{-i} \times A_{-i})$, where the support of μ_i is accordingly restricted at every step of the procedure. The second notion generalizes this procedure, yielding a correspondence $R_i^\Delta : X_i \rightrightarrows A_i$ obtained by asking that the conjectures μ_i used for x_i belong to some postulated set $\Delta_{x_i} \subseteq \Delta(\Theta_0 \times X_{-i} \times A_{-i})$. We prove that

given any state space specifying, at each state, the players' information, actions, and beliefs about the state space itself, we can view every expression as a measurable subset of it. Conversely, an event is *expressible* if it belongs to the σ-algebra generated by the expressions, when the latter are themselves viewed as events. It can be shown that expressibility of *every* event in the state space is equivalent to non-redundancy in the sense explained above. It follows that $\Theta_0 \times (X_1 \times A_1 \times H_1) \times (X_2 \times A_2 \times H_2)$ is the *unique* (up to isomorphism) state space where all events can be seen as expressions and, conversely, every expression corresponding to some (nonempty) event in some state space, can be seen as a (nonempty) event in $\Theta_0 \times (X_1 \times A_1 \times H_1) \times (X_2 \times A_2 \times H_2)$.

[11] We limit the ex ante analysis to the case where types correspond to information that can be *learned*, that is, to the case where $T_i = X_i$ for each player i. We discuss this in more detail later on in the paper. Note that any set of functions from X_i to A_i can be seen as a set of selections from the correspondence given by the union of all their graphs; this allows a comparison between ex ante solution concepts and interim solution concepts.

Δ-rationalizability is characterized by the following assumption: (i) the players are rational, (ii) their first-order beliefs satisfy the restrictions Δ, and (iii) there is common belief in (i) and (ii). In the case of no restrictions, $\Delta_{x_i} = \Delta(\Theta_0 \times X_{-i} \times A_{-i})$, condition (ii) becomes vacuously true, thus belief-free rationalizability is characterized by rationality and common belief in rationality alone.

Then we move on to *interim correlated rationalizability (ICR)* and *interim independent rationalizability (IIR)*. These two notions, like Bayesian Nash equilibrium, require a specification of a type space à la Harsanyi, describing the players' information and beliefs about θ_0 and each other's information. Formally, this is a structure $(T_i, \vartheta_i, \upsilon_i, \pi_i)_{i\in I}$ where T_i is a space of types and $\vartheta_i : T_i \to \Theta_i$, $\upsilon_i : T_i \to Y_i$ and $\pi_i : T_i \to \Delta(\Theta_0 \times T_{-i})$ are measurable functions. Thus ICR and IIR yield correspondences $ICR_i : T_i \rightrightarrows A_i$ and $IIR_i : T_i \rightrightarrows A_i$, respectively, obtained by iterated elimination, for each $t_i \in T_i$, of actions that are non-best replies, given $\vartheta_i(t_i)$, to some $\mu_i \in \Delta(\Theta_0 \times T_{-i} \times A_{-i})$, where as before, the support of μ_i is accordingly restricted at every step of the procedure. Differently from ICR, IIR requires μ_i to satisfy a conditional independence property: conditional on t_{-i}, θ_0 and a_{-i} are independent. This is reflected in the epistemic characterizations of the two notions, which are deeply different. ICR for a type t_i is characterized by the following expressible assumption: (i) the players are rational, (ii) there is common belief in rationality, and (iii) player i's information and hierarchy of beliefs, when restricted to its payoff-relevant aspects (the payoff-relevant information $\vartheta_i(t_i)$ and the induced hierarchy of beliefs over the payoff state), is compatible with the one specified by t_i. The characterization of IIR is more difficult, and we can only give it in full when the assumed type space is non-redundant. This is because in the presence of multiple types encoding the same payoff-relevant *and* payoff-irrelevant information, as well as the same hierarchy of beliefs, we do not know how to relate distinct types to distinct expressible assumptions. Whenever the type space is non-redundant, however, even if there *is* redundancy in terms of payoff-relevant information and induced hierarchies of beliefs over the payoff state, we are able to characterize IIR for a type t_i, as follows: (i) the players are rational, (ii) each player's beliefs regard the state of nature and the other player's action independent conditional on the hierarchy of beliefs of the other player, (iii) there is common belief in (i) and (ii), and (iv) the hierarchy of beliefs is the one encoded by t_i.

In the course of our analysis, we establish a few corollaries relating Δ-rationalizability to ICR and IIR. Whenever the type space has *information types*, that is, $T_i = X_i$ for each player i, Δ-rationalizability can be viewed as a generalization of ICR or IIR, provided that the assumed restrictions are, in a natural sense, those implied by the type space. The case of information types is also the focus of our last section, in which we consider *ex ante rationalizability*. In that section we introduce two new notions, *ex ante Δ-rationalizability* and *ex ante correlated rationalizability*, which we relate to the interim solution concepts analyzed earlier. Our main result in that section is that ex ante correlated rationalizability is equivalent to ICR. Thus, contrary to interim and ex ante rationalizability, which differ because of their underlying (and different) conditional independence assumptions, their correlated versions provide the same predictions.

2 Preliminaries

The basic ingredient of our analysis is a structure $(\Theta_0, (\Theta_i, Y_i, A_i, g_i)_{i\in I})$ where Θ_0 is a finite set of *states of nature* and $I = \{1, 2\}$ is the set of *players*; each player i is endowed with a finite set A_i of feasible *actions*, and the finite sets Θ_i and Y_i represent i's payoff-relevant and payoff-irrelevant private information, respectively; we call each $\theta_i \in \Theta_i$ a *payoff type* and each $x_i = (\theta_i, y_i) \in X_i := \Theta_i \times Y_i$ an *information type* of i. Accordingly, we refer to each $\theta \in \Theta := \Theta_0 \times \Theta_1 \times \Theta_2$ as a *payoff state* and to each $x \in X := X_1 \times X_2$ as an *information state*, and we assume that each player i's utility depends on the payoff state through the function $g_i : \Theta \times A \to \mathbb{R}$, where $A = A_1 \times A_2$.

2.1 Type spaces and exogenous beliefs

Interim solution concepts specify actions for each player as a correspondence of his information type and *exogenous* beliefs—interactive beliefs about the state of nature and each other's information—often modeled using Harsanyi's (1967-68) representation: a *type space* based on $\Theta_0 \times X$, or simply X*-space*, which is a tuple $(T_i, \chi_i, \pi_i)_{i\in I}$ where each T_i is a Polish space and the functions $\chi_i : T_i \to X_i$ and $\pi_i : T_i \to \Delta(\Theta_0 \times T_{-i})$ are measurable.[12] Such a model describes the players' information and, implicitly, their hierarchies of beliefs about the state of nature and each other's information; we call these X*-hierarchies* and we define them as usual, following Mertens and Zamir (1985), whose construction we now review. For every player i, let $H^1_{X,i} = \Delta(\Theta_0 \times X_{-i})$ designate the space of *first-order* X*-beliefs*, and for all $k \geq 1$ define recursively

$$H^{k+1}_{X,i} = \Big\{(\mu^1_i, \ldots, \mu^{k+1}_i) \in H^k_{X,i} \times \Delta(\Theta_0 \times X_{-i} \times H^k_{X,-i}) : \mathrm{marg}_{\Theta_0 \times X_{-i} \times H^{k-1}_{X,-i}} \mu^{k+1}_i = \mu^k_i\Big\}. \tag{1}$$

Note that, by the coherency conditions on marginal distributions, each element of the set in (1) is determined by its last coordinate; thus, whenever convenient, for all $k \geq 1$ we identify $H^k_{X,i}$ with $\Delta(\Theta_0 \times X_{-i} \times H^{k-1}_{X,-i})$, the space of k*-order* X*-beliefs* of player i. The space of X*-hierarchies* of i is

$$H_{X,i} = \Big\{(\mu^k_i)_{k\geq 1} \in \mathop{\times}_{k\geq 1} \Delta(\Theta_0 \times X_{-i} \times H^{k-1}_{X,-i}) : (\mu^1_i, \ldots, \mu^k_i) \in H^k_{X,i} \quad \forall k \geq 1\Big\}. \tag{2}$$

This space is compact metrizable (hence Polish), and there is a homeomorphism

$$\varphi_{X,i} : H_{X,i} \to \Delta(\Theta_0 \times X_{-i} \times H_{X,-i}). \tag{3}$$

The X-hierarchies described by an X-space $(T_i, \chi_i, \pi_i)_{i\in I}$ are computed recursively: for each $i \in I$ and $t_i \in T_i$, the first-order X-belief induced by t_i is defined as follows: for each $E \subseteq \Theta_0 \times X_{-i}$,

$$\eta^1_{X,i}(t_i)[E] = \pi_i(t_i)\Big[\Big\{(\theta_0, t_{-i}) \in \Theta_0 \times T_{-i} : (\theta_0, \chi_{-i}(t_{-i})) \in E\Big\}\Big].$$

[12] For any Polish space Z we write $\Delta(Z)$ for the set of all probability measures on Z, endowed with the topology of weak convergence. Throughout the paper, a product of topological spaces is always assumed endowed with the product topology, and a subspace with its relative topology. All topological spaces are always viewed also as measurable spaces (with their Borel σ-algebra).

Then, the k-order X-belief induced by t_i is thus defined: for each measurable $E \subseteq \Theta_0 \times X_{-i} \times H^{k-1}_{X,-i}$,

$$\eta^k_{X,i}(t_i)[E] = \pi_i(t_i)\Big[\Big\{(\theta_0, t_{-i}) \in \Theta_0 \times T_{-i} : (\theta_0, \chi_{-i}(t_{-i}), \eta^{k-1}_{X,-i}(t_{-i})) \in E\Big\}\Big].$$

This gives a function $\eta_{X,i} : T_i \to H_{X,i}$ satisfying, for each $t_i \in T_i$ and measurable $E \subseteq \Theta_0 \times X_{-i} \times H_{X,-i}$,

$$\varphi_{X,i}(\eta_{X,i}(t_i))[E] = \pi_i(t_i)\Big[\Big\{(\theta_0, t_{-i}) \in \Theta_0 \times T_{-i} : (\theta_0, \chi_{-i}(t_{-i}), \eta_{X,-i}(t_{-i})) \in E\Big\}\Big]. \quad (4)$$

It is well known that $\eta_{X,i}$ is not necessarily injective, that is, there may be multiple types inducing the same X-hierarchy. More generally, the X-space is said to be *non-redundant* if for all $i \in I$ and distinct $t_i, t'_i \in T_i$, either $\chi_i(t_i) \neq \chi_i(t'_i)$ or $\eta_{X,i}(t_i) \neq \eta_{X,i}(t'_i)$. In this case, for each player i the space T_i can be seen as a measurable subset of $X_i \times H_{X,i}$ and under this identification, by (4), the tuple $(T_i)_{i\in I}$ is *belief-closed* in the sense that $\varphi_{X,i}(h_{X,i})[\Theta \times T_{-i}] = 1$ for all $t_i = (x_i, h_{X,i}) \in T_i$.[13]

Hierarchies of beliefs over the payoff state

An X-space describes, in particular, the players' payoff information and their hierarchies of beliefs about the payoff state, which we call Θ*-hierarchies*. These are defined just like X-hierarchies, but letting Θ play the role of X everywhere in (1) and (2) above. Thus $H^1_{\Theta,i} = \Delta(\Theta_0 \times \Theta_{-i})$ is the space of *first-order* Θ*-beliefs* of player i, whereas a k*-order* Θ*-belief* of player i is an element of

$$H^{k+1}_{\Theta,i} = \Big\{(\mu^1_i, \ldots, \mu^{k+1}_i) \in H^k_{\Theta,i} \times \Delta(\Theta_0 \times \Theta_{-i} \times H^k_{\Theta,-i}) \ : \ \mathrm{marg}_{\Theta_0 \times \Theta_{-i} \times H^{k-1}_{\Theta,-i}} \mu^{k+1}_i = \mu^k_i\Big\},$$

a Θ*-hierarchy* of player i is an element of

$$H_{\Theta,i} = \Big\{(\mu^k_i)_{k\geq 1} \in \mathop{\times}_{k\geq 1} \Delta(\Theta_0 \times \Theta_{-i} \times H^{k-1}_{\Theta,-i}) \ : \ (\mu^1_i, \ldots, \mu^k_i) \in H^k_{\Theta,i} \quad \forall k \geq 1\Big\} \qquad \forall i \in I,$$

and a homeomorpshim analogous to (3) exists:

$$\varphi_{\Theta,i} : H_{\Theta,i} \to \Delta(\Theta_0 \times \Theta_{-i} \times H_{\Theta,-i}).$$

To see how each type in an X-space $(T_i, \chi_i, \pi_i)_{i\in I}$ induces a Θ-hierarchy, note that for each player i we can write the function χ_i as a pair of measurable functions $(\vartheta_i, \upsilon_i)$, where $\vartheta_i : T_i \to \Theta_i$ and $\upsilon_i : T_i \to Y_i$. Then for each $i \in I$ and $t_i \in T_i$ the induced first-order Θ-belief is $\eta^1_{\Theta,i}(t_i) = \mathrm{marg}_{\Theta_0 \times \Theta_{-i}} \eta^1_{X,i}(t_i)$, and recursively, the induced k-order Θ-belief is defined as follows: for each measurable $E \subseteq \Theta_0 \times X_{-i} \times H^{k-1}_{\Theta,-i}$,

$$\eta^k_{\Theta,i}(t_i)[E] = \pi_i(t_i)\Big[\Big\{(\theta_0, t_{-i}) \in \Theta_0 \times T_{-i} : (\theta_0, \vartheta_{-i}(t_{-i}), \eta^{k-1}_{\Theta,-i}(t_{-i})) \in E\Big\}\Big].$$

Note that the function $\eta_{\Theta,i} : T_i \to H_{\Theta,i}$ thus obtained fails to be injective whenever, though not only if, the function $\eta_{X,i} : T_i \to H_{X,i}$ does so. In other words, two types inducing the same X-hierarchy must induce the same Θ-hierarchy, while two types inducing the same Θ-hierarchy may

[13] It is clear that, conversely, every tuple $(E_i)_{i\in I}$ where $E_i \subseteq X_i \times H_{X,i}$ is measurable for every $i \in I$, and which is belief-closed, can be seen as a non-redundant type space.

induce distinct X-hierarchies. This fact plays an important role in the characterization of IIR below; we provide a full characterization of IIR whenever the assumed type space is non-redundant, and only in such a case; thus, we allow for redundancies in the sense of Θ-hierarchies, though we cannot extend our characterization to the case of redundancies in the sense of X-hierarchies.

Type spaces with information types

Applied models in economics often disregard (or do not include for some other reason) payoff-irrelevant information, and assume the simplest possible type spaces, those where distinct types must differ in the payoff-relevant information that they encode; in our framework, we can define such a *payoff type space* as an X-space $(T_i, \vartheta_i, \upsilon_i, \pi_i)_{i\in I}$ where for each player i, $T_i = \Theta_i$, ϑ_i is the identity function, and υ_i is constant. In many applied models that do specify payoff-irrelevant information in a less trivial way, the analogous simplification is made, by assuming that information determines X-beliefs;[14] formally, a type space *with information types* is a type space $(T_i, \chi_i, \pi_i)_{i\in I}$ where $T_i = X_i$ and χ_i is the identity for each player i. For brevity, throughout the paper we write just $(X_i, \pi_i)_{i\in I}$ to denote such an X-space.

Besides being pervasive in applications, type spaces with information types are special in at least two other respects. First, they are always non-redundant, because distinct types must differ at least in the information type that they induce. Second, they feature the following triviality property: the X-hierarchy induced by each type is determined by its induced first-order X-belief. More precisely, we have the following remark, which is used in the proofs of Corollaries 1 and 2 below.

Remark 1. Fix an X-space with information types $(X_i, \pi_i)_{i\in I}$. For every $i \in I$, $x_i \in X_i$ and $h_{X,i} \in H_{X,i}$ such that $\mathrm{marg}_{\Theta_0\times X_{-i}}\, \varphi_{X,i}(h_{X,i}) = \pi_i(x_i)$, the conditions $h_{X,i} = \eta_{X,i}(x_i)$ and

$$\varphi_{X,i}(h_{X,i})[\Theta_0 \times \cup_{x_{-i}}\{(x_{-i}, \eta_{X,-i}(x_{-i}))\}] = 1$$

are equivalent. Indeed, the former implies the latter by belief-closedness, while the converse follows at once from coherency of $h_{X,i}$.

2.2 Endogenous beliefs and interactive epistemology

Interactive epistemology views solution concepts—correspondences from information types into actions, or from types in an X-space into actions—as reduced forms of models that explicitly describe the players' (hierarchies of) *endogenous* beliefs, that is, beliefs (and beliefs about beliefs) over actions *and* payoff states and payoff-irrelevant information. Formally, the space of *first-order A-beliefs* of player i is $H_i^1 = \Delta(\Theta_0 \times X_{-i} \times A_{-i})$, the space of *k-order A-beliefs* is

$$H_i^{k+1} = \left\{(\mu_i^1, \dots, \mu_i^{k+1}) \in H_i^k \times \Delta(\Theta_0 \times X_{-i} \times A_{-i} \times H_{-i}^k) \,:\, \mathrm{marg}_{\Theta_0\times X_{-i}\times A_{-i}\times H_{-i}^{k-1}}\, \mu_i^{k+1} = \mu_i^k\right\},$$

[14]Often it is further assumed that beliefs come from a common prior, but this is irrelevant for our analysis.

and the space of *A-hierarchies* of player i is

$$H_i = \left\{ (\mu_i^k)_{k\geq 1} \in \bigtimes_{k\geq 1} \Delta(\Theta_0 \times X_{-i} \times A_{-i} \times H_{-i}^{k-1}) \,:\, (\mu_i^1,\ldots,\mu_i^k) \in H_i^k \quad \forall k \geq 1 \right\} \qquad \forall i \in I.$$

Similarly to the spaces of X-hierarchies and Θ-hierarchies, the space of A-hierarchies is also compact metrizable, and here, too, there is a homeomorphism

$$\varphi_i : H_i \to \Delta(\Theta_0 \times X_{-i} \times A_{-i} \times H_{-i}).$$

The space of A-hierarchies of player i describes all possible beliefs that i can entertain regarding the state of nature, player $-i$'s information and action, player $-i$'s belief about the state of nature and i's information and action, and so on. In particular, each A-hierarchy embodies an X-hierarchy and a Θ-hierarchy, which we can compute naturally by recursive marginalization. In what follows we let $\varrho_{X,i} : H_i \to H_{X,i}$ and $\varrho_{\Theta,i} : H_i \to H_{\Theta,i}$ designate these mappings.[15]

An *expressible assumption* (or more simply, *assumption*) about player i is a measurable subset of $X_i \times A_i \times H_i$. A *joint assumption* is a set of the form $E = E_1 \times E_2$, where for every player i, E_i is an assumption about i. All the epistemic characterizations we provide below involve *rationality* of all players, which is the joint assumption that each player chooses an action maximizing his expected payoff given his payoff type and first-order A-beliefs. Thus, we let $RAT = \Theta_0 \times RAT_1 \times RAT_2$, where

$$RAT_i = \left\{ (\theta_i, y_i, a_i, h_i) \in X_i \times A_i \times H_i \,:\, a_i \in \operatorname*{arg\,max}_{a_i' \in A_i} \; g_i(\theta_i, a_i', \mathrm{marg}_{\Theta_0 \times \Theta_{-i} \times A_{-i}} \varphi_i(h_i)) \right\}.$$ [16]

Our characterizations involve not only rationality, but also *common belief* in rationality and possibly other assumptions. Given any joint assumption $E = E_1 \times E_2$, for every $i \in I$ let

$$B_i(E) = X_i \times A_i \times \{h_i \in H_i \,:\, \varphi_i(h_i)[\Theta_0 \times E_{-i}] = 1\} \qquad \text{and} \qquad B(E) = \Theta_0 \times B_1(E) \times B_2(E).$$ [17]

Now let $B^0(E) = E$ and recursively define $B^k(E) = B(B^{k-1}(E))$ for all $k \geq 1$. Then the joint assumption of *(correct) common belief* in E is

$$CB(E) = \cap_{k\geq 0} B^k(E).$$

For each player i, we write $CB_i(E)$ for the projection of $CB(E)$ on $X_i \times A_i \times H_i$.

[15] To see how these are formally defined, define mappings $\varrho_{X,i}^k : H_i^k \to H_{X,i}^k$ and $\varrho_{\Theta,i}^k : H_i^k \to H_{\Theta,i}^k$ for every $k \geq 1$ as follows: $\varrho_{X,i}^1(h_i^1) = \mathrm{marg}_{\Theta_0 \times X_{-i}} h_i^1$ and $\varrho_{\Theta,i}^1(h_i^1) = \mathrm{marg}_{\Theta_0 \times \Theta_{-i}} h_i^1$, and recursively, for measurable $E \subseteq \Theta_0 \times X_{-i} \times H_{X,-i}^{k-1}$,

$$\varrho_{X,i}^k(h_i^k)[E] = h_i^k\left[\left\{(\theta_0, x_{-i}, a_{-i}, h_{-i}^{k-1}) \in \Theta_0 \times X_{-i} \times A_{-i} \times H_{-i}^{k-1} : (\theta_0, x_{-i}, \varrho_{X,-i}^{k-1}(h_{-i}^{k-1})) \in E\right\}\right]$$

while for measurable $E \subseteq \Theta_0 \times \Theta_{-i} \times H_{\Theta,-i}^{k-1}$,

$$\varrho_{\Theta,i}^k(h_i^k)[E] = h_i^k\left[\left\{(\theta_0, \theta_{-i}, y_{-i}, a_{-i}, h_{-i}^{k-1}) \in \Theta_0 \times X_{-i} \times A_{-i} \times H_{-i}^{k-1} : (\theta_0, \theta_{-i}, \varrho_{\Theta,-i}^{k-1}(h_{-i}^{k-1})) \in E\right\}\right].$$

[16] Hereafter we slightly abuse notation and for all $i \in I$ we denote the linear extension of g_i to $\Delta(\Theta \times A)$ also by g_i.

[17] Note that $B(\cdot)$ maps rectangular events into rectangular events. Indeed, for our purposes it is sufficient to define mutual belief for this restricted class of events (see Battigalli and Siniscalchi, 2002).

3 Epistemic characterizations of interim solution concepts

In this section we provide epistemic characterizations of belief-free rationalizability (section 3.1), Δ-rationalizability (section 3.2), interim correlated rationalizability (section 3.3) and interim independent rationalizability (section 3.4).

3.1 Belief-free rationalizability and iterated dominance

The simplest interim solution concept that we consider takes as given the economic environment alone. The solution set specifies a correspondence $R_i : X_i \rightrightarrows A_i$ for each player i, which is defined as follows: $R_i(\theta_i, y_i) = \cap_{k\geq 0} R_i^k(\theta_i, y_i)$ for all $i \in I$ and $(\theta_i, y_i) \in X_i$, where $R_i^0(\theta_i, y_i) = A_i$ and, recursively, $R_i^k(\theta_i, y_i)$ is the set of all $a_i \in A_i$ such that for some $\mu_i \in \Delta(\Theta_0 \times X_{-i} \times A_{-i})$,

$$\operatorname{supp} \mu_i \subseteq \Theta_0 \times \left\{(x_{-i}, a_{-i}) \in X_{-i} \times A_{-i} : a_{-i} \in R_{-i}^{k-1}(x_{-i})\right\},$$

$$a_i \in \arg\max_{a_i' \in A_i} \sum_{(\theta_0, \theta_{-i}, a_{-i}) \in \Theta_0 \times \Theta_{-i} \times A_{-i}} \mu_i[\{(\theta_0, \theta_{-i}, a_{-i})\} \times Y_{-i}]\, g_i(\theta_0, \theta_i, \theta_{-i}, a_i', a_{-i}).$$

This version of rationalizability, which is *belief-free* in that its computation does not need a specification of X-beliefs of any sort, is equivalent to the following *interim iterated dominance* procedure:[18] $a_i \in R_i^k(\theta_i, y_i)$ if and only if there does not exist $\alpha_i \in \Delta(R_i^{k-1}(\theta_i, y_i))$ such that

$$g_i(\theta_0, \theta_i, \theta_{-i}, \alpha_i, a_{-i}) > g_i(\theta_0, \theta_i, \theta_{-i}, a_i, a_{-i}) \qquad \forall (\theta_0, \theta_{-i}, y_{-i}) \in \Theta_0 \times X_{-i},\ \forall a_{-i} \in R_{-i}^{k-1}(\theta_{-i}, y_{-i}).$$

Note that the payoff-irrelevant information plays no role here: any two information types specifying the same payoff type have the same set of belief-free rationalizable actions. Indeed, rationality itself has nothing to do with payoff-irrelevant information, and belief-free rationalizability for an information type is the consequence of rationality and common certainty of rationality alone, given the payoff information that it specifies. Thus, belief-free rationalizability is characterized by both of the following:[19]

$$R_i(x_i) = \operatorname{proj}_{A_i} CB_i(RAT) \cap [x_i] \qquad \forall i \in I,\ \forall x_i \in X_i, \tag{5}$$

$$R_i(x_i) = \operatorname{proj}_{A_i} CB_i(RAT) \cap [\theta_i] \qquad \forall i \in I,\ \forall x_i = (\theta_i, y_i) \in X_i, \tag{6}$$

where $[\theta_i]$ and $[x_i]$ are the assumptions about player i defined as follows:

$$[\theta_i] = \{\theta_i\} \times Y_i \times A_i \times H_i\,, \qquad [x_i] = \{x_i\} \times A_i \times H_i \qquad \forall \theta_i \in \Theta_i,\ \forall x_i \in X_i.$$

[18] This extends the classical iterated dominance characterization of rationalizability in complete information games due to Pearce (1984)—see Battigalli (2003). The procedure has been used by Bergemann and Morris (2009) to define *iterative implementation* and prove that it is equivalent to robust (or type-space-independent) implementation.

[19] Note that neither directly implies the other. Indeed, (6) is equivalent to $R_i(\theta_i, y_i) = \operatorname{proj}_{A_i} \cup_{y_i' \in Y_i} CB_i(RAT) \cap [(\theta_i, y_i')]$ for all $(\theta_i, y_i) \in X_i$ and therefore it leaves open the possibility that (5) is violated for some $x_i \in X_i$. On the other hand, obtaining (6) from (5) requires the observation that the belief-free rationalizable actions of an information type only depends on its payoff type. The analogous remark applies to the alternative characterizations of Δ-rationalizability that we provide in (7) and (8) below.

3.2 Δ-rationalizability

The notion of Δ-*rationalizability* is also meant to capture strategic reasoning in the assumed economic environment with no reference to type spaces. It generalizes the belief-free approach described above—see Battigalli (2003) and Battigalli and Siniscalchi (2003, 2007). The solution concept specifies a correspondence $R_i^{\Delta} : X_i \rightrightarrows A_i$ for each player i, taking as given a profile Δ of information-dependent first-order restrictions: formally, $\Delta = ((\Delta_{x_i})_{x_i \in X_i})_{i \in I}$ where $\Delta_{x_i} \subseteq \Delta(\Theta_0 \times X_{-i} \times A_{-i})$ is a nonempty closed set for each information type x_i of each player i. The set of Δ-rationalizable actions of $(\theta_i, y_i) \in X_i$ is defined as follows: $R_i^{\Delta}(\theta_i, y_i) = \cap_{k \geq 0} R_i^{\Delta,k}(\theta_i, y_i)$, where $R_i^{\Delta,0}(\theta_i, y_i) = A_i$ and, recursively, $R_i^{\Delta,k}(\theta_i, y_i)$ is the set of all $a_i \in A_i$ such that for some $\mu_i \in \Delta_{(\theta_i, y_i)}$,

$$\begin{aligned} \operatorname{supp} \mu_i \ &\subseteq \ \Theta_0 \times \left\{(x_{-i}, a_{-i}) \in X_{-i} \times A_{-i} : a_{-i} \in R_{-i}^{\Delta,k-1}(x_{-i})\right\}, \\ a_i \ &\in \ \arg\max_{a_i' \in A_i} \sum_{(\theta_0, \theta_{-i}, a_{-i}) \in \Theta_0 \times \Theta_{-i} \times A_{-i}} \mu_i[\{(\theta_0, \theta_{-i}, a_{-i})\} \times Y_{-i}]\, g_i(\theta_0, \theta_i, \theta_{-i}, a_i', a_{-i}). \end{aligned}$$

Note that with trivial restrictions, i.e. $\Delta_{x_i} = \Delta(\Theta_0 \times X_{-i} \times A_{-i})$ for all $i \in I$ and $x_i \in X_i$, this reduces to belief-free rationalizability. As we prove below, ICR and IIR on type spaces with information types are also special cases of Δ-rationalizability. Before proceeding, let us record here the epistemic characterization of Δ-rationalizability due to Battigalli and Siniscalchi (2007, Proposition 1). Define the assumptions

$$[\Delta_{x_i}] = X_i \times A_i \times \{(\mu_i^1, \mu_i^2, \dots) \in H_i : \mu_i^1 \in \Delta_{x_i}\} \qquad \forall i \in I,\ \forall x_i \in X_i$$

and the joint assumption $[\Delta] = \Theta_0 \times [\Delta_1] \times [\Delta_2]$, where $[\Delta_i]$ is the assumption that player i satisfies the restrictions, whatever his information type, that is,

$$[\Delta_i] \ = \ \bigcup_{x_i \in X_i} ([x_i] \cap [\Delta_{x_i}]) \qquad \forall i \in I.$$

Then Δ-rationalizability is characterized by the following generalization of (5):

$$R_i^{\Delta}(x_i) \ = \ \operatorname{proj}_{A_i} CB_i(RAT \cap [\Delta]) \cap [x_i] \cap [\Delta_{x_i}] \qquad \forall i \in I,\ \forall x_i \in X_i. \tag{7}$$

Thus, Δ-rationalizability corresponds to the assumption that players are rational, their information and first-order beliefs satisfy the restrictions Δ, and there is common belief in these two facts. As a matter of fact, analogously to belief-free rationalizability, Δ-rationalizability for player i depends on his information only through the corresponding payoff type and restrictions. Thus, a generalization of (6) also holds:

$$R_i^{\Delta}(x_i) \ = \ \operatorname{proj}_{A_i} CB_i(RAT \cap [\Delta]) \cap [\theta_i] \cap [\Delta_{x_i}] \qquad \forall i \in I,\ \forall x_i = (\theta_i, y_i) \in X_i. \tag{8}$$

3.3 Interim correlated rationalizability

The solution concept of *interim correlated rationalizability* or *ICR*, introduced by Dekel, Fudenberg, and Morris (2007), applies to the Bayesian game induced by an X-space $(T_i, \vartheta_i, \upsilon_i, \pi_i)_{i \in I}$. The solution set specifies for each player i a correspondence $ICR_i : T_i \rightrightarrows A_i$ which is defined as follows:[20] $ICR_i(t_i) = \cap_{k\geq 0} ICR_i^k(t_i)$, where $ICR_i^0(t_i) = A_i$ and, recursively, $ICR_i^k(t_i)$ is the set of all $a_i \in A_i$ for which there exists a measurable function $\sigma_{-i} : \Theta_0 \times T_{-i} \to \Delta(A_{-i})$ such that

$$\operatorname{supp} \sigma_{-i}(\theta_0, t_{-i}) \subseteq ICR_{-i}^{k-1}(t_{-i}) \qquad \text{for } \pi_i(t_i)\text{-almost every } (\theta_0, t_{-i}) \in \Theta \times T_{-i}, \tag{9}$$

$$a_i \in \arg\max_{a_i' \in A_i} \int_{\Theta_0 \times T_{-i}} g_i(\theta_0, \vartheta_i(t_i), \vartheta_{-i}(t_{-i}), a_i', \sigma_{-i}(\theta_0, t_{-i}))\, \pi_i(t_i)[d\theta_0 \times dt_{-i}]. \tag{10}$$

The intuition for this solution concept is that type t_i of player i forms a probabilistic conjecture σ_{-i} on how the behavior of $-i$ depends on her type t_{-i} and on the state of nature θ_0, possibly taking into account an implicit correlation device. Indeed, as its name suggests, ICR allows the possibility that according to the probability distribution over $\Theta_0 \times T_{-i} \times A_{-i}$ induced by $\pi_i(t_i)$ and σ_{-i},[21] the state of nature θ_0 and the opponent's action a_{-i} are correlated, even after conditioning on $-i$'s type.

The conjecture σ_{-i} must rationalize a_i in the sense of (10), and it must be itself rationalizable in the sense of being supported by rationalizable actions, as specified by (9), but is otherwise unrestricted. This is reflected in the following theorem, which proves that ICR reflects rationality and common belief in rationality alone, given the Θ-hierarchies induced by the assumed type space. More precisely, given an X-space and a type t_i of player i in it, consider the following assumption: player i is rational, commonly believes in the rationality of all players, and his payoff type and Θ-hierarchy are those induced by t_i. The theorem below states that ICR for t_i captures the behavioral consequences of this assumption, and indeed of any other, stronger assumption obtained by restricting i's exogenous information or beliefs (while keeping the payoff type and Θ-hierarchy induced by t_i, of course).

For each player i, let $\mathcal{E}_i$ be the σ-algebra of *exogenous assumptions* about i, namely, the family of all (measurable) subsets $E_i \subseteq X_i \times A_i \times H_i$ which can be seen as subsets of $X_i \times H_{X,i}$ in the sense that, for some nonempty, measurable $F_i \subseteq X_i \times H_{X,i}$,

$$E_i = \Big\{(x_i, a_i, h_i) \in X_i \times A_i \times H_i : (x_i, \varrho_{X,i}(h_i)) \in F_i\Big\}.$$

Given a type space $(T_i, \vartheta_i, \upsilon_i, \pi_i)_{i\in I}$ we say that an exogenous assumption $E_i \in \mathcal{E}_i$ is *compatible* with type $t_i \in T_i$ of player i, provided that

$$\theta_i = \vartheta_i(t_i) \quad \text{and} \quad \varrho_{\Theta,i}(h_i) = \eta_{\Theta,i}(t_i) \qquad \forall (\theta_i, y_i, a_i, h_i) \in E_i. \tag{11}$$

In words, E_i is an exogenous assumption compatible with t_i if all of its elements specify the payoff type and Θ-hierarchy induced by t_i, and if it does not exclude any action or A-hierarchy whose

[20] The sets Θ_i and Y_i are singletons (and hence do not appear at all) in Dekel, Fudenberg, and Morris (2007). However, their definitions and results extend seamlessly to our framework. In particular, they prove a result (Proposition 2) similar to Theorem 1, and they also prove that any two types (possibly from different type spaces) mapping into the same Θ-hierarchy have the same ICR actions, which obtains here as an obvious consequence of Theorem 1.

[21] This is the measure ν_i such that $\nu_i[\{\theta_0\} \times E_{-i} \times \{a_{-i}\}] = \int_{E_{-i}} \sigma_{-i}(\theta_0, t_{-i})[a_{-i}]\pi_i(t_i)[\theta_0 \times dt_{-i}]$ for every $(\theta_0, a_{-i}) \in \Theta_0 \times A_{-i}$ and measurable $E_{-i} \subseteq T_{-i}$.

induced X-hierarchy is not itself excluded. Note that the largest exogenous assumption compatible with t_i is

$$\{\vartheta_i(t_i)\} \times Y_i \times A_i \times \{h_i \in H_i : \varrho_{\Theta,i}(h_i) = \eta_{\Theta,i}(t_i)\},$$

which simply says that i's payoff type and Θ-hierarchy are those specified by t_i, whereas a minimal such assumption has the following form: for some $y_i \in Y_i$ and $h_{X,i} \in H_{X,i}$ with $(\varrho_{X,i})^{-1}(h_{X,i}) \subseteq (\varrho_{\Theta,i})^{-1}(\eta_{\Theta,i}(t_i))$,

$$\{\vartheta_i(t_i)\} \times \{y_i\} \times A_i \times \{h_i \in H_i : \varrho_{X,i}(h_i) = h_{X,i}\},$$

which says that i has payoff type $\vartheta_i(t_i)$, some fixed payoff-irrelevant information y_i (possibly different from $\upsilon_i(t_i)$) and some fixed X-hierarchy $h_{X,i}$ (possibly different from $\eta_{X,i}(t_i)$) whose induced Θ-hierarchy is the same as the one induced by t_i. Now we are ready to characterize ICR.

Theorem 1. *Fix a type space $(T_i, \chi_i, \pi_i)_{i\in I}$. For all $i \in I$, $t_i \in T_i$, and $E_i \in \mathcal{E}_i$ compatible with t_i,*

$$ICR_i(t_i) = \mathrm{proj}_{A_i} CB_i(RAT) \cap E_i. \tag{12}$$

Proof. See Appendix A.1. ■

Interim correlated rationalizability and Δ-rationalizability coincide in the case of a type space with information types $(X_i, \pi_i)_{i\in I}$, whenever Δ specifies (only) the restrictions *derived* from it, that is, whenever

$$\Delta_{x_i} = \left\{ \mu_i \in \Delta(\Theta_0 \times X_{-i} \times A_{-i}) \ : \ \mathrm{marg}_{\Theta_0 \times X_{-i}} \mu_i = \pi_i(x_i) \right\} \qquad \forall i \in I,\ \forall x_i \in X_i.$$

This follows from Theorem 1, using (7), as we now show.[22]

Corollary 1. *Fix a type space with information types $(X_i, \pi_i)_{i\in I}$ and let Δ be the set of restrictions derived from it. Then*

$$ICR_i(x_i) = R_i^{\Delta}(x_i) \qquad \forall i \in I,\ \forall x_i \in X_i.$$

Proof. Since $CB_i(RAT \cap [\Delta]) = CB_i(RAT) \cap CB_i([\Delta])$, by (7) and Theorem 1 it suffices to show that $CB_i([\Delta]) \cap [x_i] \cap [\Delta_{x_i}]$ is an exogenous assumption compatible with x_i. This, in turn, follows from

$$\{x_i\} \times A_i \times \{h_i \in H_i : \varrho_{X,i}(h_i) = \eta_{X,i}(x_i)\} = CB_i([\Delta]) \cap [x_i] \cap [\Delta_{x_i}],$$

which we now prove. Since $\{x_i\} \times A_i \times \{h_i \in H_i : \mathrm{marg}_{\Theta_0 \times X_{-i}} \varphi_i(h_i) = \pi_i(x_i)\} = [x_i] \cap [\Delta_{x_i}]$ and the analogous holds for player $-i$, it suffices to show: for all $h_i \in H_i$ with $\mathrm{marg}_{\Theta_0 \times X_{-i}} \varphi_i(h_i) = \pi_i(x_i)$, the conditions $\varrho_{X,i}(h_i) = \eta_{X,i}(x_i)$ and $\sum_{x_{-i}\in X_{-i}} \varphi_i(h_i)[\Theta_0 \times ([x_{-i}] \cap [\eta_{X,-i}(x_{-i})]] = 1$ are equivalent. Indeed, this follows at once from belief-closedness and coherency (see Remark 1). ■

[22] Corollary 1 can be also proved directly. Indeed, if Δ is the set of restrictions derived from a type space with information types $(X_i, \pi_i)_{i\in I}$, then Δ_{x_i} is precisely the set of probability distributions on $\Theta_0 \times X_{-i} \times A_{-i}$ induced by $\pi_i(x_i)$ and some conjecture $\sigma_{-i} : \Theta_0 \times X_{-i} \to \Delta(A_{-i})$.

3.4 Interim independent rationalizability

The notion of *interim independent rationalizability* or *IIR*—see Ely and Pęski (2006)—also applies to the Bayesian game induced by a type space $(T_i, \vartheta_i, \upsilon_i, \pi_i)_{i\in I}$. Similarly to ICR, it specifies for each player i a correspondence $IIR_i : T_i \rightrightarrows A_i$ thus defined: $IIR_i(t_i) = \cap_{k\geq 0} IIR_i^k(t_i)$, where $IIR_i^0(t_i) = A_i$ and, recursively, $IIR_i^k(t_i)$ is the set of all $a_i \in A_i$ such that there exists a measurable function $\sigma_{-i} : T_{-i} \to \Delta(A_{-i})$ such that

$$\operatorname{supp} \sigma_{-i}(t_{-i}) \subseteq IIR_{-i}^{k-1}(t_{-i}) \qquad \text{for } \pi_i(t_i)\text{-almost every } t_{-i} \in T_{-i}, \tag{13}$$

$$a_i \in \arg\max_{a_i' \in A_i} \int_{\Theta_0 \times T_{-i}} g_i(\theta_0, \vartheta_i(t_i), \vartheta_{-i}(t_{-i}), a_i', \sigma_{-i}(t_{-i}))\, \pi_i(t_i)[d\theta_0 \times dt_{-i}]. \tag{14}$$

Remark 2. If the type space is finite, then $a_i \in IIR_i(t_i)$ if and only if a_i is rationalizable for the corresponding player/type t_i in the associated *interim strategic form*,[23] where the set of players is $T_1 \cup T_2$ and the set of available actions of each player/type t_i is A_i. Indeed, in this game the payoff to player/type t_i when choosing action a_i depends only on the actions chosen by the players/types in T_{-i}, and given a mixed action profile $\sigma_{-i} : T_{-i} \to \Delta(A_{-i})$ for such types, it is defined as

$$g_{t_i}(a_i, \sigma_{-i}) = \sum_{(\theta_0, t_{-i}) \in \Theta_0 \times T_{-i}} \pi_i(t_i)[\theta_0, t_{-i}] g_i(\theta_0, \vartheta_i(t_i), \vartheta_{-i}(t_{-i}), a_i, \sigma_{-i}(t_{-i})).$$

Thus, the set of actions that are rationalizable in the interim strategic form for player/type t_i is $ISFR_i(t_i) = \cap_{k\geq 0} ISFR_i^k(t_i)$, where $ISFR_i^0(t_i) = A_i$ and $ISFR_i^k(t_i)$ is the set of all $a_i \in A_i$ for which there is $\sigma_{-i} : T_{-i} \to \Delta(A_{-i})$ with $\operatorname{supp} \sigma_{-i}(t_{-i}) \subseteq ISFR_{-i}^{k-1}(t_{-i})$ for all $t_{-i} \in T_{-i}$ and, moreover, $a_i \in \arg\max_{a_i' \in A_i} g_{t_i}(a_i, \sigma_{-i})$. Since $IIR_i^0(t_i) = ISFR_i^0(t_i) = A_i$, an obvious induction shows that these requirements are the same as (13) and (14) above, and hence that $IIR_i(t_i) = ISFR_i(t_i)$. ■

Formally, the only difference from ICR is that the conjecture σ_{-i} used by t_i to rationalize a_i cannot depend on the state of nature; under the probability distribution on $\Theta_0 \times T_{-i} \times A_{-i}$ induced by $\pi_i(t_i)$ and σ_{-i} the conditional probabilities of $-i$'s actions given $-i$'s type do not depend on θ_0. Indeed, it is clear that the two notions coincide if there is only one state of nature, as is the case in many economic applications;[24] we record this fact in the following:

Remark 3. Assume that there is *distributed knowledge* of the payoff state, i.e. that Θ_0 is a singleton. Then $ICR_i(t_i) = IIR_i(t_i)$ for every type space $(T_i, \vartheta_i, \upsilon_i, \pi_i,)_{i\in I}$ and every $i \in I$, $t_i \in T_i$. Thus, by Theorem 1, $IIR_i(t_i) = \operatorname{proj}_{A_i} CB_i(RAT) \cap E_i$ for all $i \in I$, $t_i \in T_i$ and $E_i \in \mathcal{E}_i$ compatible with t_i. ■

[23]This is *independent* rationalizability on the interim strategic form of the Bayesian game. But, by Kuhn's (1953) equivalence result, with $I = \{1, 2\}$, correlated and independent rationalizability on the interim strategic form are equivalent (T_{-i} is like a coalition with perfect recall in the extensive form of the Bayesian game).

[24]Models with private values are an obvious example, but also many models with interdependent values satisfy this property. For example, consider "wallet games" (Klemperer, 1998), or any model where θ_i specifies player i's characteristics such as ability or riskiness, and the consequences for each player of an action profile depend on all players' characteristics.

In general, however, IIR and ICR differ, and the characterization of IIR in the latter remark fails. To be sure, the definition of IIR, just like the definition of ICR, makes no reference to the mappings $(\upsilon_i)_{i\in I}$, but unlike with ICR, where the exact specification of these mappings (and hence of anything beyond the induced Θ-hierarchies) is also entirely irrelevant for the *solution*, the IIR actions of a type do *not* depend just on its induced Θ-hierarchy. This is precisely because the independence between the state of nature and the opponent's action embodied in IIR is conditional on the opponent's *type*, not just on her Θ-hierarchy (and in the case of a redundant type space, not even on her X-hierarchy).

As we have argued earlier, we do not know how to express what it means for two different types to choose different actions, if the difference between the types *themselves* is not expressible in the given language, i.e. if the two types induce the same X-hierarchy. Therefore, our task here is to provide a full characterization of IIR for those cases where such differences can be traced to expressible features of the types, namely, for non-redundant type spaces. These can be seen as belief-closed sets of hierarchies, hence independence conditional on the opponent's type does correspond, in those cases, to an expressible assumption. Now we formalize this assumption and then state our characterization result, which says that IIR is the expression of rationality, conditional independence, and common belief thereof, given the X-hierarchies induced by the type space.

For every player i, let $H_{i,CI}$ designate the set of all $h_i \in H_i$ such that, according to the belief $\varphi_i(h_i)$, the state of nature and the action of player $-i$ are independent, conditional on every exogenous assumption about player $-i$. Formally, $h_i \in H_{i,CI}$ provided that for all $\theta_0 \in \Theta_0$ and $a_{-i} \in A_{-i}$ the condition

$$\varphi_i(h_i)[\theta_0, a_{-i} \,|\, \mathcal{E}_{-i}](\cdot) \;=\; \varphi_i(h_i)[\theta_0 \,|\, \mathcal{E}_{-i}](\cdot)\; \varphi_i(h_i)[a_{-i} \,|\, \mathcal{E}_{-i}](\cdot) \tag{15}$$

holds $\varphi_i(h_i)$-almost everywhere, with

$$\varphi_i(h_i)[\,\cdot \,|\, \mathcal{E}_{-i}](\cdot) : \Theta_0 \times X_{-i} \times A_{-i} \times H_{-i} \to \Delta(\Theta_0 \times X_{-i} \times A_{-i} \times H_{-i})$$

denoting any regular conditional probability given the measure $\varphi_i(h_i)$ and the σ-algebra $\mathcal{E}_{-i}$. Such regular conditional probability exists because $\Theta_0 \times X_{-i} \times A_{-i} \times H_{-i}$ is a Polish space (see Dudley, 2002, p. 345). Moreover, as we establish in Appendix A.2, the set $H_{i,CI}$ does not depend on the particular version of conditional probability that we choose, and furthermore, it is measurable. Thus, we can define the joint assumption $CI = \Theta_0 \times CI_1 \times CI_2$, where $CI_i = X_i \times A_i \times H_{i,CI}$ for every $i \in I$. Now let

$$[h_{X,i}] = X_i \times A_i \times \{h_i \in H_i : \varrho_{X,i}(h_i) = h_{X,i}\} \qquad \forall i \in I,\ \forall h_{X,i} \in H_{X,i}.$$

With these definitions, we can characterize IIR.

Theorem 2. *Fix a type space $(T_i, \vartheta_i, \upsilon_i, \pi_i)_{i\in I}$. For every $i \in I$ and $t_i \in T_i$,*

$$IIR_i(t_i) \;\supseteq\; \mathsf{proj}_{A_i}\, CB_i(RAT \cap CI) \cap [\vartheta_i(t_i)] \cap [\eta_{X,i}(t_i)].$$

Furthermore, if the type space is non-redundant, then

$$IIR_i(t_i) \;\subseteq\; \mathsf{proj}_{A_i}\, CB_i(RAT \cap CI) \cap [(\vartheta_i(t_i), \upsilon_i(t_i))] \cap [\eta_{X,i}(t_i)].$$

Proof. See Appendix A.3. ∎

Similarly to what we showed for ICR, we can identify IIR with Δ-rationalizability for an X-space with information types $(X_i, \pi_i)_{i\in I}$, provided that the restrictions Δ are derived from the X-space and, in addition, they embody independence between the state of nature and the opponent's action, conditional on the player's information. Let us say that $\mu_i \in \Delta(\Theta_0 \times X_{-i} \times A_{-i})$ satisfies *information-based conditional independence* if

$$\mu_i[x_{-i}] > 0 \Rightarrow \mu_i[\theta_0, a_{-i}|x_{-i}] = \mu_i[\theta_0|x_{-i}]\mu_i[a_{-i}|x_{-i}] \qquad \forall(\theta_0, x_{-i}, a_{-i}) \in \Theta_0 \times X_{-i} \times A_{-i}.$$

Let $\Delta_{i,\mathrm{CI}}$ denote this set of first-order beliefs, and say that Δ is *CI-derived* from $(X_i, \pi_i)_{i\in I}$ if

$$\Delta_{x_i} = \{\, \mu_i \in \Delta_{i,CI} \;:\; \mathrm{marg}_{\Theta_0 \times X_{-i}} \mu_i = \pi_i(x_i) \,\} \qquad \forall i \in I,\ \forall x_i \in X_i.$$ [25]

Corollary 2. *Fix a type space with information types $(X_i, \pi_i)_{i\in I}$ and let Δ be the set of restrictions CI-derived from it. Then*

$$IIR_i(x_i) = R_i^{\Delta}(x_i) \qquad \forall i \in I,\ \forall x_i \in X_i.$$

Proof. Since $CB_i(RAT\cap[\Delta]) = CB_i(RAT)\cap CB_i([\Delta])$ and $CB_i(RAT\cap CI) = CB_i(RAT)\cap CB_i(CI)$, by (7) and Theorem 2 it suffices to prove $CB_i([\Delta]) \cap [x_i] \cap [\Delta_{x_i}] = CB_i(CI) \cap [x_i] \cap [\eta_{X,i}(x_i)]$. Indeed, we have $[x_i] \cap [\Delta_{x_i}] = CI_i \cap (\{x_i\} \times A_i \times \{h_i \in H_i : \mathrm{marg}_{\Theta_0 \times X_{-i}} \varphi_i(h_i) = \pi_i(x_i)\})$, and analogously for player $-i$. Thus, the claim follows by the same argument as in the proof of Corollary 1. ∎

4 Ex ante rationalizability

In this section we show that the differences between rationalizability in the ex ante and interim strategic form of a Bayesian game are due to the different independence restrictions that are embodied in these solution concepts. This follows from a preliminary result about Δ-rationalizability that helps clarifying the conceptual issue; given any set Δ of information-dependent restrictions on beliefs, we define a notion of *ex ante* correlated Δ-rationalizability, and we show that it is in a strong sense equivalent to the interim notion of Δ-rationalizability introduced earlier. Then we define a notion of ex ante correlated rationalizability that is equivalent to ICR in the same sense.

4.1 Ex ante Δ-rationalizability

Consider the point of view of player i in an ex ante stage where he does not know x_i yet, and let S_i be the set of all functions from X_i to A_i. Then we can define a *structural ex ante strategic form*

[25] Analogously to our remark in Footnote 22, here we note that the equivalence between IIR and Δ-rationalizability (Corollary 2) can be proved directly. If Δ is the set of restrictions CI-derived from a type space with information types $(X_i, \pi_i)_{i\in I}$, then Δ_{x_i} is the set of probability distributions on $\Theta_0 \times X_{-i} \times A_{-i}$ induced by $\pi_i(x_i)$ and some conjecture $\sigma_{-i} : \Theta_0 \times X_{-i} \to \Delta(A_{-i})$ satisfying conditional independence, that is, $\sigma_{-i}(\theta_0, x_{-i}) = \sigma_{-i}(\theta_0', x_{-i})$ for all $\theta_0, \theta_0' \in \Theta_0$ and $x_{-i} \in X_{-i}$.

with two real players, 1 and 2, choosing strategies in S_1 and S_2, respectively, and a fictitious player choosing an element of $\Theta_0 \times X$, with the payoff function $\bar{g}_i : \Theta_0 \times X \times S_1 \times S_2 \to \mathbb{R}$ of each player i defined by

$$\bar{g}_i(\theta_0, \theta_1, y_1, \theta_2, y_2, s_1, s_2) = g_i(\theta_0, \theta_1, \theta_2, s_1(\theta_1, y_1), s_2(\theta_2, y_2)).$$

Now fix a set of restrictions $\Delta_i = (\Delta_{x_i})_{x_i \in X_i}$ where $\Delta_{x_i} \subseteq \Delta(\Theta_0 \times X_{-i} \times A_{-i})$ for every $x_i \in X_i$. This entails restrictions on the ex ante belief $\mu_i \in \Delta(\Theta_0 \times X \times S_{-i})$ that player i can entertain about the fictitious player's choice and the strategy of $-i$. Thus, μ_i is *consistent* with Δ_i if μ_i assigns positive probability to every x_i,[26] and moreover, conditional on x_i, it yields interim beliefs in Δ_{x_i}, that is,

$$\mu_i[x_i] > 0 \quad \text{and} \quad \bar{\mu}_i[\cdot|x_i] \in \Delta_{x_i} \qquad \forall i \in I,\ \forall x_i \in X_i,$$

where $\bar{\mu}_i$ is the probability distribution on $\Theta_0 \times X \times A_{-i}$ induced by μ_i, namely

$$\bar{\mu}_i[\theta_0, x_i, x_{-i}, a_{-i}] = \mu_i[\{(\theta_0, x_i, x_{-i})\} \times \{s_{-i} \in S_{-i} : s_{-i}(x_{-i}) = a_{-i}\}].$$

The set of *ex ante* Δ-*rationalizable* strategies is thus defined: $AR_i^{\Delta} = \cap_{k \geq 0} AR_i^{\Delta,k}$, where $AR_i^{\Delta,0} = S_i$ and $AR_i^{\Delta,k}$ is the set of all $s_i \in S_i$ such that, for some $\mu_i \in \Delta(\Theta_0 \times X \times S_{-i})$ consistent with Δ_i,[27]

$$\operatorname{supp} \mu_i \subseteq \Theta_0 \times X \times AR_{-i}^{\Delta,k-1}, \tag{16}$$

$$s_i \in \arg\max_{s_i' \in S_i} \sum_{(\theta_0, x, s_{-i}) \in \Theta_0 \times X \times S_{-i}} \mu_i[\theta_0, x, s_{-i}]\, \bar{g}_i(\theta_0, x, s_i', s_{-i}). \tag{17}$$

Note that μ_i may exhibit correlation between the fictitious player and player $-i$.

In order to relate ex ante Δ-rationalizability with Δ-rationalizability, observe that given a correspondence $F_i : X_i \rightrightarrows A_i$ and a set $S_i' \subseteq S_i$, it makes sense to consider S_i' and F_i equivalent, and write $S_i' \approx F_i$, whenever S_i' is precisely the set of selections from F_i. Thus

$$S_i' \approx F_i \quad \text{if and only if} \quad S_i' = \left\{ s_i \in S_i \ :\ s_i(x_i) \in F_i(x_i) \ \ \forall x_i \in X_i \right\}.$$

As the following result shows, this is precisely the sense in which the Δ-rationalizability correspondence $R_i^{\Delta} : X_i \rightrightarrows A_i$ and the ex ante Δ-rationalizable strategies AR_i^{Δ} are equivalent.

Proposition 1. *$AR_i^{\Delta} \approx R_i^{\Delta}$ for every $i \in I$.*

Proof. See Appendix A.4. ∎

[26]We impose this weak requirement to derive well-defined interim beliefs and avoid tedious issues concerning the differences between ex ante and interim expected payoff maximization. Alternatively, we could impose a perfection requirement (see Brandenburger and Dekel, 1987). This discussion would distract the reader's attention from the important issues.

[27]Adapting the argument Battigalli and Siniscalchi (2007) use to prove their Proposition 1, one can show that AR_i^{Δ} is the set of ex ante structural strategic form strategies of i that are consistent with (correct) common belief in rationality and in the restrictions Δ.

4.2 Ex ante correlated rationalizability

Corollary 1 and Proposition 1 yield an equivalence result for ex ante and interim correlated rationalizability in Bayesian games with information types. Before stating the result formally, let us first review the standard notion of ex ante rationalizability. We restrict our attention to the case in which Harsanyi types represent information that can be learned, that is, the case of information types. However, we remark that an equivalence result like the one stated below can be proved for *every* Bayesian game.

A strategy for the Bayesian game induced by a type space with information types $(X_i, \pi_i)_{i\in I}$ is ex ante rationalizable if it is rationalizable in the ex ante strategic form of the game. To define the ex ante strategic form, we must first specify ex ante beliefs on $\Theta_0 \times X$ consistent with the type space. Thus, we say that a prior $\Pi_i \in \Delta(\Theta_0 \times X)$ is *consistent* with $(X_i, \pi_i)_{i\in I}$ if

$$\Pi_i[x_i] > 0,^{28} \quad \text{and} \quad \Pi_i[\cdot|x_i] = \pi_i(x_i)[\cdot] \qquad \forall x_i \in X_i.$$

Once we fix a consistent prior Π_i for each player i, the *ex ante strategic form* of the induced Bayesian game is given by the expected payoff functions

$$\bar{g}_i^{\Pi_i}(s_1, s_2) \;=\; \sum_{(\theta_0, x)\in\Theta_0\times X} \Pi_i[\theta_0, x]\; \bar{g}_i(\theta_0, x, s_1, s_2) \qquad \forall i \in I.$$

It can be verified that the rationalizable strategies in this game do not depend on the particular priors Π_1, Π_2 that we fix, as long as we they are consistent with the given type space.

It is also standard to show that ex ante rationalizability implicitly relies on an *ex ante independence* assumption: a player's beliefs about (θ_0, x) and s_{-i} are given by the product measure $\Pi_i \times \mu_i$ where $\mu_i \in \Delta(S_{-i})$. Indeed, anticipating the next definition, this amounts to choosing a conjecture σ_{-i} which is a *constant* function. Ex ante independence implies interim independence, hence ex ante rationalizability implies interim independent rationalizability, or equivalently, rationalizability in the interim strategic form of the Bayesian game—see Remark 2.[29]

We now define a notion of ex ante correlated rationalizability that removes the said ex ante independence assumption. Fix a type space with information types $(X_i, \pi_i)_{i\in I}$ and priors Π_1, Π_2 consistent with it. For each player i the set of *ex ante correlated rationalizable* strategies is defined as $ACR_i = \cap_{k\geq 0} ACR_i^k$, where $ACR_i^0 = S_i$ and for all $k \geq 1$, recursively, ACR_i^k is the set of all $s_i \in S_i$

[28]As before, we include this essentially innocuous requirement just to avoid distracting the reader.

[29]The difference between ex ante and interim rationalizability is related to the difference between two notions of *extensive form rationalizability*: the more restrictive one assumes that a player has an initial conjecture about the opponent's strategy, which may be revised only after receiving some information about the opponent's behavior; the less restrictive, adopted by Pearce (1984), drops the initial conjecture and allows a player to have different conjectures at different information sets even if they only reflect information about chance moves. When we consider the extensive form of a static Bayesian game, the first solution concept yields ex ante rationalizability and the second one yields interim rationalizability. To the best of our knowledge, Battigalli (1988, pp. 719–720, Footnote 1) is the first published work pointing out the difference.

for which there exists $\sigma_{-i} : \Theta_0 \times X \to \Delta(S_{-i})$ such that

$$\operatorname{supp} \sigma_{-i}(\theta_0, x) \subseteq ACR_{-i}^{k-1} \qquad \forall (\theta_0, x) \in \Theta_0 \times X, \tag{18}$$

$$s_i \in \arg\max_{s_i' \in S_i} \sum_{(\theta_0, x, s_{-i}) \in \Theta_0 \times X \times S_{-i}} \Pi_i[\theta_0, x]\, \sigma_{-i}(\theta_0, x)[s_{-i}]\, \bar{g}_i(\theta_0, x, s_i', s_{-i}). \tag{19}$$

It can be shown that, just like with ex ante rationalizability, the ex ante correlated rationalizable actions do not depend on the priors that we choose, as long as they are consistent with $(X_i, \pi_i)_{i \in I}$.

Proposition 2. *Fix a type space with information types and priors consistent with it. Let Δ be the restrictions derived from the type space. Then*

$$ACR_i = AR_i^{\Delta} \qquad \forall i \in I.$$

Proof. Let $(X_i, \pi_i)_{i \in I}$ be a type space with information types. If $\Delta = ((\Delta_{x_i})_{x_i \in X_i})_{i \in I}$ is the set of restrictions derived from it, then for every $i \in I$ and $\mu_i \in \Delta(\Theta_0 \times X \times S_{-i})$ the conditions of consistency of μ_i with Δ_i and of $\operatorname{marg}_{\Theta_0 \times X} \mu_i$ with $(X_i, \pi_i)_{i \in I}$ are identical. Since $ACR_i^0 = AR_i^{\Delta,0} = S_i$, for every $k \geq 1$ we thus obtain, inductively: there exists μ_i consistent with Δ_i and satisfying (16) and (17) if and only if, letting $\Pi_i = \operatorname{marg}_{\Theta_0 \times X} \mu_i$, there exists σ_{-i} satisfying (18) and (19). ■

We can now prove the main result in this section.

Theorem 3. *Fix a type space with information types and priors consistent with it. Then*

$$ACR_i \approx ICR_i \qquad \forall i \in I.$$

Proof. Fix $i \in I$. By Proposition 2, $ACR_i = ACR_i^{\Delta}$. By Proposition 1, $ACR_i^{\Delta} \approx R_i^{\Delta}$. By Corollary 1, $R_i^{\Delta} = ICR_i$. Thus, $ACR_i \approx ICR_i$. ■

Thus, looking deeper into the discrepancy between ex ante and interim rationalizability, we see that it is due to the different conditional independence restrictions, not to different types being allowed or not to hold different conjectures. Indeed, once these restrictions are removed, the discrepancy disappears: ex ante correlated rationalizability treats different types just as different information sets of the same player, and yet it is fully equivalent to ICR.

5 Discussion

5.1 Extensions

***n* players.** The most natural extension of IIR to static games with more than two players assumes that each type of each player believes that, conditional on the opponents' types, the payoff state and all the opponents' actions are mutually independent, whereas the natural extension of ICR allows for general correlation. All our characterization results have straightforward generalizations to this more general framework, except for the one in Remark 3. Indeed, for this natural extension of IIR, our remark about the equivalence between IIR and ICR under distributed knowledge of the payoff state does not hold, for the same reasons why independent rationalizability is a refinement of correlated rationalizability in games with complete information.

Dynamic games. Δ-rationalizability in dynamic games with incomplete information has been studied by Battigalli (2003), Battigalli and Siniscalchi (2003, 2007) and Battigalli and Prestipino (2010). These papers discuss also how to model independence assumptions in dynamic games. They study two versions of the solution concept, one that features a forward induction principle in the spirit of Pearce (1984) and Battigalli (1997), and a weaker one that does not. Battigalli and Siniscalchi (2007) give characterizations of both versions, thus extending our characterizations in (7) and (8). Battigalli and Prestipino (2010) provide an alternative characterization of the forward-induction version of Δ-rationalizability.[30] Proposition 1 on ex ante and interim Δ-rationalizability can also be extended. Similarly, one can define versions of ICR and IIR for dynamic Bayesian games with and without forward induction. (Penta, 2009 deals with the analogue of ICR without forward induction, defining analogues for the other notions is straightforward.) For these solution concepts, we can provide appropriate extensions of Propositions 1, 2 and Theorem 3; we conjecture that an extension of Theorem 1 also holds.

5.2 Related literature

We already mentioned the relationship with the work of Battigalli (2003) and Battigalli and Siniscalchi (2003, 2007) on Δ-rationalizability. Here we just notice that none of these papers makes the difference between payoff relevant and payoff irrelevant information explicit; actually, their notation and language suggest that only payoff relevant information is considered, although this is not a formal assumption. Furthermore, the first two papers assume distributed knowledge of the payoff state, although their results do not depend on this assumption.

ICR has been introduced by Dekel, Fudenberg, and Morris (2007), who also provide some epistemic characterization results. They prove that the ICR actions of a type only depend on the induced Θ-hierarchy.[31] The most important differences between their approach and ours is that they neglect private information (like Ely and Pęski, 2006) and do not state their epistemic results as expressible characterizations, i.e. by means of events in the appropriate canonical universal type space. These differences are related. One advantage of modeling private information (including the payoff irrele-

[30] They also discuss the definition of "strong Δ-rationalizability" of Battigalli and Siniscalchi (2003, 2007), showing that it is conceptually correct and equivalent to the definition of Battigalli (2003) when (a profiles of sets of conditional probability systems) satisfies a regularity condition assumed by Battigalli-Siniscalchi, but not more generally.

[31] This allows restricting attention to ICR actions in the Θ-based universal type space, as Dekel, Fudenberg, and Morris (2006), Weinstein and Yildiz (2007), Chen, Di Tillio, Faingold, and Xiong (2010), and Penta (2009) do in their analysis of the continuity of rationalizable actions with respect to beliefs hierarchies. (Penta, 2009 considers an extensive form version of ICR.)

vant one) explicitly, is that this provides a sufficiently rich language with which we can express the property of information-based conditional independence and the related characterization of IIR. We find the analogous characterization of Dekel, Fudenberg, and Morris (2007) less instructive because it relies on an *interpretation* of the type space as an "objective" information system that cannot be expressed in a formal language. Moreover, in our richer framework we can relate IIR and ICR to Δ-rationalizability, and we can formally state the obvious but important point that ICR and IIR are equivalent with two players and distributed knowledge of the payoff state.

Ely and Pęski (2006) analyze IIR. Like Dekel, Fudenberg, and Morris (2007), their starting point is the observation that IIR is not invariant to the addition/deletion of redundant types, and therefore depends on something more than the induced Θ-hierarchies (or even X-hierarchies). Their approach to IIR is essentially orthogonal to ours. We look for conditions under which IIR actions admit an expressible characterization, whereas they change the notion of belief hierarchy in order to obtain one that identifies IIR actions. They show that, under some regularity conditions, Harsanyi types yield — beside the standard Θ-hierarchies — also richer Δ-*hierarchies* where i's first-order beliefs are elements of $\Delta(\Delta(\Theta_0 \times \Theta_{-i}))$.[32] Then they show that Δ-hierarchies identify IIR actions. It is not clear to us whether Δ-hierarchies are expressible in a meaningful sense. To elaborate further, take any type space $(T_i, \vartheta_i, \upsilon_i, \pi_i)_{i\in I}$. As Ely and Pęski (2006, p. 28) point out, letting $\pi_i(t_i|\cdot) : T_{-i} \to \Delta(\Theta_0 \times \{\vartheta_{-i}(\cdot)\})$ for each $i \in I$ and $t_i \in T_i$ be a version of the conditional probability given $-i$'s type, we obtain Δ-hierarchies: in particular, the first-order belief in the Δ-hierarchy corresponding to type t_i of player i is defined as follows: for every measurable $E \subseteq \Delta(\Theta_0 \times \Theta_{-i})$,

$$\pi_i^{\Delta,1}(t_i)[E] = \pi_i(t_i)\left[\Theta_0 \times \{t_{-i} \in T_{-i} : \pi_i(t_i|t_{-i}) \in E\}\right].$$

If the type space has information types, so that $T_{-i} = X_{-i}$, then one can express this first-order belief as uncertainty about the relevant probability measure in the array $(\pi_i(t_i|x_{-i}))_{x_{-i}\in X_{-i}}$, thus making Δ-hierarchies expressible in some sense. But if the type space does not have information types, then we are not allowed to identify T_{-i} and X_{-i}, and this interpretation cannot be offered.

Sadzik (2009) seems to take a similar route to Ely and Pęski (2006): he defines hierarchical beliefs that identify Bayesian equilibrium actions. But on closer inspection, we find his approach much more similar to ours. He enriches the environment by adding to the payoff state θ a countable sequence of payoff-irrelevant (and continuous) *signals* for each player. On this expanded space of exogenous primitive uncertainty, call it Z, he constructs a formal language and relates it to standard Z-based hierarchies, showing that they identify Bayesian equilibrium actions. We speculatively propose the following interpretation of the difference between our approach to modeling uncertainty and his: we assume that there is common awareness only of a finite number of signals and consequently put only those signals in the commonly known environment.[33] This justifies conditionally

[32] Ely and Pęski (2006) have no private information — in our framework, this would correspond to the case where X_i is a singleton for each player i. We translate their definitions into our framework in the obvious way.

[33] Of course, a player may observe payoff-irrelevant aspects of which the opponent is unaware. In this case our rationalizability analysis should (and does) neglect these aspects.

correlated beliefs: when i conditions on the information type x_{-i} of $-i$, he suspects that $-i$ may observe some other payoff irrelevant variable i is not aware of, which in turn may be correlated with θ_0, thus allowing correlation between θ_0 and a_{-i} conditional on $-i$'s information type—this is a restatement of the incomplete model interpretation of conditional correlation given by Dekel, Fudenberg, and Morris (2007). On the other hand, Sadzik (2009) puts in the environment all the "conceivable" signals, which is justified if there is common awareness of all of them.

Liu (2009) analyzes Bayesian equilibrium predictions and the role of redundant types using an approach similar to ours. In particular, he distinguishes between redundant and non-redundant Θ-based type spaces, arguing that redundant types should be used only to represent hidden uncertainty entertained by players that the modeler does not explicitly take into account. Coherently with this approach, he suggests the modeler should always use a non-redundant type space unless he is aware there may be some additional strategically relevant information he is unaware of.[34] In our framework, the additional uncertainty is represented by the set of payoff irrelevant states Y and the exogenous beliefs of players are modeled using $(\Theta \times Y)$-based type spaces. In addition, Liu (2009) also shows that the same Bayesian equilibrium predictions can be obtained both with a Θ-based redundant type space and with an appropriate $(\Theta \times Y)$-based non-redundant type space. Instead of addressing Bayesian Equilibrium predictions, we use this richer uncertainty space, to highlight the connections among different definitions of rationalizability and to investigate the role of expressible independence restrictions.

A Appendices

To ease notation in the proofs below, given a joint assumption $E = E_1 \times E_2$, for each $k \geq 0$ we define $MB^k(E) = \cap_{0\leq n\leq k} B^n(E)$ and, for each player i, we denote the projection of $MB^k(E)$ on $X_i \times A_i \times H_i$ by $MB_i^k(E)$. Note that $MB_i^0(E) = E_i$ and $MB_i^k(E) = E_i \cap B_i(MB^{k-1}(E))$, while $CB_i(E) = \cap_{k\geq 0} MB_i^k(E)$. For each player i we let $U_i = X_i \times A_i \times H_i$ and $U_{X,i} = X_i \times H_{X,i}$. Moreover, we define $MB_i^{-1}(RAT) = U_i$, so that letting $MB^{-1}(RAT) = \Theta_0 \times U_1 \times U_2$ we have $MB_i^k(RAT) = RAT_i \cap B_i(MB_{-i}^{k-1}(RAT))$ for all $k \geq 0$. Finally, we make use of the mapping $\overline{\varrho}_{\Theta,i} : H_{X,i} \to H_{\Theta,i}$ defined as $\overline{\varrho}_{\Theta,i} = \varrho_{\Theta,i} \circ \varrho_{X,i}^{-1}$.

A.1 Proof of Theorem 1

Fix an X-space $(T_i, \vartheta_i, \upsilon_i, \pi_i)_{i\in I}$. Define

$$[u_{X,i}] = \{(x_i, a_i, h_i) \in U_i \ : \ (x_i, \varrho_{X,i}(h_i)) = u_{X,i}\} \qquad \forall i \in I, \ \forall u_{X,i} \in U_{X,i}.$$

[34]He also provides a necessary and sufficient condition on the space Θ (called "separativity") to identify a Θ-based redundant type space with a $(\Theta \times Y)$-based non-redundant type space through a mapping that preserves Θ-hierarchies. Given the finiteness assumption, this condition is satisfied in our framework.

Part I

In this part we prove the following:

$$ICR_i^k(t_i) \supseteq \mathrm{proj}_{A_i}\, MB_i^{k-1}(RAT) \cap [u_{X,i}] \quad \forall i \in I,\ \forall t_i \in T_i,\ \forall u_{X,i} \in \{\vartheta_i(t_i)\} \times Y_i \times (\overline{\varrho}_{\Theta,i})^{-1}(\eta_{\Theta,i}(t_i)).$$

This is enough to establish $ICR_i(t_i) \supseteq \mathrm{proj}_{A_i}\, CB_i(RAT) \cap E_i$ for every $E_i \in \mathcal{E}_i$ compatible with $\eta_{\Theta,i}(t_i)$, because every such E_i is a union of events of the form $[u_{X,i}]$ as above (and the union of their projections on A_i is the projection of their union).

The proof is by induction in k. The claim is trivially true for $k = 0$. Now let $n \geq 1$, assume that the claim is true for $k = n - 1$, and fix any $i \in I$, $t_i \in T_i$, $a_i \in A_i$, $y_i \in Y_i$ and $h_i \in H_i$ such that $\varrho_{\Theta,i}(h_i) = \eta_{\Theta,i}(t_i)$ and $(\vartheta_i(t_i), y_i, a_i, h_i) \in MB_i^{n-1}(RAT)$. Let $\varsigma_{-i} : \Theta_0 \times U_{-i} \to \Delta(\Theta_0 \times U_{-i})$ be any conditional distribution (see e.g. Dudley, 2002, pp. 269-270) given $\varphi_i(h_i)$ and the σ-algebra generated by the mapping

$$(\theta_0, \theta_{-i}, y_{-i}, a_{-i}, h_{-i}) \mapsto (\theta_0, \theta_{-i}, \varrho_{\Theta,-i}(h_{-i})).$$

Since ς_{-i} is measurable with respect to this σ-algebra, we can view it as a function with $\Theta_0 \times U_{\Theta,-i}$ as its domain. Thus, we can define $\sigma_{-i} : \Theta_0 \times T_{-i} \to \Delta(A_{-i})$ by letting

$$\sigma_{-i}(\theta_0, t_{-i}) = \mathrm{marg}_{A_{-i}}\, \varsigma_{-i}(\theta_0, \vartheta_{-i}(t_{-i}), \eta_{\Theta,-i}(t_{-i})) \qquad \forall (\theta_0, t_{-i}) \in \Theta_0 \times T_{-i}. \tag{20}$$

Note that $(\vartheta_i(t_i), y_i, a_i, h_i) \in MB_i^{n-1}(RAT) \subseteq B_i(\Theta_0 \times MB_{-i}^{n-2}(RAT))$ implies

$$\mathrm{supp}\, \varsigma_{-i}(\theta_0, u_{-i}) \subseteq \{\theta_0\} \times MB_{-i}^{n-2}(RAT) \qquad \text{for } \varphi_i(h_i)\text{-almost every } (\theta_0, u_{-i}) \in \Theta_0 \times U_{-i}$$

and hence by (20), using the induction hypothesis and $\varrho_{\Theta,i}(h_i) = \eta_{\Theta,i}(t_i)$,

$$\mathrm{supp}\, \sigma_{-i}(\theta_0, t_{-i}) \subseteq ICR_{-i}^{n-1}(t_{-i}) \qquad \text{for } \pi_i(t_i)\text{-almost every } (\theta_0, t_{-i}) \in \Theta_0 \times T_{-i}.$$

It is clear that $(\vartheta_i(t_i), y_i, a_i, h_i) \in MB_i^{n-1}(RAT) \subseteq RAT_i$ and $\varrho_{\Theta,i}(h_i) = \eta_{\Theta,i}(t_i)$ imply, using (20),

$$a_i \in \arg\max_{a_i'} \int_{\Theta_0 \times T_{-i}} g_i(\theta_0, \vartheta_i(t_i), \vartheta_{-i}(t_{-i}), a_i', \sigma_{-i}(\theta_0, t_{-i}))\pi_i(t_i)[d\theta_0 \times dt_{-i}].$$

Thus, $a_i \in ICR_i^n(t_i)$.

Part II

Here we prove the following:

$$ICR_i^k(t_i) \subseteq \mathrm{proj}_{A_i}\, MB_i^{k-1}(RAT) \cap [u_{X,i}] \quad \forall i \in I,\ \forall t_i \in T_i,\ \forall u_{X,i} \in \{\vartheta_i(t_i)\} \times Y_i \times (\overline{\varrho}_{\Theta,i})^{-1}(\eta_{\Theta,i}(t_i)).$$

This is enough to establish $ICR_i(t_i) \subseteq \mathrm{proj}_{A_i}\, CB_i(RAT) \cap E_i$ for every $E_i \in \mathcal{E}_i$ compatible with $\eta_{\Theta,i}(t_i)$, because $MB_i^{k-1}(RAT) \cap [u_{X,i}]$ is a decreasing sequence of nonempty compact sets converging to $CB_i(RAT) \cap [u_{X,i}]$, which is therefore a nonempty subset of $CB_i(RAT) \cap E_i$ whenever $[u_{X,i}] \subseteq E_i$.

The proof is by induction in k. As $\varrho_{X,i}$ is onto for each $i \in I$, the claim is clearly true for $k = 0$. Now let $n \geq 1$, assume that the claim is true for $k = n - 1$, and fix $i \in I$, $t_i \in T_i$, $y_i \in Y_i$, $h_{X,i} \in H_{X,i}$ with $\overline{\varrho}_{\Theta,i}(h_{X,i}) = \eta_{\Theta,i}(t_i)$, and $a_i \in ICR_i^n(t_i)$. Then there is $\sigma_{-i} : \Theta_0 \times T_{-i} \to \Delta(A_{-i})$ with

$$\operatorname{supp} \sigma_{-i}(\theta_0, t_{-i}) \subseteq ICR_{-i}^{n-1}(t_{-i}) \qquad \text{for } \pi_i(t_i)\text{-almost every } (\theta_0, t_{-i}) \in \Theta \times T_{-i}, \tag{21}$$

$$a_i \in \underset{a_i' \in A_i}{\arg\max} \int_{\Theta_0 \times T_{-i}} g_i(\theta_0, \vartheta_i(t_i), \vartheta_{-i}(t_{-i}), a_i', \sigma_{-i}(\theta_0, t_{-i}))\, \pi_i(t_i)[d\theta_0 \times dt_{-i}]. \tag{22}$$

Note that by the induction hypothesis $ICR_{-i}^{n-1}(t_{-i}) = ICR_{-i}^{n-1}(t'_{-i})$ for all $t_{-i}, t'_{-i} \in T_{-i}$ such that $\chi_{-i}(t_{-i}) = \chi_{-i}(t'_{-i})$ and $\eta_{X,-i}(t_{-i}) = \eta_{X,-i}(t'_{-i})$. Thus, without loss of generality we may assume $\sigma_{-i}(\theta_0, t_{-i}) = \sigma_{-i}(\theta_0, t'_{-i})$ for all $\theta_0 \in \Theta_0$ in each such case. Thus, by (21) and again by the induction hypothesis, there is $\tilde{\sigma}_{-i} : \Theta_0 \times U_{X,-i} \to \Delta(U_{-i})$ with

$$\operatorname{marg}_{A_{-i}} \tilde{\sigma}_{-i}(\theta_0, \chi_{-i}(t_{-i}), \eta_{X,-i}(t_{-i})) = \sigma_{-i}(\theta_0, t_{-i}) \qquad \forall (\theta_0, t_{-i}) \in \Theta_0 \times T_{-i}, \tag{23}$$

$$\operatorname{supp} \tilde{\sigma}_{-i}(\theta_0, u_{X,-i}) \subseteq MB_{-i}^{n-2}(RAT) \cap [u_{X,-i}] \qquad \forall u_{X,-i} \in U_{X,-i}. \tag{24}$$

Let ν_i be the probability distribution on $\Theta_0 \times U_{-i}$ induced by $\varphi_{X,i}(\eta_{X,i}(t_i))$ and $\tilde{\sigma}_{-i}$, that is,

$$\nu_i[\{\theta_0\} \times E_{-i}] = \int_{U_{X,-i}} \tilde{\sigma}_{-i}(\theta_0, u_{X,-i})[E_{-i}]\, \varphi_{X,i}(\eta_{X,i}(t_i))[\theta_0 \times du_{X,-i}]$$

for every $\theta_0 \in \Theta_0$ and measurable $E_{-i} \subseteq U_{-i}$. Since φ_i is onto, there exists $h_i \in H_i$ such that $\varphi_i(h_i) = \nu_i$. By construction, $\varphi_{X,i}(\varrho_{X,i}(h_i)) = \varphi_{X,i}(h_{X,i})$ and hence $\varrho_{X,i}(h_i) = h_{X,i}$ since $\varphi_{X,i}$ is injective. In particular, $\varrho_{\Theta,i}(h_i) = \eta_{\Theta,i}(t_i)$, which implies $(\vartheta_i(t_i), y_i, a_i, h_i) \in RAT_i$ by (22) and (23). Moreover, by (24), $\varphi_i(h_i)[\Theta_0 \times MB_{-i}^{n-2}(RAT)] = 1$. Thus, $(\vartheta_i(t_i), y_i, a_i, h_i) \in MB_i^{n-1}(RAT)$.

A.2 Measurability of the conditional independence assumption

Since φ_i is a homeomorphism, in order to prove that the set $H_{i,CI}$ is measurable it suffices to show that the set of all probability distributions on $\Theta_0 \times X_{-i} \times A_{-i} \times H_{-i}$ satisfying (15) is measurable. This, in turn, follows from the lemma below, letting $Z = \Theta_0 \times X_{-i} \times A_{-i} \times H_{-i}$, $\mathcal{F}_3 = \mathcal{E}_{-i}$, $\mathcal{F}_1 = \{\Theta_0' \times X_{-i} \times A_{-i} \times H_{-i} : \Theta_0' \subseteq \Theta_0\}$ and $\mathcal{F}_2 = \{\Theta_0 \times X_{-i} \times A'_{-i} \times H_{-i} : A'_{-i} \subseteq A_{-i}\}$. (Note that $\mathcal{E}_{-i}$ is indeed countably generated, since $X_{-i} \times H_{X,-i}$ is compact metric.) From the proof of the lemma it also follows that the set $H_{i,CI}$ does not depend on the particular version of conditional probability that we choose when defining it, as we stated above.

Lemma 1. *Fix a Polish space Z with its Borel σ-algebra $\mathcal{F}$. Let $\mathcal{F}_1 \subseteq \mathcal{F}$ and $\mathcal{F}_2 \subseteq \mathcal{F}$ be two finite sub-algebras, and let $\mathcal{F}_3 \subseteq \mathcal{F}$ be a countably generated sub-σ-algebra. Let $\mathcal{F}_{23}$ denote the σ-algebra generated by $\mathcal{F}_2 \cup \mathcal{F}_3$, and for each $\mu \in \Delta(Z)$ and $E \in \mathcal{F}$ fix arbitrarily regular versions $\mu[E|\mathcal{F}_3](\cdot)$ and $\mu[E|\mathcal{F}_{23}](\cdot)$ of the conditional probability of E given $\mathcal{F}_3$ and $\mathcal{F}_{23}$, respectively. Then for every $\mu \in \Delta(Z)$ the following conditions are equivalent:*

$$\mu\Big[\{z \in Z : \mu[E_1 \cap E_2|\mathcal{F}_3](z) = \mu[E_1|\mathcal{F}_3](z)\mu[E_2|\mathcal{F}_3](z) \;\; \forall E_1 \in \mathcal{F}_1, \forall E_2 \in \mathcal{F}_2\}\Big] = 1\,; \tag{25}$$

$$\mu\Big[\{z \in Z : \mu[E_1|\mathcal{F}_3](z) = \mu[E_1|\mathcal{F}_{23}](z) \;\; \forall E_1 \in \mathcal{F}_1\}\Big] = 1\,. \tag{26}$$

Furthermore, the set of all $\mu \in \Delta(Z)$ satisfying (25) (or equivalently (26)) is measurable.

Proof. The equivalence between (25) and (26) is well known—see e.g. Billingsley (1995, p. 456). Now let us verify that the set of all $\mu \in \Delta(Z)$ satisfying (25) is measurable. For every $p \in [0,1]$, $E \in \mathcal{F}$ and $E_3 \in \mathcal{F}_3$, observe that the two conditions

$$\mu[\{z \in E_3 : \mu[E|\mathcal{F}_3](z) \geq p\}] = \mu(E_3) \tag{27}$$

and

$$\mu[E \cap E_3'] \geq p\mu[E_3'] \qquad \forall E_3 \supseteq E_3' \in \mathcal{F}_3 \tag{28}$$

are equivalent,[35] and denote by $M(p, E, E_3)$ the set of all $\mu \in \Delta(Z)$ satisfying (27). By the said equivalence, each such set is measurable. Moreover, letting $\mathcal{A}$ denote the algebra generated by any countable family of events generating $\mathcal{F}_3$, the set of all $\mu \in \Delta(Z)$ satisfying (25) can be written as

$$\bigcap_{p,q \in \mathbb{Q}_{[0,1]}} \bigcap_{E_1 \in \mathcal{F}_1} \bigcap_{E_2 \in \mathcal{F}_2} \bigcap_{E_3 \in \mathcal{A}} \Big[\Big(\Delta(Z) \setminus \big(M(p, E_1, E_3) \cap M(q, E_2, E_3)\big)\Big) \cup M(pq, E_1 \cap E_2, E_3)\Big], \tag{29}$$

where $\mathbb{Q}_{[0,1]}$ denotes the set of rational numbers between 0 and 1. The set in (29) is measurable, so the proof is complete. ∎

A.3 Proof of Theorem 2

Part I

Fix an X-space $(T_i, \vartheta_i, \upsilon_i, \pi_i)_{i \in I}$. Here we prove that

$$IIR_i^k(t_i) \;\supseteq\; \mathsf{proj}_{A_i}\, MB_i^{k-1}(RAT \cap CI) \cap [\vartheta_i(t_i)] \cap [\eta_{X,i}(t_i)] \qquad \forall i \in I,\ \forall t_i \in T_i,\ \forall k \geq 0.$$

The claim is trivially true for $k = 0$. Now let $n \geq 1$, assume the claim is true for $k = n - 1$, and fix any $i \in I$, $t_i \in T_i$, $a_i \in A_i$, $y_i \in Y_i$ and $h_i \in H_i$ such that $\varrho_{X,i}(h_i) = \eta_{X,i}(t_i)$ and $(\vartheta_i(t_i), y_i, a_i, h_i) \in MB_i^{n-1}(RAT \cap CI)$. Let $\varsigma_{-i} : \Theta_0 \times U_{-i} \to \Delta(\Theta_0 \times U_{-i})$ be any conditional distribution given $\varphi_i(h_i)$ and the σ-algebra generated by the sets $\{\theta_0\} \times E_{-i}$, where $\theta_0 \in \Theta_0$ and $E_{-i} \in \mathcal{E}_{-i}$. As $(\vartheta_i(t_i), y_i, a_i, h_i) \in MB_i^{n-1}(RAT \cap CI) \subseteq CI_i$, it follows that $\varphi_i(h_i) \in H_{i,CI}$ and hence that $\mathsf{marg}_{A_{-i}}\, \varsigma_{-i}(\theta_0, u_{-i}) = \mathsf{marg}_{A_{-i}}\, \varsigma_{-i}(\theta_0', u_{-i})$ for $\varphi_i(h_i)$-almost every $(\theta_0, t_{-i}) \in \Theta_0 \times U_{-i}$ and every $\theta_0' \in \Theta_0$. (This follows at once from the equivalence between (25) and (26) in Lemma 1.) Moreover, since ς_{-i} is measurable with respect to the said σ-algebra, we can view it as a function with $\Theta_0 \times U_{X,-i}$ as its domain. Thus, using the fact that $\varrho_{X,i}(h_i) = \eta_{X,i}(t_i)$, there exists a well defined, measurable $\sigma_{-i} : T_{-i} \to \Delta(A_{-i})$ satisfying the following:

$$\sigma_{-i}(t_{-i}) \;=\; \mathsf{marg}_{A_{-i}}\, \varsigma_{-i}(\theta_0, \chi_{-i}(t_{-i}), \eta_{X,-i}(t_{-i})) \qquad \text{for } \pi_i(t_i)\text{-almost every } (\theta_0, t_{-i}) \in \Theta_0 \times T_{-i}. \tag{30}$$

[35] By definition of conditional probability, $\int_{E_3} \mu[E|\mathcal{F}_3](z)\mu(dz) = \mu[E \cap E_3]$ for every $E_3 \in \mathcal{F}_3$. Thus, (27) implies (28) and, conversely, (28) implies that $\{z \in Z : \mu[E|\mathcal{F}_3](z) < p\}$ is a $\mathcal{F}_3$-measurable event of μ-probability zero, i.e. (27).

Note that $(\vartheta_i(t_i), y_i, a_i, h_i) \in MB_i^{n-1}(RAT \cap CI) \subseteq B_i(\Theta_0 \times MB_{-i}^{n-2}(RAT \cap CI))$ implies

$$\operatorname{supp} \varsigma_{-i}(\theta_0, u_{-i}) \subseteq \{\theta_0\} \times MB_{-i}^{n-2}(RAT \cap CI) \quad \text{for } \varphi_i(h_i)\text{-almost every } (\theta_0, u_{-i}) \in \Theta_0 \times U_{-i} \tag{31}$$

and hence, by the induction hypothesis,

$$\operatorname{supp} \sigma_{-i}(t_{-i}) \subseteq IIR_{-i}^{n-1}(t_{-i}) \qquad \text{for } \pi_i(t_i)\text{-almost every } t_{-i} \in T_{-i}.$$

Clearly, $(\vartheta_i(t_i), y_i, a_i, h_i) \in MB_i^{n-1}(RAT \cap CI) \subseteq RAT_i$ and $\varrho_{X,i}(h_i) = \eta_{X,i}(t_i)$ imply, using (30),

$$a_i \in \arg\max_{a_i'} \int_{\Theta_0 \times T_{-i}} g_i(\theta_0, \vartheta_i(t_i), \vartheta_{-i}(t_{-i}), a_i', \sigma_{-i}(t_{-i}))\pi_i(t_i)[d\theta_0 \times dt_{-i}].$$

This proves that $a_i \in IIR_i^n(t_i)$.

Part II

Fix a non-redundant X-space $(T_i, \vartheta_i, \upsilon_i, \pi_i)_{i\in I}$. We prove that

$$IIR_i^k(t_i) \subseteq \operatorname{proj}_{A_i} MB_i^{k-1}(RAT \cap CI) \cap [(\chi_i(t_i), \eta_{X,i}(t_i))] \qquad \forall i \in I,\ \forall t_i \in T_i,\ \forall k \geq 0.$$[36]

The claim is true for $k = 0$ because $\varrho_{X,i}$ is onto. Now let $n \geq 1$, assume that the claim is true for $k = n - 1$, and fix $i \in I$, $t_i \in T_i$ and $a_i \in IIR_i^n(t_i)$. Fix a measurable $\sigma_{-i} : T_{-i} \to \Delta(A_{-i})$ with

$$\operatorname{supp} \sigma_{-i}(t_{-i}) \subseteq IIR_{-i}^{n-1}(t_{-i}) \qquad \text{for } \pi_i(t_i)\text{-almost every } t_{-i} \in T_{-i}, \tag{32}$$

$$a_i \in \arg\max_{a_i' \in A_i} \int_{\Theta_0 \times T_{-i}} g_i(\theta_0, \vartheta_i(t_i), \vartheta_{-i}(t_{-i}), a_i', \sigma_{-i}(t_{-i}))\, \pi_i(t_i)[d\theta_0 \times dt_{-i}]. \tag{33}$$

By the induction hypothesis, non-redundancy and (32), there exists $\tilde{\sigma}_{-i} : U_{X,-i} \to \Delta(X_{-i} \times A_{-i} \times H_{-i})$ satisfying the following:

$$\operatorname{marg}_{A_{-i}} \tilde{\sigma}_{-i}((\chi_{-i}(t_{-i}), \eta_{X,-i}(t_{-i}))) = \sigma_{-i}(t_{-i}) \qquad \forall t_{-i} \in T_{-i}, \tag{34}$$

$$\operatorname{supp} \tilde{\sigma}_{-i}(u_{X,-i}) \subseteq MB_{-i}^{n-2}(RAT \cap CI) \cap [u_{X,-i}] \qquad \forall u_{X,-i} \in U_{X,-i}. \tag{35}$$

Let ν_i be the probability distribution on $\Theta_0 \times U_{-i}$ induced by $\varphi_{X,i}(\eta_{X,i}(t_i))$ and $\tilde{\sigma}_{-i}$, that is,

$$\nu_i[\{\theta_0\} \times E_{-i}] = \int_{H_{X,-i}} \tilde{\sigma}_{-i}(u_{X,-i})[E_{-i}]\, \varphi_{X,i}(h_{X,i})[\theta_0 \times du_{X,-i}]$$

for every $\theta_0 \in \Theta_0$ and measurable $E_{-i} \subseteq U_{-i}$. Since φ_i is onto, there exists $h_i \in H_i$ with $\varphi_i(h_i) = \nu_i$. By construction, $\nu_i \in H_{i,CI}$ and $\varphi_i(\varrho_{X,i}(h_i)) = \varphi_{X,i}(\eta_{X,i}(t_i))$, hence $(\vartheta_i(t_i), \chi_i(t_i), a_i, h_i) \in CI_i$ and, since $\varphi_{X,i}$ is injective, $\varrho_{X,i}(h_i) = \eta_{X,i}(t_i)$. Thus $(\vartheta_i(t_i), \chi_i(t_i), a_i, h_i) \in RAT_i$ by (33) and (34), while (35) implies $\varphi_i(h_i)[\Theta_0 \times MB_{-i}^{n-2}(RAT \cap CI)] = 1$. Therefore, $(\vartheta_i(t_i), y_i, a_i, h_i) \in MB_i^{n-1}(RAT \cap CI)$.

[36]This is enough to establish $IIR_i(t_i) \subseteq \operatorname{proj}_{A_i} CB_i(RAT \cap CI) \cap [(\chi_i(t_i), \eta_{X,i}(t_i))]$ because $MB_i^{k-1}(RAT \cap CI) \cap [u_{X,i}]$ is a decreasing sequence of nonempty compact sets converging to $CB_i(RAT \cap CI) \cap [u_{X,i}]$, which is therefore nonempty.

A.4 Proof of Proposition 1

We prove by induction that $ACR_i^{\Delta,k} \approx R_i^{\Delta,k}$ for every $i \in I$ and $k \geq 0$. This trivially holds for $k = 0$. Let $n \geq 1$ and assume that it is true for $k = n - 1$. In order to prove it for $k = n$, fix $i \in I$ and $s_i \in S_i$. If $s_i \in ACR_i^{\Delta,n}$ then there exists $\mu_i \in \Delta(\Theta_0 \times X \times S_{-i})$ consistent with Δ_i such that

$$\operatorname{supp} \mu_i \subseteq \Theta_0 \times X \times AR_{-i}^{\Delta,n-1}, \tag{36}$$

$$s_i \in \arg\max_{s_i' \in S_i} \sum_{(\theta_0,x,s_{-i}) \in \Theta_0 \times X \times S_{-i}} \mu_i[\theta_0, x, s_{-i}]\, \bar{g}_i(\theta_0, x, s_i', s_{-i}). \tag{37}$$

Since μ_i is consistent with Δ_i, for all $x_i \in X_i$ we have $\mu_i[x_i] > 0$ and, letting ν_{x_i} be the conditional probability given x_i induced on $\Theta_0 \times X_{-i} \times A_{-i}$ by μ_i, also $\nu_{x_i} \in \Delta_{x_i}$, hence by the induction hypothesis (36) and (37), respectively, imply that for every $x_i = (\theta_i, y_i) \in X_i$,

$$\operatorname{supp} \nu_{x_i} \subseteq \Theta_0 \times \left\{(x_{-i}, a_{-i}) \in X_{-i} \times A_{-i} : a_{-i} \in R_{-i}^{\Delta,n-1}(x_{-i})\right\}, \tag{38}$$

$$s_i(x_i) \in \arg\max_{a_i \in A_i} \sum_{(\theta_0,\theta_{-i},y_{-i},a_{-i}) \in \Theta_0 \times X_{-i} \times A_{-i}} \nu_{x_i}[\theta_0, \theta_{-i}, y_{-i}, a_{-i} \,|\, x_i]\, g_i(\theta_0, \theta_i, \theta_{-i}, a_i, a_{-i}). \tag{39}$$

This proves that s_i is a selection from $R_i^{\Delta,n}$. Conversely, if the latter is true, then for each $x_i \in X_i$ there exists $\nu_{x_i} \in \Delta_{x_i}$ such that (38) and (39) hold. Let λ_i be an arbitrary full-support probability distribution on X_i, and let μ_i denote the probability distribution on $\Theta_0 \times X \times S_{-i}$ defined as

$$\mu_i[\theta_0, x_i, x_{-i}, s_{-i}] = \lambda_i[x_i]\, \nu_{x_i}[\theta_0, x_{-i}, s_{-i}(x_{-i})] \qquad \forall (\theta_0, x_i, x_{-i}, s_{-i}) \in \Theta_0 \times X_i \times X_{-i} \times S_{-i}.$$

By construction, using the induction hypothesis, (38) and (39) guarantee that μ_i satisfies (36) and (37), respectively. Thus, $s_i \in AR_i^{\Delta,n}$.

References

Battigalli, Pierpaolo (1988), "Implementable strategies, prior information and the problem of credibility in extensive games." *RISEC (International Review of Economics and Business)*, 35, 705–733.

Battigalli, Pierpaolo (1997), "On rationalizability in extensive games." *Journal of Economic Theory*, 74, 40–61.

Battigalli, Pierpaolo (2003), "Rationalizability in infinite, dynamic games with incomplete information." *Research in Economics*, 57, 1–38.

Battigalli, Pierpaolo and Andrea Prestipino (2010), "Transparent restrictions on beliefs and forward induction reasoning in games with asymmetric information." Mimeo, Bocconi University and New York University.

Battigalli, Pierpaolo and Marciano Siniscalchi (2002), "Strong belief and forward induction reasoning." *Journal of Economic Theory*, 106, 356–391.

Battigalli, Pierpaolo and Marciano Siniscalchi (2003), "Rationalization and incomplete information." *Advances in Theoretical Economics*, 3. Available at http://www.bepress.com/bejte/advances/vol3/iss1/art3.

Battigalli, Pierpaolo and Marciano Siniscalchi (2007), "Interactive epistemology in games with payoff uncertainty." *Research in Economics*, 61, 165–184.

Bergemann, Dirk and Stephen Morris (2005), "Robust implementation: The role of large type spaces." Mimeo, Yale University (previous version: 2003).

Bergemann, Dirk and Stephen Morris (2007), "Belief-free incomplete information games." Cowles Foundation Discussion Paper 1619.

Bergemann, Dirk and Stephen Morris (2009), "Robust implementation in direct mechanisms." *Review of Economic Studies*, 76, 1175–1204.

Billingsley, Patrick (1995), *Probability and Measure*. John Wiley & Sons.

Brandenburger, Adam and Eddie Dekel (1987), "Rationalizability and correlated equilibria." *Econometrica*, 55, 1391–1402.

Carlsson, Hans and Eric van Damme (1993), "Global games and equilibrium selection." *Econometrica*, 61, 989–1018.

Chen, Yi-Chun, Alfredo Di Tillio, Eduardo Faingold, and Siyang Xiong (2010), "Uniform topologies on types." *Theoretical Economics*, 5, 445–478.

Dekel, Eddie, Drew Fudenberg, and Stephen Morris (2006), "Topologies on types." *Theoretical Economics*, 1, 275–309.

Dekel, Eddie, Drew Fudenberg, and Stephen Morris (2007), "Interim correlated rationalizability." *Theoretical Economics*, 2, 15–40.

Dudley, Richard M. (2002), *Real Analysis and Probability*. Cambridge University Press, second edition.

Ely, Jeffrey C. and Marcin Pęski (2006), "Hierarchies of belief and interim rationalizability." *Theoretical Economics*, 1, 19–65.

Harsanyi, John (1967-68), "Games with incomplete information played by bayesian players, i-iii." *Management Science*, 14, 159–182, 320–334, 486–502.

Heifetz, Aviad and Dov Samet (1998), "Topology-free typology of beliefs." *Journal of Economic Theory*, 82, 324–341.

Klemperer, Paul (1998), "Auctions with almost common values: The 'wallet game' and its applications." *European Economic Review*, 42, 757–770.

Kuhn, Harold W. (1953), "Extensive games and the problem of information." In Harold W. Kuhn and A. W. Tucker, editors, "Contributions to the Theory of Games II," Princeton University Press, pp. 193–216.

Liu, Qingmin (2009), "On redundant types and bayesian formulation of incomplete information." *Journal of Economic Theory*, 144, 2115–2145.

Mertens, Jean-Francois and Shmuel Zamir (1985), "Formulation of bayesian analysis for games with incomplete information." *International Journal of Game Theory*, 14, 1–29.

Morris, Stephen and H. S. Shin (2007), "Common belief foundations of global games." Mimeo, Princeton University.

Pearce, David G. (1984), "Rationalizable strategic behavior and the problem of perfection." *Econometrica*, 52, 1029–1050.

Penta, Antonio (2009), "Higher order beliefs in dynamic environments." Mimeo, University of Pennsylvania.

Sadzik, Tomasz (2009), "Beliefs revealed in bayesian equilibrium." Mimeo, New York University.

Tan, T. and S. Werlang (1988), "The bayesian foundation of solution concepts of games." *Journal of Economic Theory*, 45, 370–391.

Weinstein, Jonathan and Muhamet Yildiz (2007), "A structure theorem for rationalizability with application to robust predictions of refinements." *Econometrica*, 75, 365–400.