Higher Order Beliefs in Dynamic Envinronments.*

Antonio Penta†

Dept. of Economics,
University of Pennsylvania

this version: February 7, 2009

Abstract

This paper explores the role of higher order beliefs in dynamic environments, and extends most of Weinstein and Yildiz's (*Econometrica*, 2007) results to these settings. To this end, a solution concept for Bayesian games in extensive form is introduced: *Interim Sequential Rationalizability* ($\mathcal{ISR}$). It is shown that, if the space of uncertainty is sufficiently rich, $\mathcal{ISR}$ is generically unique on the universal type space of Mertens and Zamir (*Int.Journ.of Game Th.*,1985). If for type t $\mathcal{ISR}$ is not unique, any of the multiple outcomes is uniquely-$\mathcal{ISR}$ along a sequence converging to t. Furthermore, $\mathcal{ISR}$ is upper hemicontinuous (u.h.c.). It is thus the strongest u.h.c. solution concept on the universal space. Furthermore, $\mathcal{ISR}$ is *type space invariant* (behavior is completely determined by hierarchies of beliefs, irrespective of the type space used to model them); in environments in which each player knows his own payoffs, $\mathcal{ISR}$ also satisfies a stronger robustness property, called *model invariance.*

Keywords: dynamic games – hierarchies of beliefs – higher order beliefs – robustness – uniqueness

1 Introduction

A fundamental tenet of main-stream game theory is the assumption that the rules of the game, including its payoff and information structure, are commonly known by players. Typically,

*preliminary and incomplete version: please, do not circulate without permission of the author; comments are very appreciated. (*email:* penta@sas.upenn.edu)

†I would like to thank my advisor, G. Mailath, for the dedication, the encouragement, and above all for the constant and valuable criticisms. I also thank L. Alaoui, Y. Chen, E. Dekel, A. Friedenberg, K. Kim, J. Li, Q. Liu, M.Pinter and A. Postlewaite. Special thanks go to P. Battigalli, to whom I'm particularly indebted. Finally, I thank seminar and conference audiences at: UPenn, Bocconi, the 2008 NASM of the ES (Pittsburgh) and the 2008 ESEM (Milan).

such common knowledge (CK) assumptions are only meant to be approximately satisfied in the "actual" situation, and should therefore be regarded as mere simplifying modelling assumptions. The significance of our predictions then relies crucially on their robustness to possible misspecifications of players' beliefs. Yet, it is a well-known fact that a model's predictions can be largely upset by arbitrary small perturbations of players' higher order beliefs (e.g. Rubinstein, 1989).

Recent work by Weinstein and Yildiz (2007, WY hereafter) explored the premises of the refinements literature when all CK assumptions on the payoffs of a static game are relaxed, proving a somewhat negative result: Whenever a model has multiple rationalizable outcomes, any of these is uniquely rationalizable for an arbitrarily close model of beliefs. On the other hand, in the space of hierarchies of beliefs, models are generically dominance-solvable.[1] WY build on the literature on *global games*: A sufficiently rich space of uncertainty (containing dominance regions for each player's strategy) is assumed; then, an "infection argument" is applied to obtain the generic uniqueness result.[2] This result generalizes an important insight from the global games' literature: Multiplicity can be seen as the direct consequence of the CK assumptions that are implicit in standard models. Relaxing such assumptions, CK of rationality is (generically) sufficient to explain agents' coordination of expectations, delivering a unique rationalizable outcome. On the other hand, when there is multiplicity, WY's results imply that there is no way to refine it away: If no CK restrictions on payoffs are imposed, *rationalizability* is the strongest solution concept that delivers *robust* predictions.[3]

WY's analysis does not apply to dynamic games: This is an important limitation, as many refinements have been introduced precisely to account for extensive form considerations. The extension to dynamic games though is not just a technical matter: It raises several conceptual problems that require special care.[4]

In dynamic games, players' information about the environment is not entirely exogenous:

[1] WY work with the universal type space of Mertens and Zamir (1985), which is endowed with the product topology. The notion of genericity adopted in WY is a weak one: the uniqueness result is proved for an open and dense subset of the universal type space.

[2] The classical reference on global games is Carlsson and Van Damme (1993); Frankel et al. (2003) generalize many of the results; Morris and Shin (2003) survey the literature.

[3] More formally, if all CK-assumptions are relaxed, Interim Correlated Rationalizability is the strongest upper hemicontinuous solution concept on the universal type space.

[4] WY suggest to extend their work to dynamic settings maintaining a normal form approach and introducing *trembles*. They also point out possible problems that a tremble-based approach may determine. The present paper instead adopts an *extensive form* approach. This is not simply to avoid the use of trembles: The problems with trembles are just technical; rather, the most critical conceptual problems are raised by the normal form approach. For instance, independent work by Chen (2008) also dispensed with trembles, but maintained a normal form-approach. As it will be discussed below (section 6), his approach cannot account for environments in which *sequential rationality* maintains some restrictive power, a subset of the environments

As the game unfolds, players' may extract information about the environment from the history of play, updating their initial beliefs; furthermore, some information endogenously becomes CK (e.g. public histories), possibly serving as a coordination device and favoring multiplicity. This mechanism is at work in many studies on dynamic global games, which have shown how imposing the global games information structure on dynamic environments may not deliver the familiar uniqueness results.[5] On a more general level, these observations cast some shadow on the possibility of drawing tempting analogies from static to dynamic environments: For instance, it is not clear whether the pervasive multiplicity of equilibria in signalling games can be imputed to the implicit CK assumptions, as relaxing such assumptions in a dynamic setting may not suffice to obtain the familiar uniqueness result.

Another special feature of dynamic environments is the possibility that players observe unexpected events: In general, players hold *joint* beliefs about their opponents' behavior and the features of the environment they don't know; upon observing an unexpected event, players may update their beliefs in any way that is consistent with the new piece of evidence, but they would maintain whatever *knowledge* of the environment they had. The distinction between *knowledge* and *certainty* (probability-one belief) is thus fundamental, as it affects the set of beliefs players may entertain after unexpected events. This distinction is immaterial for the purpose of WY's analysis in static settings, but becomes crucial in dynamic environments, and must be accounted for in the specification of the model, raising novel questions of robustness.

The discussion above can be summarized by three questions, left open by the existing literature: What solution concept does deliver *robust predictions*, when no CK-restrictions on payoffs are imposed in dynamic environments? Can multiplicity, also in dynamic environments, be imputed to the commonly made CK assumptions? If so, what solution concept does deliver the familiar uniqueness result, when such assumptions are relaxed?

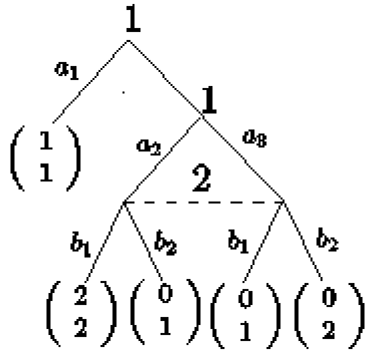
To answer these questions, a solution concept for Bayesian games in extensive form is introduced, *Interim Sequential Rationalizability* ($\mathcal{ISR}$ hereafter). $\mathcal{ISR}$ is characterized by the epistemic assumptions of *sequential rationality* and *initial common certainty of sequential rationality*, and extends to Bayesian games ideas from solution concepts introduced by Ben-Porath (1997) and Battigalli and Siniscalchi (2007).[6] The intuition behind $\mathcal{ISR}$ can be grasped by the following example.

Example 1 (Interim Sequential Rationalizability). Consider the complete informa-

considered here.

[5]See e.g., Angeletos et al. (2006, 2007) and Angeletos and Werning (2006). Section 6 discusses that literature.

[6]The epistemic characterization of $\mathcal{ISR}$ is provided in the appendix. Section 6 discusses the relations of $\mathcal{ISR}$ with existing solution concepts.



tion game in figure 1. Strategy a_3 is dominated by a_1. Thus, if *at the beginning of the game* player 2 thinks that 1 is rational, he assigns zero probability to a_3 being played. For example, 2 could assign probability one to a_1, so that the next information set is unexpected. We can consider two different hypothesis now:

> [**H.1**] 2 believes that 1 is rational even after an unexpected move; or
>
> [**H.2**] 2 believes that 1 is rational as long as he is not surprised, but he is willing to consider that 1 is not rational if he observes an unexpected move.

If [**H.1**] is true, in the subgame player 2 would still assign zero probability to a_3, and play b_1 if rational. If 1 believes [**H.1**] and that 2 is rational, he would expect b_1 to be played. Then, if 1 is also rational, he would play a_2. This is the logic of Pearce's (1984) *Extensive Form Rationalizability (EFR)*, which delivers (a_2, b_1) as the unique outcome in this game.

Now, let's maintain that 2 is rational, but assume [**H.2**] instead: once surprised, player 2 is willing to consider that 1 is not rational, and hence in the subgame he may assign probability one to a_3 having been played, which would justify b_2. If at the beginning 2 assigned positive probability to a_2, then the subgame would not be unexpected, and player 2 would still assign zero probability to a_3, making b_1 the unique best response. Thus, if [**H.2**] is true, either b_1 or b_2 may be played by a rational player 2. If 1 believes that 2 is rational and that [**H.2**] is true, he cannot rule out either b_1 or b_2, and so both a_1 and a_2 may be played by a rational player 1.□

As illustrated by the example, the epistemic assumptions characterizing $\mathcal{ISR}$ entail very weak beliefs of players in their opponents' rationality. Consequently, $\mathcal{ISR}$ may seem an unreasonably weak solution concept. Nonetheless, one of the main results of the paper shows that when no CK restrictions on payoffs are imposed, $\mathcal{ISR}$ is the strongest upper hemicontinuous solution concept for dynamic games: any refinement of $\mathcal{ISR}$ would deliver non-robust predictions.[7] This result is closely related to the work by Dekel and Fudenberg (1990), who considered a similar robustness question.[8] It is worth though to emphasize that it holds once *all* CK assumptions are being relaxed: although assumptions of CK are intrinsically strong ones, in some cases analysts may have reasons to believe that some CK assumptions are actually satisfied. In that case, the robustness exercise performed here may be unnecessarily demanding, and stronger predictions may be robust when only some (as opposed to all) CK assumptions are relaxed.[9]

A generic uniqueness result for $\mathcal{ISR}$ is also proved, with respect to which the weakness of $\mathcal{ISR}$ is rather a *strength*. This result shows that imputing multiplicity to CK assumptions is legitimate in dynamic environments as well. Furthermore, very weak epistemic assumptions (generically) suffice for players to achieve coordination of expectations in dynamic environments, with no need to invoke sophisticated backward or forward induction reasoning.

The results discussed so far investigated the relation between a model's predictions and players' hierarchies of beliefs, i.e. the behavior of a solution concept on the universal type space. In applied work, the modelling activity typically involves the selection of a subset of all the possible hierarchies of beliefs to focus on. To the extent that "small mistakes" can be made in such selection, the concern for *robust predictions* provides a foundation for the adoption of $\mathcal{ISR}$ as a solution concept. The modelling activity though is not limited to the selection of a subset of hierarchies: Following Harsanyi (1967-68), it is standard to represent a given set of hierarchies of beliefs by means of (non-universal) *type spaces*; the choice of which type space to use to represent a given set of hierarchies of beliefs may potentially affect the predictions of a solution concept. Invariance with respect to such modelling choices is thus another robustness property (*type space-invariance*): $\mathcal{ISR}$ is proven to satisfy this property as well.[10] Other important modelling choices concern players' information structures, which determine what players *know* and what they can at most be *certain* of. As argued above, this distinction is particularly important in dynamic settings, leading to a novel notion of robustness, called *model invariance*. $\mathcal{ISR}$ is shown to be model-invariant in some environments, not in others. A more thourough discussion of these robustness questions is postponed to section 2.

[7] Example 3 in section 2 shows that, despite its intuitive appeal, the predictions of *EFR* for the game in figure 1 are *not* robust. A related non-robustness result for *EFR* was proved by Battigalli and Siniscalchi (2007). I thank Pierpaolo Battigalli for pointing this out.

[8] A discussion of the connections with that work is postponed to section 6.

[9] The same observation applies to WY's as well. What robust predictions can be made when only *some* CK-assumptions are relaxed is an interesting question for future research. Penta (2009) starts investigating the problem in the context of static games.

[10] Type space-dependence in static settings has been studied by Ely and Peski (2006) and Dekel et al. (2007). The issue is discussed at some lenght in sections 2, 4.2 and 6.

The rest of the paper is organized as follows: section 2 contains several examples to illustrate the main concepts and results. Section 3 introduces the game theoretic framework and $\mathcal{ISR}$. Section 4 provides some robustness results: upper hemicontinuity, type space-invariance and model-invariance. Section 5 explores the structure of $\mathcal{ISR}$ on the universal type space, and proves that under a suitable *richness condition* it is generically unique and that any refinement of $\mathcal{ISR}$ is not robust. Proofs are in the appendices, which also contain the epistemic characterization of $\mathcal{ISR}$ and (for environments in which players' know their own payoffs) a characterization of $\mathcal{ISR}$ as Dekel and Fudenberg's (1990) $S^{\infty}W$-procedure applied to the interim normal form.

2 Relaxing CK-assumptions and Robustness in dynamic games

This section is meant as a non-technical introduction to the main concepts and ideas developed in the paper. The starting point is the observation that standard models assume that a certain payoff structure is CK among the players, thus imposing restrictions on the entire hierarchies of beliefs. The robustness question is then: What predictions retain their validity, when all CK assumptions are relaxed and agents' hierarchies of beliefs are only "close" to the one specified in the original model? In the context of static environments, WY proved that rationalizability is the strongest robust solution concept: Whenever a model has multiple rationalizable outcomes, *any* of these is uniquely rationalizable in a sufficiently "close" model of beliefs. The following example, borrowed from WY, illustrates the point:

Example 2 (Weistein and Yildiz, 2007). Let payoffs be as in the following matrix, where θ is a real number:

	Attack	Not
Attack	θ,θ	$\theta-1,0$
Not	$0,\theta-1$	$0,0$

Let $\theta^{*}=2/5$ and $T^{CK}=\left\{t^{CK}\right\}$ denote the model in which θ^{*} is CK. This delivers a coordination game, with $(Attack, Attack)$ and (Not, Not) as the two pure strategy Nash Equilibria, hence both actions are rationalizabile for each player. In this complete information model, only Not is a risk-dominant action for each player. If no CK assumptions are imposed though, Risk-dominance (and so any refinement of rationalizability) does not deliver predictions that are robust to possible misspecifications of the agents'hierarchies of beliefs. Models with beliefs arbitrarily close to t^{CK}'s can be constructed, for which either action is uniquely rationalizable.

Let Θ^{*} be a parameter space that contains θ^{*}, $-2/5$ and $6/5$, so that any strategy is dominant in some state, and consider type space T with set of types $T_1=\{-1,1,3,...\}$ and $T_2=\{0,2,4,...\}$, and beliefs as follow: type -1 puts probability one on state $\theta'\neq\theta^{*}$ and type 0; types $k=0,1,2,...$ all put probability $p\in(0,1)$ on $(\theta^{*},k-1)$ and probability $(1-p)$

on $(\theta^*, k+1)$. The sequence of types profiles $\{(k, k+1)\}_{k\in\mathbb{N}}$ converges to *common certainty* of θ^* as $k \to \infty$: Type $k=1$ is certain of θ^*, and also puts probability one on the opponent's type being 0 or 2, who are also certain of θ^*. Hence type 1 is consistent with mutual certainty of θ^*, but not with common certainty: Type 1 puts positive probability on type 0, which puts positive probability on type -1, which is certain of $\theta' \neq \theta^*$. Similarly, any type k is consistent with k iterations of mutual certainty of θ^*, so that the first k orders of beliefs of type k are the same as type t^{CK}. But as long as $p \in (0,1)$, any finite k is not consistent with common certainty of θ^*: *Common certainty* is only approached letting $k \to \infty$.[11]

Now, suppose that $\theta' = -2/5$ and $p = 2/3$: Type -1 plays Not, as he is certain that it is dominant; type 0 is certain of θ^*, but puts probability 2/3 on type -1, who plays Not, hence playing Not is his unique rationalizable strategy. The argument can be iterated, so that for each $k > 1$, despite there are k levels of mutual certainty of θ^*, playing Not is the unique rationalizable action. It is easy to check that the same reasoning can be repeated to obtain *Attack* as uniquely rationalizable, simply by letting $\theta' = 6/5$.□

WY's program involves the following steps: 1) Start with a standard model (such as t^{CK}), which makes implicit CK assumptions on the payoff structure and players' hierarchies of beliefs. 2) To relax *all* the CK assumptions, this model is embedded in a larger one, with an underlying space of uncertainty Θ^*: Being part of the specification of the model, any such space would be commonly known by players. For instance, letting $\Theta^* = \{-2/5, \theta^*\}$ would not allow to relax the assumption that *Attack* is *not* dominant. While letting $\{-2/5, \theta^*, 6/5\} \subseteq \Theta^*$ as in example 2, allows to relax the assumptions that either action is not dominant. Hence, relaxing *all* CK assumptions, essentially means to consider all possible hierarchies of beliefs players may have about a *sufficiently rich* space of uncertainty. Let Θ^* denote such "rich" fundamental space. In example 1, type t^{CK} was "embedded" in such larger model as a type (call it t^*) corresponding to *common certainty* of θ^*. 3) Once the CK assumptions of the original model are embedded as common certainty assumptions in the "universal model", the robustness of a solution concept can be formulated as a continuity property in this space. Hence, the properties of sequences converging to the common certainty type t^* could be used to address the robustness of the predictions of the complete information model t^{CK}.

The possibility of envisioning t^{CK} as t^* is central to the argument, but rests on the interchangeability of *knowledge* and *certainty*. Modelling *knowledge* as *certainty* is innocuous for the purpose of WY's analysis, but the distinction becomes crucial when the analysis is extended to dynamic settings. As a consequence, the properties of a solution concept on the "universal model" do not provide an exhaustive answer to the original robustness question: WY's program must be deconstructed, and new concepts introduced.

[11] This convergence is in the product topology of hierarchies of beliefs.

2.1 Dynamic environments.

Modelling incomplete information. A dynamic game is defined by an *extensive form* $\langle N, \mathcal{H}, \mathcal{Z}\rangle$ ($N = \{1, ..., n\}$ is the set of players; $\mathcal{H}$ and $\mathcal{Z}$ the sets of partial and terminal histories, respectively) and *players' payoffs*, defined over the terminal histories. As for the static case, incomplete information is modelled parametrizing the payoff functions on a *rich* space of uncertainty Θ^*, letting $u_i : \mathcal{Z} \times \Theta^* \to \mathbb{R}$.

In general, let Θ^* be written as

$$\Theta^* = \Theta_0^* \times \Theta_1^* \times ... \times \Theta_n^*.$$

For each $i = 1, ..., n$, Θ_i^* is the set of player i's *payoff types*, i.e. possible pieces of information that player i may have about the payoff state; Θ_0^* instead represents any residual uncertainty that is left after pooling all players' information. The interpretation is that when the *payoff state* is $\theta = (\theta_0, \theta_1, ..., \theta_n)$, player i *knows* that $\theta \in \Theta_0^* \times \{\theta_i\} \times \Theta_{-i}^*$, where $\Theta_{-i}^* = \times_{j \in N \setminus \{i\}} \Theta_j^*$. The tuple $\left\langle \Theta_0^*, (\Theta_i^*, u_i)_{i \in N} \right\rangle$, where $u_i : \mathcal{Z} \times \Theta^* \to \mathbb{R}$ for each $i \in N$, represents players' information about everyone's preferences, and is referred to as *information structure.*[12] Special cases of interest are those of *private values (PV)*, in which each u_i depends on Θ_i^* only, and the case in which the u_i's depend on Θ_0^* only and players have *no information (NI)* about payoffs (i.e., without loss of generality, $\Theta^* = \Theta_0^*$). In a *PV-environment*, each player's payoffs only depend on what he *knows*. Hence, in a *PV-environment*, it is CK that everybody *knows* his own payoffs: Uncertainty may only concern the opponents' payoffs, and higher order beliefs. In contrast, in *NI-environments* players have no knowledge of the payoff state, and all CK-assumptions can be relaxed. In particular, each player does not *know* his own preferences over the terminal histories: He merely holds beliefs about that.[13]

This distinction is immaterial in WY's analysis of static games, but it becomes crucial in dynamic environments: In the *NI-case*, even if i puts probability one on θ_0 (i.e. i is *certain* that he has preferences $u_i(\theta_0) : \mathcal{Z} \to \mathbb{R}$), he may revise what he thinks his preferences are, if he observes something unexpected (e.g. unexpected moves of the opponents). This is not possible in *PV-environments*, in which i *knows* his own preferences over the terminal nodes. (Example 5 below illustrates one implication of such distinction).

[12] Notice that, the standard definition, an *information structure* specifies players' partitions and priors over the set of states (e.g. Dekel and Gul, 1997). Here, players' beliefs, i.e. their priors, are not specified. Appending players' *beliefs* to an information structure, in the terminology adopted here, delivers a *model* (see below).

[13] As discussed above, the notion of *richness* that is adopted qualifies what CK assumptions are being relaxed. For example, if $\Theta_0^* = \Theta^* = ([0,1]^n)^{\mathcal{Z}}$, then *all* CK-assumptions are relaxed, as in this case Θ^* represents all the possible preference profiles over $\mathcal{Z}$. This specification of Θ^* satisfies the richness condition introduced in section 5.

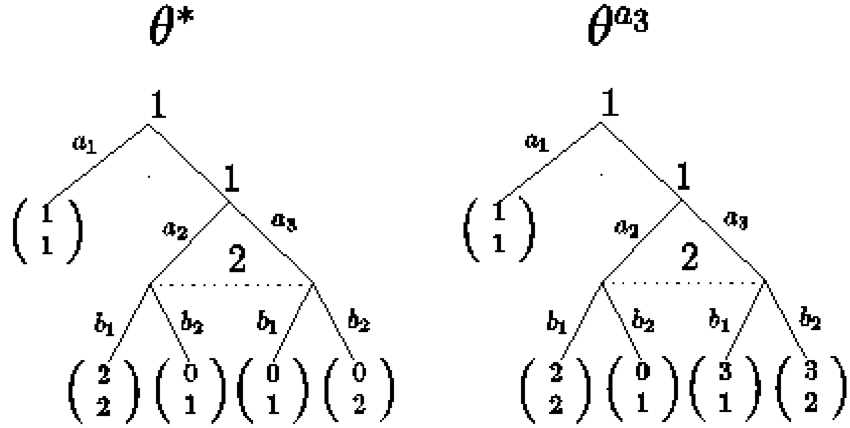
The Interim Approach. An *information structure* and an *extensive form* define a *dynamic game with payoff uncertainty*, but do not complete the description of the strategic situation: Players' beliefs about what they don't know must be specified, i.e. beliefs about $\Theta_0^* \times \Theta_{-i}^*$ (*first order beliefs*), point beliefs about $\Theta_0^* \times \Theta_{-i}^*$ and the opponents' first order beliefs (*second order beliefs*), and so on. The Θ^**-based universal type space*, $\mathcal{U}(\Theta^*)$, can be thought of as the set of all such hierarchies of beliefs (Mertens and Zamir, 1985). Each element $t_i = (\theta_i, e_i) \in \mathcal{U}_i(\Theta^*)$ is a complete description of player i: His *information* θ_i (what he knows), and his *epistemic type* e_i (his beliefs about what he doesn't know, $\Theta_0^* \times \Theta_{-i}^* \times E_{-i}^*$).

It is important to stress one point: Players' hierarchies of beliefs (or *types*) are purely subjective states describing a player's view of the strategic situation he is facing. As such, they enter the analysis as a *datum* and should be regarded in isolation (i.e. player by player and type by type). Nothing prevents players' views of the world to be inconsistent with each other (i.e. to assign any probability to opponents' types other than the actual one). As analysts, we have nothing to say about these beliefs (they are part of the environment, exogenous variables); it is *given* such beliefs that we can apply game theoretic reasoning to make predictions about players' behavior (the endogenous variables). The name *"Interim" Sequential Rationalizability* is meant to emphasize this point.

Robustness in the Universal Model. To explore what predictions retain their validity when *all* CK-assumptions are relaxed, we must specify a *rich* space of uncertainty Θ^* and look at players' types in the universal space $\mathcal{U}(\Theta^*)$: Assuming CK of $\langle \Theta^*, \mathcal{U}(\Theta^*) \rangle$ entails no further loss of generality; $\langle \Theta^*, \mathcal{U}(\Theta^*) \rangle$ will thus be referred to as the *universal model.* A solution concept assigns to each player's hierarchy a set of strategies.

In modelling a strategic situation, applied theorists typically select a subset of the possible hierarchies to focus on. To the extent that the "true" hierarchies are understood to be only close to the ones considered in a specific model, the concern for robustness of the theory's predictions translates into a continuity property of the solution concept correspondence. In this paper a solution concept is "robust" if it never rules out strategies that are not ruled out for arbitrarily close hierarchies of beliefs: This is equivalent to requiring upper hemicontinuity of the solution concept correspondence on $\mathcal{U}(\Theta^*)$.

Example 3 (Non-Robustness of EFR-predictions). Consider the situation in figure 2, letting $\Theta^* = \{\theta^*, \theta^{a_3}\}$: in state θ^* payoffs are the same as in example 1; at state θ^{a_3} only 1's payoffs are changed, so that a_3 is strictly dominant (hence, Θ^* allows to relax *CK* that a_3 is *not dominant*). Suppose that player 1 knows the true state, while 2 doesn't. This is a PV-environment, as each player knows his own payoffs (and this is CK), and can be represented letting $\Theta_1^* = \{\theta_1^*, \theta_1^{a_3}\}$ and Θ_2^* be a singleton (it can thus be ignored).Let $t^* \in \mathcal{U}(\Theta^*)$ denote the type profile representing *common certainty* of θ^*. (It is not CK, because t_2^* only *knows*



his own payoffs. He puts probability one on t_1^*, with payoffs as in state θ^*, but he doesn't know that. t_1^* instead *knows* everybody's payoffs). A reasoning similar to that in example 1 implies that $\{a_1, a_2\}$ and $\{b_1, b_2\}$ are the sets of $\mathcal{ISR}$ strategies for t_1^* and t_2^*: Written $\mathcal{ISR}(t^*) = \{a_1, a_2\} \times \{b_1, b_2\}$.

To illustrate the robustness question just discussed, a sequence of types $\{t^m\} \subseteq \mathcal{U}(\Theta^*)$ will be constructed, converging to t^*, such that (a_1, b_2) is the unique $\mathcal{ISR}$-outcome for each t^m: since it is the *unique* $\mathcal{ISR}$-outcome along the sequence, any refinement of $\mathcal{ISR}$ that rules out (a_1, b_2) for type t^* would not be u.h.c. on $\mathcal{U}(\Theta^*)$, i.e. not "robust". (Notice that (a_1, b_2) was ruled out by *EFR* in example 2.)

Fix $\varepsilon \in \left(0, \frac{1}{6}\right)$ and let $p \in \left(0, \frac{\varepsilon}{(1-2\varepsilon)}\right)$. Consider the set of type profiles $T_1^\varepsilon \times T_2^\varepsilon \subseteq \mathcal{U}(\Theta^*)$ s.t.: $T_1^\varepsilon = \{-1^{a_3}, 1^*, 1^{a_3}, 3^*, 3^{a_3}, 5^*, 5^{a_3}...\}$ and $T_2^\varepsilon = \{0, 2, 4, ...\}$. Types $k^\star$ ($k = -1, 1, 3, ...$, $\star = *, a_3$) are player 1's types who *know* that the true state is $\theta^\star$; 2's types only know their own payoffs, which are constant across states, but don't know the opponent's type. Suppose that beliefs are described as follows. Type -1^{a_3} puts probability one on facing type 0; type 0 assigns probability $\frac{1}{1+p}$ to type -1^{a_3}, and complementary probability to types 1^* and 1^{a_3}, with weights $(1-\varepsilon)$ and ε, respectively. Similarly, for all $k = 2, 4, ...$ player 2's type k puts probability $\frac{1}{1+p}$ on 1's types $(k-1)^*$ and $(k-1)^{a_3}$, with weights $(1-\varepsilon)$ and ε respectively, and complementary probability $\frac{p}{1+p}$ on the $(k+1)$-types, with weight $(1-\varepsilon)$ on $(k+1)^*$ and ε on $(k+1)^{a_3}$. For all other types of player 1, with $k = 1, 3, ...$, and $\star = *, a_3$, type $k^\star$ puts probability $\frac{1}{1+p}$ on 2's type $k-1$, and complementary probability on 2's type $k+1$. Similarly to example 1, the increasing sequence of even k's and odd k^*'s converges to t^* as we let ε approach 0. It will be shown that player 2's types $0, 2, 4, ...$ only play b_1, while 1's types $1^*, 3^*, ...$ only play a_1: All types k^{a_3} ($k = -1, 1, 3, ...$) would play a_3, for they *know* it is dominant. Type 0 puts probability $\frac{1}{1+p}$ on type -1, who plays a_3; given these initial beliefs, type 0's conditional conjectures after In must put probability at least $\frac{1}{1+p}$ on a_3 being played, which makes b_2 optimal for him. Type 1^* also puts probability $\frac{1}{1+p}$ on type 0, who plays b_2, thus a_1 is the unique best response. Type 2's initial beliefs are such that type 1^* plays a_1 and types 1^{a_3} and 3^{a_3} play a_3. Hence, the probability of a_3 being played, conditional on In being observed, must be no smaller than

$$\begin{aligned}\Pr\left(\theta^{a_3}|\text{not } 1^*\right) &= \frac{\varepsilon}{1-\left(\frac{1}{1+p}\right)(1-\varepsilon)} \\ &= \frac{(1+p)\,\varepsilon}{p+\varepsilon}\end{aligned}$$

Given that $p < \frac{\varepsilon}{(1-2\varepsilon)}$, this probability is greater than $\frac{1}{2}$. Hence, playing b_2 is the unique best response, irrespective of type 2's conjectures about 3^*'s behavior. Given this, type 3^* also plays a_1. The reasoning can be iterated, so that for all types $1^*, 3^*, 5^*, \dots$, a_1 is the unique $\mathcal{ISR}$ strategy, while for types $0, 2, 4, \dots$ of player 2, strategy b_2 is.[14] $\square$

Upper hemicontinuity in the universal model addresses a specific robustness question: robustness, with respect to small "mistakes" in the modelling choice of which subset of players' hierarchies to consider. Clearly, a solution concept that never rules out anything is robust, but not very interesting: one way to solve this trade-off is then to look for a strongest robust solution concept. It will be shown (proposition 4) that when all CK-assumptions are relaxed (i.e. an even larger Θ^* is considered), for any type $t \in \mathcal{U}(\Theta^*)$ and any $s \in \mathcal{ISR}(t)$, there exists a sequence $t^m \to t$ such that $\{s\} = \mathcal{ISR}(t^m)$ for any m. Furthermore, $\mathcal{ISR}$ is u.h.c. on $\mathcal{U}(\Theta^*)$ (proposition 1). Hence, $\mathcal{ISR}$ is a *strongest robust solution concept.*[15]

Type Spaces, Models and Invariance. When applied theorists choose a subset of Θ^*-hierarchies to focus on, they typically represent them by means of (non universal) Θ^**-based type spaces*, rather than elements of $\mathcal{U}(\Theta^*)$.

Definition 1 *A* Θ^*-based type space *is a tuple*

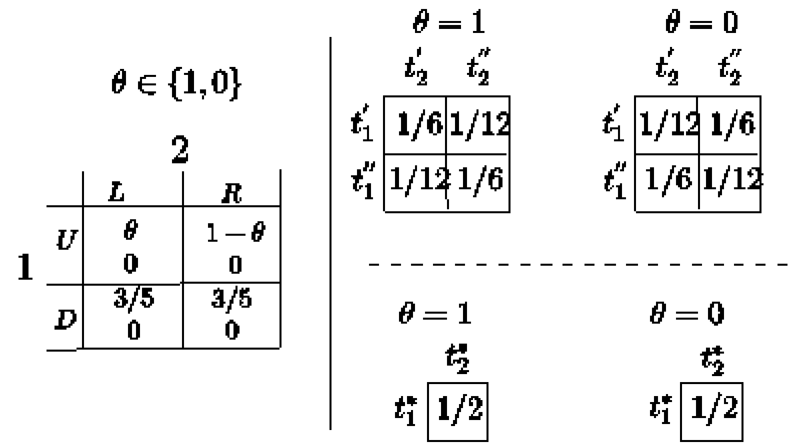
$$\mathcal{T}_{\Theta^*} = \left\langle \Theta_0^*, \left(\Theta_i^*, E_i, T_{i,\Theta^*}, \tau_i\right)_{i\in N}\right\rangle$$

such that $T_{i,\Theta^*} = \Theta_i^* \times E_i$ *and* $\tau_i : T_{i,\Theta^*} \to \Delta\left(\Theta_0^* \times T_{-i,\Theta^*}\right)$

Each type t_i in T_{i,Θ^*} corresponds to a Θ^*-hierarchy for player i. Representing a hierarchy as a type in a (non-universal) type space $\mathcal{T}_{\Theta^*}$ rather than an element of $\mathcal{U}(\Theta^*)$ does not change the CK-assumptions on the *information structure*, but it does impose CK-assumptions on

[14] The main intuition between the non-robustness of EFR in example 3 is the same as the one behind the main results of Dekel and Fudenberg (1990). A major difference between that and the present work though is given by the generic uniqueness result.

[15] $\mathcal{ISR}$ is *the* strongest among the class of solution concepts that satisfy *initial common certainty of sequential rationality.*



players' hierarchies of beliefs, and their correlation with the states of nature θ_0. A solution concept is *type space-invariant* if the behavior prescribed for a given hierarchy does not depend on whether it is represented as an element of $\mathcal{U}(\Theta^*)$ or of a different $\mathcal{T}_{\Theta^*}$. Thus, *type space-invariance* is also a robustness property: robustness, respect to the introduction of the extra CK-assumptions on players' beliefs imposed by non-universal type spaces. The following example, borrowed from Dekel et al. (2007), illustrates the point in the context of a static game.

Example 4 (Type Space-Invariance). Consider the game in figure 3. Suppose that players have no information about the states $\Theta^* = \{0, 1\}$, that they both put $\frac{1}{2}$ probability on each state and that this is common knowledge.In figure 3 two type spaces are used to model this situation: In the bottom type space T^*, each player has only one type, t_i^*, which puts probability $\frac{1}{2}$ on each pair $\left(0, t_{-i}^*\right)$ and $\left(1, t_{-i}^*\right)$. In the top type space T', each player has two types, t_i' and t_i''. The two matrices represent the common prior on T'. Notice that each type in T_i' corresponds to the same Θ^*-hierarchy as t^*, i.e. represents the beliefs that the two states are equally likely and that this is common knowledge. A *type space-invariant* solution concept would deliver the same predictions for t_i^*, t_i' and t_i'': That is, the behavior is completely determined by the Θ^*-hierarchies, irrespective of the type space used to represent them. In particular, modelling hierarchies by means of types in smaller type spaces or as elements of $\mathcal{U}(\Theta^*)$ doesn't change the outcomes.

Applying rationalizability to the interim normal forms obtained by the two type spaces in figure 3 (a solution concept known as *Interim Independent Rationalizability, IIR*), one obtains that both actions are rationalizable for types t_1' and t_1'', while only D is rationalizable for t_1^*. Hence, *IIR* is *not* type space-invariant.[16]□

[16] Ely and Peski (2006) and Dekel et al. (2007) observed that *IIR* is not type space-invariant. Dekel et al. (2007) thus developed the concept of *Interim Correlated Rationalizability (ICR)*, proving that it is *type*

Proposition 2 shows that $\mathcal{ISR}$ is *type space-invariant* under all information structures.

In writing down a game, as analysts we typically make CK-assumptions not only on players' beliefs, but also on payoffs. That is, not only we use *non-universal* Θ^**-based type spaces*, but we also impose CK of a smaller space of uncertainty $\Theta \subseteq \Theta^*$, and work with Θ*-based type spaces.*

Definition 2 *A* model *of the environment is a* Θ*-based type space, where* Θ *is such that* $\Theta_k \subseteq \Theta_k^*$ *for each* $k = 0, ..., n$.

Each type in a *model* induces a Θ-hierarchy, and hence a Θ^*-hierarchy. A solution concept is *model invariant* if the behavior is completely determined by the Θ^*-hierarchies irrespective of the model they are represented in. *Model invariance* is a stronger robustness property than *type space-invariance*, as it also requires robustness to the introduction of extra CK-assumptions on the information structure.[17] For concreteness, consider the following example.

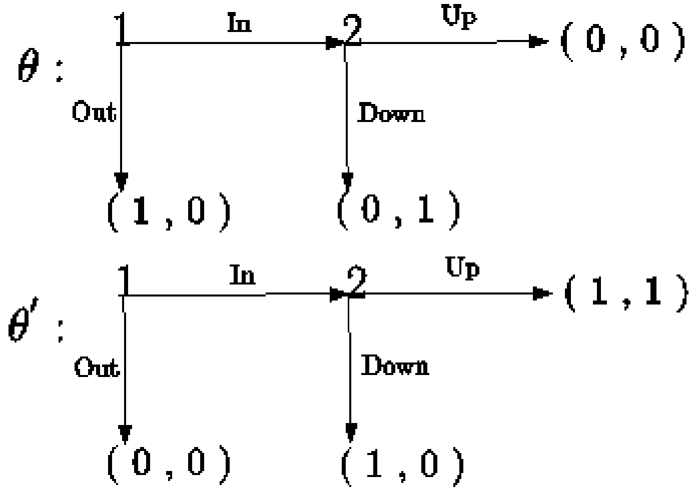
Example 5 (Model Invariance). It is tempting to consider the common certainty type $t^* \in \mathcal{U}(\Theta^*)$ in example 3 as "equivalent", in some sense, to the CK-type in example 1: They share the same hierarchy of beliefs over Θ^*, afterall. Admittedly, example 3 was designed to hint at that, but the careful reader may have noticed that such "equivalence" was never explicitly mentioned in example 3. The reason is that such interpretation is not always legitimate. *Model-invariance* addresses specifically the question whether envisioning "CK-types" such as that in example 1 as "common certainty-types" such as t^* in example 3 affects the predictions of a solution concept.

To illustrate the point, consider a *NI-environment* $\left\langle \Theta^*, (u_i)_{i=1,2} \right\rangle$, where $\Theta_0^* = \{\theta, \theta'\}$ and Θ_1 and Θ_2 are singletons. The extensive form and the payoffs in each state are represented in figure 4.

Let $\mathcal{T} = \left\langle \Theta, (T_i, \tau_i)_{i=1,2} \right\rangle$ and $\mathcal{T}' = \left\langle \Theta', (T_i', \tau_i')_{i=1,2} \right\rangle$ be two *models* such that $\Theta = \{\theta\}$, $T_i = \{t_i\}$ and $\tau_i(t_i)[\theta, t_{-i}] = 1$ for $i = 1, 2$ and similarly for $\mathcal{T}'$ where $\Theta' = \{\theta'\}$. Thus, t and t' represent *common knowledge* of θ and θ', respectively. Also consider the Θ^*-based type space $\mathcal{T}'' = \left\langle \Theta^*, (T_i'', \tau_i'')_{i=1,2} \right\rangle$ s.t. $T_i'' = \left\{t_i^a, t_i^b\right\}$, $\tau_i''(t_i^a)\left[\theta, t_{-i}^a\right] = 1$ and $\tau_i''\left(t_i^b\right)\left[\theta', t_{-i}^b\right] = 1$: Types profiles t^a and t^b represent *common certainty* of θ and θ', respectively; thus they share the same Θ^*-hierarchy as types t and t'. But differently from types t and t', t_i^a and t_i^b do not *know* θ or θ'.

space-invariant.

[17] A Θ-based type space differs from a Θ^*-based type space in which there is *common certainty* of Θ for the fact that it imposes *common knowledge* of Θ. This distinction is inconsequential in static settings, in which knowledge can be replaced by probability-one belief, but it is crucial in dynamic settings, in which the notion of *model-invariance* is indeed very demanding.



$\mathcal{ISR}$ applied to the two complete information games delivers the backward induction outcome, i.e. $\mathcal{ISR}(t) = \{Out\} \times \{Down\}$ and $\mathcal{ISR}(t') = \{In\} \times \{Up\}$. In contrast, both Up and $Down$ are consistent with $\mathcal{ISR}$ for type t_2^a: At the beginning of the game, type t_2^a is certain (but does not *know*) that the payoff state is θ, and that he is facing type t_1^a,for which Out is dominant: Hence, the second node is unexpected. At this point, t_2^a is surprised, and may as well put probability one on payoff state θ', and rationally play $Down$.□

Example 5 illustrated the importance of distinguishing *knowledge* from *certainty* in dynamic games, and also showed that $\mathcal{ISR}$ is not *model-invariant* in *NI-environments*. Proposition 3 in section 4.3 shows that $\mathcal{ISR}$ is *model-invariant* in *PV-environments*. (Hence, from the point of view of $\mathcal{ISR}$, it was legitimate to envision the CK-type of example 1 as the common belief type t^* in example 3, afterall).

***WY's program* and Robustness in dynamic environments.** Three notions of robustness have thus been introduced: 1) *model invariance*: robustness with respect to the relaxation of CK-assumptions on the information structure; 2) *type space invariance*: robustness with respect to the relaxation of CK-assumptions on players' hierarchies of beliefs; 3) *u.h.c. in the universal model*: robustness with respect to misspecification of the hierarchies of beliefs.

The first two coincide in static environments, and WY's thus focused on (3) only.[18] These distinctions instead are crucial for dynamic environments. Each of these concepts addresses a well-defined robustness question, and each has intrinsic interest. But only if both (1) and (2) are satisfied, the CK assumptions implicit in a standard game can be embedded in the universal model, and WY's program entirely applied.

[18] *Type space-invariance* for Interim Correlated Rationalizability was proved by Dekel et al. (2007).

3 Game Theoretic Framework

The analysis that follows applies to dynamic *environments*, defined by an *extensive form* and a *model*, i.e. a specification of players' interactive beliefs about everyone's preferences. The present work is concerned with robustness of solution concepts to different specifications of *players' model*, i.e. the *extensive form* will be maintained fixed throughout, and the *model* varied. These concepts are formally introduced next:

Extensive Forms: An extensive form is defined by a tuple

$$\Gamma = \left\langle N, \bar{\mathcal{H}}, \mathcal{H}, \mathcal{Z}, (A_i)_{i\in N} \right\rangle$$

where $N = \{1, ..., n\}$ is the set of players; for each player i, A_i is the (finite) set of his possible actions; a finite collection $\bar{\mathcal{H}}$ of histories (concatenations of action profiles), partitioned into the set of terminal histories $\mathcal{Z}$ and the set of partial histories $\mathcal{H}$ (which includes the *empty history* ϕ). As the game unfolds, the partial history h that has just occurred becomes public information and perfectly recalled by all players. At some stages there may be simultaneous moves. For each partial history $h \in \mathcal{H}$ and player $i \in N$, let $A_i(h)$ denote the (finite) set of actions available to player i at history h, and let $A(h) = \times_{i\in N} A_i(h)$ and $A_{-i}(h) = \times_{j\in N\setminus\{i\}} A_j(h)$.[19] Without loss of generality, $A_i(h)$ is assumed to be non-empty: player i is *inactive* at h if $|A_i(h)| = 1$; he is *active* otherwise. If there is only one active player at each h, the game has *perfect information*.[20]

Pure strategies of player i assign to each partial history $h \in \mathcal{H}$ an action in the set $A_i(h)$. Let S_i denote the set of *reduced strategies* (plans of actions) of player i. Two strategies correspond to the same reduced strategy $s_i \in S_i$ if and only if they are realization-equivalent to s_i, that is, they preclude the same collection of histories and for every non precluded history h they select the same action, $s_i(h)$.[21] Each profile of reduced strategies s induces a unique terminal history $\mathbf{z}(s) \in \mathcal{Z}$. For each $h \in \mathcal{H}$, let $S_i(h)$ be the set of s_i allowing history h (meaning that there is some s_{-i} such that h is a prefix of $\mathbf{z}(s_i, s_{-i})$). $\mathcal{H}(s_i)$ is the set of histories not precluded by s_i: $\mathcal{H}(s_i) = \{h \in \mathcal{H} : s_i \in S_i(h)\}$.

[19] Formally: $A_i(h) = \left\{a_i \in A_i : \exists a_{-i} \in A_{-i} \text{s.t.} (h, (a_i, a_{-i})) \in \bar{\mathcal{H}}\right\}$

[20] To allow for random moves, a fictitious player 0 can be introduced: at each history h, player 0 can take an action $a_0 \in A_0(h)$, to be interpreted as a move of nature.

[21] The definition of strategies requires the specification of players' behavior also at histories that are precluded by the strategy itself. This is necessary for *equilibrium concepts* based on sequential rationality: in the equilibrium, the behavior specified by s_i at histories $h \notin H(s_i)$ represents the opponents' beliefs about i's behavior in case he has deviated from s_i and h is reached (see Rubinstein, 1991). In this paper a *non-equilibrium approach* is considered, players' beliefs about the opponents' behavior at each history are explicitely modelled. Hence attention can be restricted to players' *plans of actions* (or reduced strategies).

Players' Models:

Players have preferences over the terminal histories, represented by payoff functions $u_i : \mathcal{Z} \to \mathbb{R}$ for each i. Attaching a preference profile $(u_i)_{i \in N}$ to an extensive form delivers a *multistage game with observable actions* (with complete information).[22] To model *incomplete information*, payoff functions are parametrized on a fundamental space of uncertainty Θ^* such that

$$\Theta^* = \Theta_0^* \times \Theta_1^* \times ... \times \Theta_n^*,$$

and $u_i : \mathcal{Z} \times \Theta^* \to \mathbb{R}$.[23] Elements of Θ^* are referred to as *payoff states.* For each $i = 1, ..., n$, Θ_i^* is the set of player i's *payoff types*, i.e. possible information that player i may have about the payoff state. Θ_0^* instead represents the *states of nature* (any residual uncertainty that is left after pooling all players' information). Each Θ_k^* ($k = 0, 1, ..., n$) is assumed non-empty, Polish and convex, and each $u_i : \mathcal{Z} \times \Theta^* \to \mathbb{R}$ continuous.[24] The tuple $\langle \Theta^*, (u_i)_{i \in N} \rangle$ represents the *fundamental information structure*, and describes players' information about everyone's payoffs. Tuples $\langle \Theta, (u_i)_{i \in N} \rangle$ such that $\Theta = \times_{k=0}^n \Theta_k$ and $\Theta_k \subseteq \Theta_k^*$ is Polish for each k are referred to as (Θ^*-based) *information structures.*

For ease of reference, two special cases are defined:

Definition 3 Private Values (PV): *for each* $i \in N$*,* $u_i : \mathcal{Z} \times \Theta_i \to \mathbb{R}$

Definition 4 No Information (NI): *for each* $i \in N$*,* $u_i : \mathcal{Z} \times \Theta_0 \to \mathbb{R}$

A *model* (or Bayesian model) is a tuple $\langle \Theta, \mathcal{T}, (\hat{u}_i)_{i \in N} \rangle$ such that: (i) Θ is a (Θ^*-based) *information structure*; (ii) $\mathcal{T}_\Theta$ is a Θ-*based type space*, that is:

$$\mathcal{T} = \langle \Theta, (T_i, \boldsymbol{\theta}_i, \tau_i)_{i \in N} \rangle$$

such that for each $i \in N$, T_i is a compact set of i's *types*, and $\boldsymbol{\theta}_i : T_i \to \Theta_i$ (onto and continuous) assigns to each type a *payoff type* and $\tau_i : T_i \to \Delta(\Theta_0 \times T_{-i})$ (continuous) assigns to each type a *belief* about the states of nature and the opponents' types; (iii) $\hat{u}_i : \mathcal{Z} \times \Theta_0 \times T \to \mathbb{R}$ are such that for each $(z, \theta_0, t) \in \mathcal{Z} \times \Theta_0 \times T$, $\hat{u}_i(z, \theta_0, t) = u_i(z, \theta_0, \boldsymbol{\theta}(t))$.

Attaching a *model* $\langle \Theta, \mathcal{T}, (\hat{u}_i)_{i \in N} \rangle$ to the extensive form Γ delivers a *Bayesian multistage game with observable actions.*

[22] Cf. Fudenberg and Tirole (1991), §3.3.

[23] This representation is without loss of generality: for example, taking the underlying space of uncertainty $\Theta^* \equiv [0,1]^{\mathcal{Z}}$ and $u_i(z, \theta) = \theta(z)$

[24] Convexity is only required for the results in section 5. Anyway, it can be dropped for Θ_0^* without affecting any result (included those of section 5).

Hierarchies, Type Spaces and the Universal Model:

Fix an information structure Θ. Players' hierarchies are defined as follows: for each $i \in N$ let $Z_i^0 = \Theta_{-i}$ and for $k \geq 1$ define $Z_i^k = Z_i^{k-1} \times \Delta\left(Z_{-i}^{k-1}\right)$. An element of $\Delta\left(Z_i^{k-1}\right)$ is a Θ-*based* k-*order belief*, one of $\times_{k\geq 1}\Delta\left(Z_i^{k-1}\right)$ a Θ-*based beliefs hierarchy* or Θ-*hierarchy*. For each i, let $T_{i,\,\Theta}^*$ denote the set of i's collectively coherent Θ-hierarchies (Mertens and Zamir, 1985).[25] Each type in a model $\left\langle \Theta, \mathcal{T}, (\hat{u}_i)_{i\in N}\right\rangle$ induces a Θ-hierarchy: For each $t_i \in T_i$, let $\hat{\pi}_i^0(t_i) = \boldsymbol{\theta}_i(t_i)$; construct mappings $\hat{\pi}_i^k : T_i \to \Delta(Z_i^{k-1})$ recursively for all $i \in I$ and $k \geq 1$, s.t. $\hat{\pi}_i^1(t_i)$ is the pushforward of $\tau_i(t_i)$ given by the map from $\Theta \times T_{-i}$ to Z_i^0, such that

$$(\theta_0, t_{-i}) \mapsto \left(\theta_0, \hat{\pi}_{-i}^0(t_{-i})\right),$$

and $\hat{\pi}_i^k(t_i)$ is the pushforward of $\tau_i(t_i)$ given by the map from $\Theta_0 \times T_{-i}$ to Z_i^{k-1} such that

$$(\theta_0, t_{-i}) \mapsto \left(\theta_0, \hat{\pi}_{-i}^0(t_{-i}), \hat{\pi}_{-i}^1(t_{-i}), \ldots, \hat{\pi}_{-i}^{k-1}(t_{-i})\right).$$

The mappings $\hat{\pi}_i^* : T_i \to T_i^*$ thus constructed, i.e.

$$t_i \mapsto \hat{\pi}_i^*(t_i) = \left(\hat{\pi}_i^0(t_i), \hat{\pi}_i^1(t_i), \hat{\pi}_i^2(t_i), \ldots\right),$$

assign to each type in a Θ-based type space, the corresponding Θ-*hierarchy of beliefs*.

From Mertens and Zamir (1985) and Brandenburger and Dekel (1993) we know that when the $T_{i,\,\Theta}^*$'s are endowed with the product topology, there is a homeomorphism

$$\boldsymbol{\phi}_i : T_{i,\,\Theta}^* \longrightarrow \Delta\left(\Theta_{-i} \times T_{-i,\,\Theta}^*\right)$$

that preserves beliefs of all orders: for all $t_i^* = (\pi_i^1, \pi_i^2, \ldots) \in T_{i,\,\Theta}^*$,

$$marg_{Z_i^{k-1}}\ \boldsymbol{\phi}_i(t_i^*) = \pi_i^k \quad \forall k \geq 1.$$

Hence, the tuple $\mathcal{T}_\Theta^* = \left\langle \Theta, \left(T_{i,\,\Theta}^*, \boldsymbol{\theta}_i^*, \tau_i^*\right)_{i\in N}\right\rangle$ with $\tau_i^* = \phi_i$ is a type space. It will be referred to as the Θ-*based Universal Type Space*. The mappings $\hat{\pi}_i^* : T_i \to T_i^*$ constitute the *canonical belief morphism* from $\mathcal{T}_\Theta$ to $\mathcal{T}_\Theta^*$.

A (Θ-based) *finite type* is any element $t_i \in T_{i,\,\Theta}^*$ such that $\left|supp\ \tau_{t_i}^*\right| < \infty$. The set of finite types is denoted by $\hat{T}_{i,\,\Theta} \subseteq T_{i,\,\Theta}^*$. A model is *finite* if $|\Theta \times T_\Theta|$ is finite.

The pair $\left\langle \Theta^*, \mathcal{T}_{\Theta^*}^*, (\hat{u}_i)_{i\in N}\right\rangle$ is referred to as the *universal model*.

[25] Formally, let $H_i^1 = \Delta\left(Z_i^0\right)$, and for all $k \geq 1$,

$$H_i^{k+1} = \left\{ \begin{array}{c} (\pi_i^1, \ldots, \pi_i^{k+1}) \in H_i^k \times \Delta\left(Z_i^k\right) : \\ \mathrm{mrg}_{Z_i^{k-1}}\ \pi_i^{k+1} = \pi_i^k \text{ and } \pi_i^{k+1}\left[H_{-i}^k\right] = 1 \end{array} \right\}$$

Let $T_i^* = \left\{\left(\pi_i^1, \pi_i^2, ...\right) \in \times_{k\geq 1}\Delta\left(Z_i^{k-1}\right) : (\pi_i^1, \ldots, \pi_i^k) \in H_i^k \text{ for all } k \geq 1\right\}$

3.1 Interim Sequential Rationalizability

Fix a model $\left\langle \Theta, \mathcal{T}, (\hat{u}_i)_{i\in N}\right\rangle$, and consider the induced Bayesian game

$$\Gamma^{\mathcal{T}} = \left\langle N, \bar{\mathcal{H}}, \Theta, (T_i, \boldsymbol{\theta}_i, \tau_i, \hat{u}_i)_{i\in N}\right\rangle$$

Player i's *conjectures* are represented by Conditional Probability Systems (CPS), i.e. arrays of conditional beliefs, one for each history, (denoted by $\mu^i = \{\mu^i(\cdot|h)\}_{h\in\mathcal{H}} \in \Delta^{\mathcal{H}}(\Theta_0 \times T_{-i} \times S_{-i})$) such that: *(i)* μ^i is consistent with Bayes' rule whenever possible, and *(ii)* for each $h \in \mathcal{H}$, $\mu^i(\cdot|h) \in \Delta(\Theta_0 \times T_{-i} \times S_{-i}(h))$.[26]

To avoid confusion between the *exogenous* and the *endogenous* beliefs, in the following the term *beliefs* will be restricted to the former and represented by players' types; the term *conjectures* will be used to the latter and represented by the CPSs just introduced.

For each type $t_i \in T_i$, his *consistent conjectures* are

$$\Phi_i(t_i) = \left\{\mu^i \in \Delta(\Theta_0 \times T_{-i} \times S_{-i}) : marg_{\Theta_0\times T_{-i}}\mu^i(\cdot|\phi) = \tau_i(t_i)\right\}$$

That is, t_i's consistent conjectures agree with his beliefs on the environment at the beginning of the game. The set of *Sequential Best Responses* for type t_i to conjectures $\mu^i \in \Delta^{\mathcal{H}}(\Theta_0 \times T_{-i} \times S_{-i})$, denoted by $r_i(\mu^i|t_i)$, is defined as:

$$\begin{aligned} &s_i \in r_i\left(\mu^i|t_i\right) \text{ if and only if } \forall h \in \mathcal{H}(s_i) \\ &s_i \in \arg\max_{s_i'\in S_i(h)} \int_{\Theta_0\times T_{-i}\times S_{-i}} \hat{u}_i\left(\mathbf{z}(s_i, s_{-i}), \theta_0, t_{-i}, t_i\right) d\mu^i(\theta_0, t_{-i}, s_{-i}|h)\end{aligned}$$

Definition 5 *A strategy $s_i \in S_i$ is* sequentially rational for type t_i*, written $s_i \in r_i(t_i)$, if there exists $\mu^i \in \Phi_i(t_i)$ such that $s_i \in r_i(\mu^i|t_i)$.*

The notion of *sequential rationality* is stronger than (normal-form) *rationality*, which would only require that a player optimizes with respect to his initial conjectures, hence putting no restrictions on behavior at zero-probability histories. Notice also that working with reduced strategies, the restrictions only concern histories that are reachable by s_i.

Interim Sequential Rationalizability ($\mathcal{ISR}$) is introduced next:

Definition 6 ($\mathcal{ISR}$) *For each $i \in N$, let $\mathcal{ISR}_i^{\mathcal{T},0} = T_i \times S_i$. For each $k = 0, 1, ...,$ and $t_i \in T_i$, let $\mathcal{ISR}_i^{\mathcal{T},k}(t_i) = \left\{s_i \in S_i : (t_i, s_i) \in \mathcal{ISR}_i^{\mathcal{T},k}\right\}$, $\mathcal{ISR}^{\mathcal{T},k} = \times_{i=1,...,n}\mathcal{ISR}_i^{\mathcal{T},k}$ and $\mathcal{ISR}_{-i}^{\mathcal{T},k} = \times_{j\neq i,0}\mathcal{ISR}_j^{\mathcal{T},k}$. Define recursively, for $k = 1, 2, ...,$ and $t_i \in T_i$*

$$\mathcal{ISR}_i^{\mathcal{T},k}(t_i) = \left\{ \hat{s}_i \in \mathcal{ISR}_i^{\mathcal{T},k-1}(t_i) : \begin{array}{c} \exists \mu^i \in \Phi_i(t_i) \;\; s.t. \\ (1)\; \hat{s}_i \in r_i(\mu^i|t_i) \\ (2)\; \operatorname{supp}(\mu^i(\cdot|\phi)) \subseteq \Theta_0 \times \mathcal{ISR}_{-i}^{\mathcal{T},k-1} \end{array} \right\}$$

[26] See Battigalli and Siniscalchi (1999).

Finally: $\mathcal{ISR}^{\mathcal{T}} := \bigcap_{k \geq 0} \mathcal{ISR}^{\mathcal{T},k}$

$\mathcal{ISR}$ consists of an iterated deletion procedure: for each type t_i, reduced strategy s_i survives the k-th round of deletion if and only if s_i is sequentially rational for type t_i, with respect to conjectures μ^i that, *at the beginning of the game*, assign zero probability to pairs (t_{-i}, s_{-i}) that are inconsistent with the previous rounds of deletion. If history h is given zero probability by the conditional conjectures held the preceding node, i's conjectures at h may put positive probability on any triple $(\theta_0, t_{-i}, s_{-i}) \in \Theta_0 \times T_{-i} \times S_{-i}(h)$. The lack of restrictions on the conjectures held at unexpected histories rules out elements of forward induction reasoning (see example 2).[27] Notice also that $\mathcal{ISR}$ considers players' conjectures that allow for correlation in the opponents' strategies in the Bayesian game.[28] A more thorough discussion of the solution concept and its relation with the literature is postponed to section 6.

Example 2 (reprise): Example 2 already illustrated how the solution concept works. We repeat the argument just to familiarize with the procedure and the notation: a_3 is dominated by a_1, hence it is deleted at the first round. Given this, $\mathcal{ISR}$ restricts 2's *initial conjectures* to put zero probability on a_3, that is $\mu^2(a_3|\phi) = 0$. No further restrictions are imposed: in particular, conjectures can be $\hat{\mu}^2$ s.t. $\hat{\mu}^2(a_1|\phi) = 1$ and $\hat{\mu}^2(a_3|In) = 1$, which makes b_2 the unique sequential best response to $\hat{\mu}^2$. Given that neither b_1 nor b_2 is deleted at the second round, also a_1 cannot be deleted, and the procedure stops at $\mathcal{ISR} = \{a_1, a_2\} \times \{b_1, b_2\}$.■

An epistemic characterization for $\mathcal{ISR}$ is provided in appendix C, in terms of *"sequential rationality and initial common certainty of sequential rationality"* (proposition 6). Appendix D instead provides an alternative characterization of the solution concept, as Dekel and Fudenberg's (1990) procedure of one round of deletion of weakly dominated strategies followed by iterated deletion of strictly dominated strategies in the interim reduced normal form of the game.

4 Robustness(-es)

This section gathers properties of $\mathcal{ISR}$ that can be interpreted as robustness properties (see section 2). These are, respectively: upper hemicontinuity, type space-invariance and model-invariance. The first two hold for any information structure $\left\langle \Theta, (u_i)_{i \in N} \right\rangle$, while the latter holds in PV-environments, not in NI-environments.

[27] Example 2 also provides the main insight for the epistemic characterization provided in appendix C.

[28] Similarly to the distinction between *Interim Independent* and *Correlated Rationalizability* (Dekel et al., 2007), one could think of refining $\mathcal{ISR}$ so that players' conjectures on the opponents' behavior are measurable with respect to their types. Given the results in section 5, such refinement would not be robust.

4.1 Upper Hemicontinuity

As discussed in section 2, the upper hemicontinuity of $\mathcal{ISR}$ on the universal type space addresses a specific robustness question: the fact that $\mathcal{ISR}$ is u.h.c. means that any behavior that $\mathcal{ISR}$ rules out for a given Θ-hierarchy is also ruled out for all nearby hierarchies. Specifically, in the product topology, suppose that as analysts we know players hierarchies only up to a finite order k: if a solution concept is not u.h.c. it means that we can never rule out that by refining our model of beliefs of order higher than k, the solution concept allows behavior that is ruled out in the original model.[29]

Proposition 1 *For each $t \in T^*_\Theta$ and sequence $\{t^m\} : t^m \to t$ and for $\{s^m\} \subseteq S$ s.t. $s^m \to \hat{s}$ and $s^m \in \mathcal{ISR}(t^m)$ for all m, $\hat{s} \in \mathcal{ISR}(t)$.*
Proof. *(See appendix)* ■

4.2 Type Space Invariance

The intuition behind the type space-invariance of $\mathcal{ISR}$ is the same as for that of ICR in Dekel et al. (2007). The problem of type space-dependence originates in the different possibility of correlation between types and payoff states θ_0 allowed by different type spaces, representing the same set of hierarchies. Solution concepts, such as ICR and $\mathcal{ISR}$, that do not impose any condition of independence on players' conjectures about the opponents' strategies, already allow all the possible correlation and therefore are not affected by these differences across type spaces.[30] Hence, they are type space-invariant.

Proposition 2 *Let $\mathcal{T}$ and $\mathcal{T}'$ be two Θ-based type spaces. If $t_i \in T_i, t'_i \in T'_i$ are s.t. $\hat{\pi}^*(t_i) = \hat{\pi}^*(t'_i) \in T^*_{i,\Theta}$, then $ISR^{\mathcal{T}}(t_i) = \mathcal{ISR}^{\mathcal{T}'}(t'_i)$. Indeed, for each k, if for all $l \leq k$, $\hat{\pi}^l_i(t_i) = \hat{\pi}^l_i(t_i)$ then $\mathcal{ISR}^{\mathcal{T},k-1}_i(t_i) = \mathcal{ISR}^{\mathcal{T}',k-1}_i(t'_i)$*
Proof. *The result is an immediate implication of the epistemic characterization of $\mathcal{ISR}$ given in the appendix (proposition 6).* ■

4.3 Model Invariance

The model-dependence of $\mathcal{ISR}$ in NI-environments was shown in example 5 in section 2. The general intuition is the following: the only restrictions that $\mathcal{ISR}$ puts on the conjectures held at zero-probability histories come from the definition of the model (that they be concentrated on $\Theta_0 \times T_{-i}$): once surprised, type t_i may assign positive probability to pairs (θ_0, t_{-i}) that were initially given zero probability by the beliefs $\tau_i(t_i)$. Moving from smaller to larger models

[29] WY extensively discuss the interpretation of the product topology.

[30] The refinement of $\mathcal{ISR}$ mentioned in footnote 28 would not be type space-invariant.

(e.g. from a Θ-based type space, to a Θ^*-based type space) puts less and less restrictions on such possible conjectures. In *NI-environments*, larger Θ_0 also means more freedom to specify a player's beliefs about his own payoffs, thereby changing the set of sequential best responses. In contrast, in *PV-environments* players know their own payoffs: even if their beliefs are completely upset, they don't alter a type's preferences over the terminal nodes. This provides the intuition for the model-invariance result in PV-settings.

Let $\left\langle \Theta, \mathcal{T}_\Theta, (\hat{u}_i)_{i\in N}\right\rangle$ be a *model*. Any type $t_i \in T_{i,\,\Theta}$ induces a Θ-hierarchy $\hat{\pi}_i^*(t_i) = \left(\boldsymbol{\theta}_i(t_i), \hat{\pi}_i^1(t_i), ...\right) \in T_{i,\,\Theta}^*$. Since $\Theta = \times_{k=0}^n \Theta_k$ is such that $\Theta_k \subseteq \Theta_k^*$ for all k, $\hat{\pi}_i^*(t_i)$ can be naturally embedded in the Θ^*-based universal type space, and be seen as a Θ^*-based hierarchy. Let $\beta_i : T_{i,\,\Theta}^* \to T_{i,\,\Theta^*}^*$ denote such embedding and let $\kappa_i^* \equiv \beta_i \circ \hat{\pi}_i^*$.

Proposition 3 *Assume* private values*: For any finite model* $\left\langle \Theta, \mathcal{T}_\Theta, (\hat{u}_i)_{i\in N}\right\rangle$ *and any type* $t_i \in T_{i,\,\Theta}$, $\mathcal{ISR}_i^{\mathcal{T}_\Theta}(t_i) = \mathcal{ISR}_i^{\mathcal{T}_{\Theta^*}^*}(\kappa_i^*(t_i))$.
Proof. *(See appendix)* ■

5 The structure of $\mathcal{ISR}$ in the *Universal Model*

For their analysis, WY assume the existence of strict dominance regions for each of the players' strategies. This condition cannot be satisfied by the reduced normal form of a dynamic game, in which payoffs are defined over the terminal histories: No reduced strategy can be strictly dominated by another that only differs at nodes that can be prevented from being reached by some strategies of the opponents. An obvious candidate to solve the problem is to introduce *trembles*. It is important though (as WY point out) to pursue an analysis that does not involve trembles. In this way, the results can be used to address robustness questions for also for tremble-based solution concepts (e.g. *sequential equilibrium*), of major interest in dynamic games. This is realized here adopting a solution concept for the extensive form, $\mathcal{ISR}$. Being based on the natural notion of rationality for dynamic environments, this approach also has the advantage of addressing the question at hand more directly. Correspondingly, the appropriate richness condition for dynamic games entails a weakening of the notion of dominance.

Definition 7 *Strategy* s_i *is* conditionally dominant *at* $\theta \in \Theta^*$ *if* $\forall h \in \mathcal{H}(s_i)$, $\forall s_i' \in S_i(h)$, $\forall s_{-i} \in S_{-i}(h)$

$$s_i(h) \neq s_i'(h) \Rightarrow u_i(\mathbf{z}(s_i, s_{-i}), \theta) > u_i(\mathbf{z}(s_i', s_{-i}), \theta)$$

Richness Condition (RC): $\forall s \in S$, $\exists \theta^s = \left(\theta_0^s, \theta_i^s, \theta_{-i}^s\right) \in \Theta^*$: $\forall i \in N$, s_i is *conditionally dominant* at θ^s.

The main result in this section states that whenever a type profile $\hat{t}$ has multiple $\mathcal{ISR}$-outcomes, any of these is uniquely $\mathcal{ISR}$ for a sequence of players' types converging to $\hat{t}$

(proposition 4). An immediate implication is that any refinement of $\mathcal{ISR}$ (e.g. *EFR*) is not robust (See example 3 in section 2).

5.1 Sensitivity of Multiplicity to higher order beliefs

Proposition 4 *Under the* richness condition*, for any finite type profile* $\hat{t} \in \hat{T}_{\Theta^*}$ *and any* $s \in \mathcal{ISR}\left(\hat{t}\right)$*, there exists a sequence of finite type profiles* $\left\{\hat{t}^m\right\} \subseteq \hat{T}_{\Theta^*}$ *s.t.* $\hat{t}^m \to \hat{t}$ *as* $m \to \infty$ *and* $\mathcal{ISR}\left(\hat{t}^m\right) = \{s\}$ *for each* m*.*

Furthermore, for each m*,* $\hat{t}^m$ *belongs to a finite belief-closed subset of types,* $T^m \subseteq T^*_{\Theta^*}$*, such that for each* m *and each* $t \in T^m$ *,* $|\mathcal{ISR}(t)| = 1$*.*

Proposition 4 implies that any refinement of $\mathcal{ISR}$ is not u.h.c.; since by proposition 1, $\mathcal{ISR}$ is u.h.c. the following is true.

Corollary 1 $\mathcal{ISR}$ *is the strongest u.h.c. solution concept for dynamic Bayesian games*

The last part of the proposition, stating that the types in the sequence belong to a finite, belief-closed subset of types, is of particular interest in the context of PV-environments, in which $\mathcal{ISR}$ is also *model-invariant*. As discussed above, in these settings WY's program can be applied entirely, which means that such belief-closed set of types can be considered as *models*, in the sense of definition 2.

The proof of proposition 4 requires a substantial investment in additional concepts and notation. To facilitate the reader that is not interested in the technicalities, all these are confined to the next subsection. The argument is only sketched here.

The proof exploits a refinement of $\mathcal{ISR}$, $\mathcal{SSR}$, in which strategies that are never strict *sequential best responses* are deleted at each round. The proof is articulated in two main steps, which parallel WY's argument: in the first (lemma 2), it is shown that if $s_i \in \mathcal{ISR}_i(t_i)$ for finite type t_i, then s_i is also $\mathcal{SSR}$ for some type close to t_i; in the second (lemma 3), it is shown that by perturbing beliefs further, any $s_i \in \mathcal{SSR}_i(t'_i)$ can be made uniquely $\mathcal{ISR}$ for a type close to t'_i. The main points of departure from WY's are due to the necessity of breaking the ties between strategies at unreached information sets: this is necessary to obtain uniqueness in the converging sequence. Although notationally involved, the idea is very simple. Consider the sequence constructed in example 3: to obtain b_2 as the unique $\mathcal{ISR}$ for player 2, given that 1 would play a_1, it was necessary to perturb player 2's beliefs introducing, with arbitrarily small probability, the possibility that he is facing types $1^{a_3}, 3^{a_3}, ...$, i.e. types who know that a_3 is dominant. These "dominance types" play the role of trembles, and allow to break the tie between b_1 and b_2; but they do this in an indirect way, through the perturbation of the belief structure, in a way that naturally leads to completely mixed conjectures.

5.1.1 Proof of Proposition 4

Definition 8 *Let* $\mathcal{SSR}_i^0 = T^*_{i,\ \Theta^*} \times S_i$*. For each* $k = 0, 1, ...,$ *and* $t_i \in T^*_{i,\ \Theta^*}$*; let* $\mathcal{SSR}_i^k(t_i) = \left\{s_i \in S_i : (t_i, s_i) \in \mathcal{SSR}_i^k\right\}$*,* $\mathcal{SSR}^k = \times_{i=1,...,n}\mathcal{SSR}_i^k$ *and* $\mathcal{SSR}_{-i}^k = \times_{j\neq i,0}\mathcal{SSR}_j^k$*. Define recursively, for* $k = 1, 2, ...,$ *and* $t_i \in T_i$

$$\mathcal{SSR}_i^k(t_i) = \left\{ \hat{s}_i \in \mathcal{SSR}_i^k(t_i) : \begin{array}{r} \exists \mu^i \in \Phi_i(t_i) \ \ s.t. \\ (1).\ r_i(\mu^i|t_i) = \{\hat{s}_i\} \\ (2).\ \text{supp}(\mu^i(\cdot|\phi)) \subseteq \Theta_0^* \times \mathcal{SSR}_{-i}^{k-1} \\ (3).\ if\ t_{-i} \in \text{supp}(\text{marg}_{T^*_{-i,\ \Theta^*}}\mu^i(\cdot|\phi)) \\ and\ s_{-i} \in \mathcal{SSR}_{-i}^{k-1}(t_{-i}), \\ then:\ s_{-i} \in \text{supp}(\text{marg}_{S_{-i}}\mu^i(\cdot|\phi)) \end{array} \right\}$$

Finally: $\mathcal{SSR} = \bigcap_{k\geq 0} \mathcal{SSR}^k$

The following lemma states the standard fixed-point property for $\mathcal{SSR}$.

Lemma 1 *Let* $\{V_j\}_{j\in N}$ *be s.t. for each* $i \in N$*,* $V_i \subseteq S_i \times T_i$ *and* $\forall s_i \in V_i(t_i)$*,* $\exists \mu^i \in \Phi_i(t_i)$ *:*

- *(i)* $\text{supp}(\mu^i(\cdot|\phi)) \subseteq \times_{j\neq i}V_j$
- *(ii)* $\{s_i\} = r_i(\mu^i|t_i)$

Then: $V_i(t_i) \subseteq \mathcal{SSR}_i(t_i)$

Exploiting the richness condition, let $\bar{\Theta} \subset \Theta^*$ be a finite set of dominance states, s.t. $\forall s \in S$, $\exists!\theta^s \in \bar{\Theta}$ at which s is conditionally dominant. For each $s \in S$, let $\bar{t}^s \in T^*_{\Theta^*}$ be s.t. for each i, $\boldsymbol{\theta}_i(\bar{t}^s) = \theta_i^s$ and $\tau_i(\bar{t}_i^s)\left[\theta_0^s, \bar{t}_{-i}^s\right] = 1$. Let $\bar{T} = \{\bar{t}^s : s \in S\}$, and let $\bar{T}_j$ and $\bar{T}_{-j}$ denote the corresponding projections: $\bar{T}$ is a finite set of types representing common belief of θ^s, for each $s \in S$. Elements of $\bar{T}_i$ will be referred to as *dominance-types*, and will play the role of the k^{a_3}-types in example 3.

For each i and $s_i \in S_i$, let $\bar{T}_{-i}(s_i)$ be s.t. $\forall s_{-i} \in S_{-i}, \exists!\bar{t}_{-i} \in \bar{T}_{-i}(s_i)$ s.t. $\bar{t}_{-i} = \bar{t}_{-i}^{(s_i,s_{-i})}$. Notice that for each $\bar{t}_i^s \in \bar{T}_i$, $\{s_i\} = \mathcal{SSR}_i^1(\bar{t}_i^s)$, because s_i is the unique sequential best reply to any conjecture consistent with condition 3 for $\mathcal{SSR}$ in definition 8.

Lemma 2 *Under the* richness condition*, for any finite type* $t_i \in \hat{T}_{i,\ \Theta^*}$*, for any* $s_i \in \mathcal{ISR}_i(t_i)$*, there exists a sequence of finite types* $\{\iota^m(t_i, s_i)\}_{m\in\mathbb{N}}$*, such that:*

- *(i)* $\iota^m(t_i, s_i) \to t_i$ *as* $m \to \infty$
- *(ii)* $\forall m,\ s_i \in \mathcal{SSR}_i(\iota^m(t_i, s_i))$ *and* $\iota^m(t_i, s_i) \in \hat{T}_i$

- *(iii)* $\forall m$, *conjectures* $\mu^{s_i,m} \in \Phi(\iota^m(t_i,s_i))$ *s.t.* $\{s_i\} = r_i(\mu^{s_i,m}|\iota^m(t_i,s_i))$ *satisfy* $\bar{T}_{-i}(s_i) \subseteq \text{supp}\Big(\text{marg}_{T}$

Proof. *(See appendix)* ■

Lemma 3 *Under the* richness condition*, for each finite type* $\hat{t}_i \in \hat{T}_{i,\,\Theta^*}$*, for each* k*, for each* $s_i \in \mathcal{SSR}_i^k(\hat{t}_i)$ *such that the conjectures* $\mu^{s_i} \in \Phi(\hat{t}_i)$*:* $\{s_i\} = r_i(\mu^{s_i}|\hat{t}_i)$ *satisfy* $\bar{T}_{-i}(s_i) \subseteq \text{supp}\Big(\text{marg}_{T^*_{-i,\,\Theta^*}}\mu^{s_i}$ $\exists\, \tilde{t}_i \in \hat{T}_i$ *s.t.*

1. *For each* $k' \le k$, $\hat{\pi}^{k'}(\hat{t}_i) = \hat{\pi}^{k'}(\tilde{t}_i)$
2. $\mathcal{ISR}_i^{k+1}(\tilde{t}_i) = \{s_i\}$
3. $\tilde{t}_i \in \tilde{T}_i^{\tilde{t}_i}$ *for some finite belief closed set of types* $\tilde{T}^{\tilde{t}_i} = \times_{j\in N}\tilde{T}_j^{\tilde{t}_i}$ *such that* $\left|\mathcal{ISR}^{k+1}(t)\right| = 1$ *for each* $t \in \tilde{T}^{\tilde{t}_i}$.

Hence, for any such $s_i \in \mathcal{SSR}_i(\hat{t}_i)$ *there exists a sequence of finite types* $t_{i,m} \to \hat{t}_i$ *s.t.* $\mathcal{ISR}_i(t_{i,m}) = \{s_i\}$.
Proof. *See appendix.* ■

Given the lemmata above, the proof of proposition 4 is immediate:

Proof of Proposition 4: Take any $\hat{t} \in \hat{T}$ and any $s \in \mathcal{ISR}(\hat{t})$. For each i, from lemma 2 there exists a sequence $\{t_i^m\} \subseteq \hat{T}_{i,\,\Theta^*}$ of finite types s.t. $t_i^m \to \hat{t}_i$ and for each i, $s_i \in \mathcal{SSR}_i(t_i^m)$ for each m, for conjectures μ^{s_i} as in the thesis of lemma 2 and in the hypothesis of lemma 3. Then we can apply lemma 3 to the types t_i^m for each m: for $s_i \in \mathcal{SSR}_i(t_i^m)$, for each k, there exists a sequence $\left\{\tilde{t}_i^{m,k}\right\}_{k\in\mathbb{N}}$ s.t. $\tilde{t}_i^{m,k} \to t_i^m$ for $k \to \infty$ s.t. $\mathcal{ISR}_i\left(\tilde{t}_i^{m,k}\right) = \{s_i\}$. Because the universal type-space is metrizable, there exists a sequence $k_m \to \infty$ with $t_i^{m,k_m} \to \hat{t}_i$. Set $\hat{t}_i^m = t_i^{m,k_m}$, so that $\hat{t}^m \to \hat{t}$ as $m \to \infty$ and $\mathcal{ISR}(\hat{t}^m) = \{s\}$ for each m.■

5.2 Genericity of Uniqueness

In this section it is proved that uniqueness holds for an open and dense set of types in the universal type space.

The proof also uses the following known result:

Lemma 4 (Mertens and Zamir (1985)) *The set* $\hat{T}_{\Theta^*}$ *of finite types is dense in* $T^*_{\Theta^*}$*, i.e.*

$$T^*_{\Theta^*} = cl\left(\hat{T}_{\Theta^*}\right)$$

Proposition 5 *Under the* richness assumption*, the set*

$$\mathcal{U} = \{t \in T^* : |\mathcal{ISR}(t)| = 1\}$$

is open and dense in $T^*_{\Theta^*}$. *Moreover, the unique* $\mathcal{ISR}$ *outcome is locally constant, in the sense that* $\forall t \in U$ *such that* $ISR(t) = \{s\}$*, there exists an open neighborhood of types,* $\mathcal{N}_\delta(t)$*, such that* $ISR(t') = \{s\}$ *for all* $t' \in \mathcal{N}_\delta(t)$.

Proof: ($\mathcal{U}$ **is dense)** To show that $\mathcal{U}$ is dense, notice that by proposition 2, for any for any $\hat{t} \in \hat{T}$ there exists a sequence $\{\hat{t}^m\} \subseteq \hat{T}^m$ s.t. $\hat{t}^m \to \hat{t}$ and $\mathcal{ISR}(\hat{t}^m) = \{s\}$ for some $s \in \mathcal{ISR}(\hat{t})$. By definition, $\hat{t}^m \in U$ for each m. Hence, $\hat{t} \in cl(U)$, thus $\hat{T} \subseteq cl(\mathcal{U})$. But we know that $cl(\hat{T}) = T^*$, therefore $cl(\mathcal{U}) \supseteq cl(\hat{T}) = T^*$. Hence U is dense.

($\mathcal{U}$ **is open and** $\mathcal{ISR}$ **locally constant in** $\mathcal{U}$**)** Since (proposition 2) $\mathcal{ISR}$ is u.h.c., for each $t \in \mathcal{U}$, there exists a neighborhood $\mathcal{N}_\delta(t)$ s.t. for each $t' \in \mathcal{N}_\delta(t)$, $\mathcal{ISR}(t') \subseteq \mathcal{ISR}(t)$. Since ISR$(t) = \{s\}$ for some s, and ISR$(t') \neq \emptyset$, it follows trivially that $\mathcal{ISR}(t') = \{s\}$, hence $\mathcal{N}_\delta(t) \subseteq \mathcal{U}$. Therefore U is open. By the same token, we also have that $\mathcal{ISR}(t') = \{s\}$ for all $t' \in \mathcal{N}_\delta(t)$, i.e. the unique $\mathcal{ISR}$ outcome is locally constant.■

Corollary 2 *Generic uniqueness of* $\mathcal{ISR}$ *implies generic uniqueness of any equilibrium refinement. In particular, of any Perfect-Bayesian Equilibrium outcome.*

For each $s \in S$, let $\mathcal{U}^s = \left\{t \in \hat{T}_{\Theta^*} : \mathcal{ISR}(t) = \{s\}\right\}$. From proposition 5 we know that these sets are open. Let the boundary be $bd(\mathcal{U}^s) = cl(\mathcal{U}^s) \backslash \mathcal{U}^s$.

Corollary 3 *Under the richness condition, for each* $t \in \hat{T}_{\Theta^*}$:$|\mathcal{ISR}(t)| > 1$ *if and only if* $\exists$ $s, s' \in \mathcal{ISR}(t) : s \neq s'$ *such that* $t \in bd(\mathcal{U}^s) \cap bd(\mathcal{U}^{s'})$

Summing up, the results of this section conclude that $\mathcal{ISR}$ is a generically unique and locally constant solution concept, that yields multiple solutions at, and only at, the boundaries where the concept changes its prescribed behavior. The structure of $\mathcal{ISR}$ is therefore analogous to that proved by WY for *ICR*. Given the results from this and the previous section, the analogues of the remaining results in WY can be obtained in a straightforward manner for $\mathcal{ISR}$: in particular, proposition 4 also holds if one imposes the the common prior assumption.

6 Related Literature and Concluding Remarks

On NI- and PV-settings. In NI-settings, *all* CK-assumptions are relaxed. In particular, the assumption that players know their own preferences. Under an equivalent richness condition, independent work by Chen (2008) studied the structure of *ICR* for dynamic NI-environments, extending WY's results. Together with the upper hemicontinuity of $\mathcal{ISR}$, Chen's results imply that the two solution concepts coincide on the universal model in these settings.[31] Outside of the realm of NI-environments, $\mathcal{ISR}$ generally refines *ICR* (imposing *sequential rationality* restrictions). Hence, $\mathcal{ISR}$ is the solution concept that extends WY's results to dynamic environments in all settings. For the special NI-case, it coincides with *ICR*: This means that when *all* CK-assumptions are relaxed, thereby included that players know their own payoffs, *sequential rationality* loses its restrictive power. The intuition is simple: In NI-environments players don't know their own payoffs, they merely have beliefs about them. Once an unexpected information set is reached, Bayes' rule does not restrict players' beliefs, which can be set arbitrarily. Under the richness conditions, there are essentially no restrictions on players' beliefs about their own preferences, so that any behavior can be justified. Hence, the only restrictions that retain their bite are those imposed by (normal form) rationality alone. This is the same intuition behind the non-model invariance of $\mathcal{ISR}$ in NI-settings (example 5).

To the extent that the interest in studying extensive form games comes from the notion of sequential rationality, PV-settings, in which the assumption that *players know their own payoffs* is maintained, seem to be the most significant for dynamic environments. As shown by proposition 3, $\mathcal{ISR}$ is *model invariant* in these settings. In appendix D it is also shown that, in PV-environments, $\mathcal{ISR}$ is (generically) equivalent to Dekel and Fudenberg's (1990) $S^{\infty}W$ procedure applied to the interim normal form.[32]

On the Related Solution concepts. $\mathcal{ISR}$ generalizes to games with incomplete and imperfect information the solution concept developed by Ben-Porath (1997) to characterize the behavioral implications of *(initial) common certainty of rationality.* Ben-Porath (1997) also proves that for games with *complete* and *perfect information* with payoffs in generic position, his solution concept is equivalent to Dekel and Fudenberg's (1990) $S^{\infty}W$-procedure. The results in appendices C and D thus generalize Ben -Porath's (1997) to incomplete and imperfect information.

Battigalli and Siniscalchi (2007) notion of *weak* Δ-*rationalizability* is also closely related to $\mathcal{ISR}$: their solution concept is not defined for Bayesian games, but for games with payoff uncertainty. For Bayesian games with information types $\mathcal{ISR}$ can be shown to be equivalent to *weak* Δ-*rationalizability*, where the Δ-restrictions on first order beliefs are those derived from the type space.[33]

[31] I thank Eddie Dekel for this observation.

[32] The $S^{\infty}W$-procedure consists of one round of deletion of weakly dominated strategies followed by iterated deletion of strongly dominated strategies.

[33] An analogous result in Battigalli et al. (2008) relates Δ-*rationalizability* and *interim correlated rationalizability* in games with information types.

On the Robustness(-es). A closely related paper, that addresses the question of robustness with a very similar spirit, is Dekel and Fudenberg (1990): In that paper, they characterize $S^{\infty}W$ as being the solution concept that is "robust" with respect to the possibility that players entertain small doubts about their opponents' payoff functions. The robustness result for $\mathcal{ISR}$ is in the same spirit. Dekel and Fudenberg (1990) maintain the assumption that players *know* their own payoffs: this corresponds to the PV-case, for which it is shown that indeed $\mathcal{ISR}$ coincides with $S^{\infty}W$ applied to the interim normal form (appendix D).

Type space invariance, in the context of normal form games, has been addressed by the literature: Ely and Peski (2006) and Dekel et al. (2007) pointed out the type space-dependence of *Interim Independent Rationalizability (IIR).* Based on this observation, Ely and Peski (2006) showed that the relevant measurability condition for *IIR* is not in terms of the Θ-*hierarchies*, but in terms of $\Delta(\Theta)$-hierarchies. Dekel et al. (2007) instead introduced the concept of *Interim Correlated Rationalizability (ICR)* and showed that it is *type space invariant.*[34] To the best of my knowledge, the problem of *model invariance* was not addressed by the literature.[35]

On the Impact of higher order beliefs on multiplicity. The result of proposition 5 and the epistemic characterization of $\mathcal{ISR}$ imply that, generically, *initial common certainty of sequential rationality* is sufficient to achieve coordination of expectations. Generically, multiplicity is driven by the CK-assumptions of our models. As discussed in the introduction, this parallels what we knew from WY and the literature on static global games. A growing literature is exploring to what extent the main insights from the theory of global games can be extended to dynamic environments. These contributions are mainly from an applied perspective, and do not pursue a systematic analysis of these problems. Consequently, the conclusions are variegated: for instance, Chamley (1999) and Frankel and Pauzner (2000) obtain the familiar uniqueness result in different setups, under different sets of assumptions. On the other hand, few recent papers that explore the role of learning in *dynamic global games* seem to question the general validity of these results: Work by Angeletos et al. (2006, 2007) and Angeletos and Werning (2006) apply the global games' information structure to dynamic environments, and obtain non-uniqueness results that contrast with the familiar

[34] *Interim Independent Rationalizability* is Pearce's (1984) rationalizability in the interim normal form. In turn, it coincides with Pearce's (1984) *Extensive Form Rationalizability (EFR)* applied to the extensive form of a static Bayesian game. *Interim Correlated Rationalizability* can be seen as a version of Pearce's (1984) EFR that allows for more correlation in beliefs. Battigalli et al. (2008) explore the epistemic foundations and the relations between Δ-Rationalizability, ICR, IIR, and other solution concepts for Bayesian games in normal forms.

[35] The "extension" and "contraction" properties considered by Friedenberg and Meyer (2007) are similar in spirit.

ones in static settings. The origin of such multiplicity lies in a tension between the global games' information structure and the dynamic structure: by relaxing CK-assumptions, the former favors uniqueness; in dynamic games, some information endogenously becomes CK (e.g. public histories), thus mitigating the impact of the information structure.

Many aspects of the papers mentioned above make the comparison with the present framework difficult, and a careful analysis and understanding of the relations between the two approaches is an interesting open question for future research.[36] One important difference underlying the contrasting results in term of uniqueness is certainly the fact that, in those papers, not *all* CK assumptions are relaxed. For instance, in the paper by Angeletos et al. (2007), the assumption of CK that the stage game does not depend on the previous history is maintained throughout. As mentioned in the introduction, important work for future research would be to investigate the robustness and uniqueness questions when only *some* CK assumptions are relaxed.[37] This line of research may help shed some light on the relations between the present work and the literature on dynamic global games.

7 Appendix A: proofs from section 4

Proof of proposition 1: The proof is by induction. The initial step is trivial, for $\mathcal{ISR}^0$ is vacuously u.h.c.. For the inductive step, suppose that $\mathcal{ISR}^{k-1}$ is u.h.c., and let $\{t_i^m\}$ be s.t. $t_i^m \to t_i$, $\{s_i^m\}$ s.t. $s_i^m \to \hat{s}_i$ and $s_i^m \in \mathcal{ISR}_i^k(t_i^m)$ for all m. Then, for each m, there exists $\mu^{i,m} \in \Delta^{\mathcal{H}}\left(\Theta_0 \times T^*_{-i,\,\Theta} \times S_{-i}\right)$ s.t.

$$\begin{aligned}&(1)\ s_i^m \in r_i\left(\mu^{i,m}|t_i^m\right)\\&(2)\ \tau^*_{t_i^m} = marg_{\Theta_0 \times T^*_{-i,\,\Theta}}\mu^{i,m}\left(\cdot|\phi\right)\\&(3)\ supp(\mu^{i,m}\left(\cdot|\phi\right)) \subseteq \Theta_0 \times \mathcal{ISR}^{k-1}_{-i}\end{aligned}$$

We want to show that $\hat{s}_i \in \mathcal{ISR}_i^k(t_i)$, i.e. that $\exists \hat{\mu}^i \in \Delta^{\mathcal{H}}\left(T^*_{-i} \times S_{-i}\right)$ s.t.

$$\begin{aligned}&(1')\ \hat{s}_i \in r_i\left(\hat{\mu}^i|\hat{t}_i\right)\\&(2')\ \tau^*_{t_i} = marg_{\Theta_0 \times T^*_{-i}}\hat{\mu}^i\left(\cdot|\phi\right)\\&(3')\ supp(\hat{\mu}^i\left(\cdot|\phi\right)) \subseteq \Theta_0 \times \mathcal{ISR}^{k-1}_{-i}\end{aligned}$$

Consider the sequence $\{\mu^{i,m}\} \subseteq \Delta^{\mathcal{H}}\left(\Theta_0 \times T^*_{-i,\,\Theta} \times S_{-i}\right)$. Since $\{\mu^{i,m}\}$ is in a compact, there exists a subsequential limit $\hat{\mu}^i$. By continuity of τ^*, $\hat{\mu}^i$ satisfies condition (2'), since (2) holds for each $\mu^{i,m}$ and $t_i^m \to t_i$. Furthermore, since the best-response correspondence $r_i(\cdot)$ is u.h.c. and $s_i^m \in r_i\left(\mu^{i,m}|t_i^m\right)$ for each m, also condition (1') is satisfied. Given that $\mu^{i,m} \to \hat{\mu}^i$, the upper hemicontinuity of $\mathcal{ISR}^{k-1}_{-i}$ (from the inductive hypothesis) suffices for (3').■

[36] An important difference is that the game in Angeletos et al. (2007) has infinite horizon. Extending the analysis of this paper to multistage games with infinite horizon is an important question, left to future research.

[37] In Penta (2008) these questions are investigated in the context of static games.

Proof of proposition 3: The proof is by induction: Let $t_i^* \in T_{i,\Theta^*}^*$. Clearly, if $\boldsymbol{\theta}_i(t_i) = \hat{\pi}_i^0(t_i^*)$, $\mathcal{ISR}_i^{\mathcal{T}_\Theta,1}(t_i) = \mathcal{ISR}_i^{\mathcal{T}_{\Theta^*}^*,1}(t_i^*)$. As inductive hypothesis, assume that $(\hat{\pi}_i^n(t_i))_{n=0}^{k-1} = (\hat{\pi}_i^n(t_i^*))_{n=0}^{k-1}$ implies that $\mathcal{ISR}_i^{\mathcal{T}_\Theta,k}(t_i) = \mathcal{ISR}_i^{\mathcal{T}_{\Theta^*}^*,k}(t_i^*)$, and suppose that $(\hat{\pi}_i^n(t_i))_{n=0}^{k} = (\hat{\pi}_i^n(t_i^*))_{n=0}^{k}$. It will be shown that $(\hat{\pi}_i^n(t_i))_{n=0}^{k} = (\hat{\pi}_i^n(t_i^*))_{n=0}^{k}$ implies that $\mathcal{ISR}_i^{\mathcal{T}_\Theta,k+1}(t_i) = \mathcal{ISR}_i^{\mathcal{T}_{\Theta^*}^*,k+1}(t_i^*)$. Under the inductive hypothesis, $s_{-i} \in \mathcal{ISR}_i^{\mathcal{T}_\Theta,k}(t_{-i})$ for some $t_{-i} \in supp\ \tau_i(t_i)$ if and only if $s_{-i} \in \mathcal{ISR}_{-i}^{\mathcal{T}_{\Theta^*}^*,k}(t_{-i}^*)$ for some $t_{-i}^* \in supp\ \tau_i(t_i^*)$. In PV-environments only the conjectures about S_{-i} are payoff relevant for player i (θ_{-i}'s don't affect i's payoffs, and Θ_0 is a singleton.). Thus, under the inductive hypothesis, $\exists \mu^i \in \Phi_i(t_i)$ s.t. $supp(marg_{T_{-i,\Theta}\times S_{-i}}\mu^i) \subseteq \mathcal{ISR}_{-i}^{\mathcal{T}_\Theta,k+1}$ and $s_i \in r_i(\mu^i|t_i)$ if and only if $\exists \hat{\mu}^i \in \Phi_i(t_i)$ s.t. $supp(marg_{T_{-i,\Theta^*}^*\times S_{-i}}\hat{\mu}^i) \subseteq \mathcal{ISR}_{-i}^{\mathcal{T}_{\Theta^*}^*,k+1}$ and $s_i \in r_i(\hat{\mu}^i|t_i^*)$. (Remember the only restrictions on the conjectures over S_{-i} imposed by $\mathcal{ISR}$ are at the begininng of the game). Hence $\mathcal{ISR}_i^{\mathcal{T}_\Theta,k+1}(t_i) = \mathcal{ISR}_i^{\mathcal{T}_{\Theta^*}^*,k+1}(t_i^*)$.■

8 Appendix B: proofs from section 5

proof of lemma 2: **(Part I:)** Fix $t_i \in \hat{T}_{i,\Theta^*}$. For each $k \neq i$, let Θ_k' be the finite set of payoff states that receive positive probability by $t_i \in T_j$, $j \neq k$, and let $\Theta_i = \Theta_i' \cup \bar{\Theta}_i$. ($\Theta_i$ is finite because t_i is a finite type and $\bar{\Theta}_i$ is finite). $\forall s_i \in \mathcal{ISR}(t_i)$, $\exists \mu^{s_i} \in \Phi_i(t_i)$ s.t. (1) $s_i \in r_i(\mu^{s_i}|t_i)$ and (2) $supp(\mu^{s_i}(\cdot|\phi)) \subseteq \mathcal{ISR}_{-i}$. Given a probability space $(\Omega, \mathcal{B})$ and a set $A \in \mathcal{B}$, denote by $\upsilon_{[A]}$ the uniform probability distribution concentrated on A. For each $\varepsilon \in [0,1]$, consider the set of types profiles $\times_{i\in N} T_i^\varepsilon \subseteq T_{\Theta^*}^*$ s.t. each T_i^ε consists of all the types $\bar{t}_i \in \bar{T}_i$ and of types $\bar{\iota}_i(t_i, s_i, \varepsilon)$ s.t.:

$$\boldsymbol{\theta}_i(\bar{\iota}_i(t_i,s_i,\varepsilon)) = \varepsilon\theta_i^s + (1-\varepsilon)\,\boldsymbol{\theta}_i(t_i)$$

and

$$\tau_i^\varepsilon(\bar{\iota}_i(t_i,s_i,\varepsilon)) = \varepsilon\upsilon_{\left[\{\theta_0^s\}\times\bar{T}_{-i}(s_i)\right]} + (1-\varepsilon)\left[\mu^{s_i}(\cdot|\phi)\circ\hat{\iota}_{-i,\varepsilon}^{-1}\right]$$

where $\bar{T}_{-i} \subseteq T_{-i}^\varepsilon$ is the subset of *dominance-types* profiles defined above, and

$$\hat{\iota}_{-i,\varepsilon} : \Theta_0 \times T_{-i} \times S_{-i} \to \Theta_0 \times T_{-i}^\varepsilon \text{ is s.t.}$$
$$\hat{\iota}_{-i,\varepsilon}(\theta_0, s_{-i}, t_{-i},) = (\theta_0, \bar{\iota}_{-i}(t_{-i}, s_{-i}, \varepsilon))$$

By construction, with probability ε, type $\bar{\iota}_i(t_i,s_i,\varepsilon)$ is certain that s_i is conditionally dominant, and puts positive probability to all of the opponents' *dominance types* in $\bar{T}_{-i}$. Define $\gamma : \Theta_0 \times T_{-i}^\varepsilon \to \Theta_0 \times T_{-i}^\varepsilon \times S_{-i}$ s.t.:

$$\forall \bar{\iota}_{-i}(t_{-i},s_{-i},\varepsilon) \in T_{-i}^\varepsilon :$$
$$\gamma(\theta_0, \bar{\iota}_{-i}(t_{-i},s_{-i},\varepsilon)) = (\theta_0, \bar{\iota}_{-i}(t_{-i},s_{-i},\varepsilon), s_{-i})$$
$$\text{and for every } \bar{t}_{-i}^s \in \bar{T}_{-i} \subseteq T_{-i}^\varepsilon,$$
$$\gamma\left(\theta_0, \bar{t}_{-i}^s\right) = \left(\theta_0, \bar{t}_{-i}^s, s_{-i}\right)$$

Consider the conjectures $\hat{\mu}^i \in \Delta^{\mathcal{H}}\left(\Theta_0 \times T^{\varepsilon}_{-i} \times S_{-i}\right)$ defined by:

$$\hat{\mu}^i(\cdot|\phi) = \left(\tau^{\varepsilon}_i(\bar{\iota}_i(t_i, s_i, \varepsilon)) \circ \gamma^{-1}\right) \in \Delta\left(\Theta_0 \times T^{\varepsilon}_{-i} \times S_{-i}\right)$$

For any $\varepsilon > 0$, the conjectures $\hat{\mu}^i$ are such that $\bar{T}_{-i} \subseteq supp\left(marg_{T^{\varepsilon}_{-i}}\hat{\mu}^i(\cdot|\phi)\right)$. From the definition of γ, it follows that $supp\left(marg_{S_{-i}}\hat{\mu}^i(\cdot|\phi)\right) = S_{-i}$, so that the entire CPS $\left(\hat{\mu}^i(\cdot|h)\right)_{h\in\mathcal{H}}$ can be obtained via Bayes' Rule. This also implies that $\hat{\mu}^i$ satisfies condition (3) in the definition of $\mathcal{SSR}$. Furthermore, by construction, $\hat{\mu}^i \in \Phi_i(\bar{\iota}_i(t_i, s_i, \varepsilon))$, and $\forall\varepsilon > 0$, $\forall h \in \mathcal{H}$, $\exists\eta^{\varepsilon,h} \in (0,1)$ s.t. $\eta^{\varepsilon,h} \to 0$ as $\varepsilon \to 0$ and

$$\begin{aligned} marg_{\Theta_0\times T^*_{-i}\times S_{-i}}\hat{\mu}^i(\cdot|h) = &\, \eta^{\varepsilon,h}\upsilon_{\left[\{\theta^s_0\}\times\bar{T}^h_{-i}(s_i)\times S_{-i}(h)\right]} \\ &+ \left(1-\eta^{\varepsilon.h}\right) marg_{\Theta_0\times S_{-i}}\mu^{s_i}(\cdot|h) \end{aligned}$$

where $\bar{T}^h_{-i}(s_i) = \left\{\bar{t}^{(s_i,s_{-i})}_{-i} : s_{-i} \in S_{-i}(h)\right\}$. Hence, the conditional conjectures $\hat{\mu}^i(\cdot|h)$ of type $\bar{\iota}_{-i}(t_{-i}, s_{-i}, \varepsilon)$ are a mixture: with probability $\left(1-\eta^{\varepsilon,h}\right)$ they agree with $\mu^i(\cdot|\phi)$, which made s_i sequential best response; with probability $\eta^{\varepsilon,h}$ they are concentrated on payoff states $\{\theta^{s_i}_0\} \times \left\{\boldsymbol{\theta}_{-i}(t_{-i}) : t_{-i} \in \bar{T}^h_{-i}(s_i)\right\}$, which together with the fact that $\boldsymbol{\theta}_i(\bar{\iota}_i(t_i, s_i, \varepsilon)) = \varepsilon\theta^{s_i}_i + (1-\varepsilon)\boldsymbol{\theta}_i(t_i)$ breaks the all ties in favor of s_i so that $r_i\left(\hat{\mu}^i|\bar{\iota}_i(t_i, s_i, \varepsilon)\right) = \{s_i\}$. Thus, $s_i \in \mathcal{SSR}_i(\bar{\iota}_i(t_i, s_i, \varepsilon))$, so that (ii) and (iii) in the lemma are satisfied for all $\varepsilon > 0$.

The remainder of the proof guarantees that also part (i) in the lemma holds, and it is identical to WY's counterpart.

(**Part II:**) it will be shown that $\hat{\pi}^*_i(\bar{\iota}_i(t_i, s_i, \varepsilon)) \to \hat{\pi}^*_i(t_i)$ as $\varepsilon \to 0$. By construction, $\tau_i(\bar{\iota}_i(t_i, s_i, \varepsilon))$ are continuous in ε, hence $\hat{\pi}^*_i(\bar{\iota}_i(t_i, s_i, \varepsilon)) \to \hat{\pi}^*_i(\bar{\iota}_i(t_i, s_i, 0))$ as $\varepsilon \to 0$ (Brandenburger and Dekel, 1993). It suffices to show that $\hat{\pi}^*_i(\bar{\iota}_i(t_i, s_i, 0)) = \hat{\pi}^*_i(t_i)$ for each t_i and i. This is proved by induction. The payoff types and the first order beliefs are the same. For the inductive step, assume that $\left(\hat{\pi}^l_i(\bar{\iota}_i(t_i, s_i, 0))\right)^{k-1}_{l=0} = \left(\hat{\pi}^l_i(t_i)\right)^{k-1}_{l=0}$. We will show that $\hat{\pi}^k_i(\bar{\iota}_i(t_i, s_i.0)) = \hat{\pi}^k_i(t_i)$. Define $D^{k-1}_{-i} = \left\{\left(\hat{\pi}^l_{-i}(t_{-i})\right)^{k-1}_{l=0} : t_{-i} \in T_{-i}\right\}$. Under the inductive hypothesis, it can be shown (see WY) that

$$marg_{\Theta_0\times D^{k-1}_{-i}}\left[\mu^i(\cdot|\phi) \circ \hat{\iota}^{-1}_{-i}\right] = marg_{\Theta_0\times D^{k-1}_{-i}}\mu^i(\cdot|\phi) \qquad (\diamondsuit)$$

Therefore:

$$\begin{aligned} \hat{\pi}^k_i(\bar{\iota}_i(t_i, s_i, 0)) &= \upsilon_{\left[\hat{\pi}^{k-1}_{-i}(\bar{\iota}_i(t_i,s_i,0))\right]} \times marg_{\Theta_0\times D^{k-1}_{-i}}\left[\mu^{s_i}(\cdot|\phi) \circ \hat{\iota}^{-1}_{-i}\right] \\ &= \upsilon_{\left[\hat{\pi}^{k-1}_{-i}(\bar{\iota}_i(t_i,s_i,0))\right]} \times marg_{\Theta_0\times D^{k-1}_{-i}}\mu^{s_i}(\cdot|\phi) \\ &= \upsilon_{\left[\hat{\pi}^{k-1}_{-i}(t_{-i})\right]} \times marg_{\Theta_0\times D^{k-1}_{-i}}\mu^{s_i}(\cdot|\phi) \\ &= \upsilon_{\left[\hat{\pi}^{k-1}_{-i}(t_{-i})\right]} \times marg_{\Theta_0\times D^{k-1}_{-i}}\tau_{t_i} \\ &= \hat{\pi}^k_i(t_i) \end{aligned}$$

where the first equality is the definition of k-th level belief; the second from ($\diamondsuit$); the third from the inductive hypothesis; the fourth from the fact that $\mu^{s_i} \in \Phi(t_i)$; the last one again by definition.■

Proof of Lemma 3: The proof is by induction: For $k = 0$, let $\tilde{t}_i$ be s.t. $\boldsymbol{\theta}_i(\tilde{t}_i) = \boldsymbol{\theta}_i(\hat{t}_i)$ and $\tau_i(\tilde{t}_i) = \upsilon_{[\{\theta^s\}\times \bar{T}_{-i}(s_i)]}$. Clearly, $\mathcal{ISR}_i^1(\tilde{t}_i) = \{s_i\}$ and condition (1) is satisfied. Fix $k > 0$, write each $t_{-i} = (\lambda, \kappa)$ where $\lambda = \left\{\hat{\pi}_{-i}^{k'}(t_{-i})\right\}_{k'=1}^{k-1}$ and $\kappa = \left\{\hat{\pi}_{-i}^{k'}(t_{-i})\right\}_{k'=k}^{\infty}$. Let $L_{-i}^{k-1} = \left\{\lambda : \exists\kappa \text{ s.t. } (\lambda,\kappa) \in T_{-i}^*\right\}$. As inductive hypothesis, assume that: for each finite $t_{-i} = (\lambda,\kappa)$ and $s_{-i} \in \mathcal{SSR}_{-i}^{k-1}(t_{-i})$ s.t. $\bar{T}_{-i}(s_i) \subseteq supp\left(marg_{\hat{T}_{-i}}\mu^{s_i}(\cdot|\phi)\right)$, there exists finite $t_{-i}^{s_{-i}} = \left(\lambda, \kappa^{s_{-i},\lambda}\right)$ s.t. $\mathcal{ISR}_i^k(t_i^{s_i}) = \{s_i\}$. Take any $s_i \in \mathcal{SSR}_i^k(\hat{t}_i)$ s.t. $\bar{T}_{-i}(s_i) \subseteq supp\left(marg_{\hat{T}_{-i}}\mu^{s_i}(\cdot|\phi)\right)$: we will construct a type $\tilde{t}_i$ s.t. for each $k' \le k$, $\hat{\pi}_i^{k'}(\hat{t}_i) = \hat{\pi}_i^{k'}(\tilde{t}_i)$, $\mathcal{I}\text{SR}_i^{k+1}(\tilde{t}_i) = \{s_i\}$: By definition, if $s_i \in \mathcal{SSR}_i^k(\hat{t}_i)$, $\exists\mu^{s_i} \in \Delta^{\mathcal{H}}\left(\Theta_0^* \times T_{-i}^* \times S_{-i}\right)$ s.t.

$$
\begin{aligned}
&(1)\ \tau_i(\hat{t}_i) = marg_{\Theta\times T_{-i}}\mu^{s_i}(\cdot|\phi)\\
&(2).\ supp\left(\mu^{s_i}(\cdot|\phi)\right) \subseteq \Theta_0^* \times \mathcal{SSR}_{-i}^{k-1}\\
&(3)\ \{s_i\} = r_i\left(\mu^{s_i}|\hat{t}_i\right)
\end{aligned}
$$

Using the inductive hypothesis, define the mapping

$$
\begin{aligned}
&\varphi : \bigcup_{h\in\mathcal{H}}\left[supp\left(marg_{\Theta_0^*\times L_{-i}^{k-1}\times S_{-i}}\mu^{s_i}(\cdot|h)\right)\right] \to \Theta_0^* \times T_{-i}^*\\
&\text{such that: } \varphi(\theta,\lambda,s_{-i}) = \left(\theta, \left(\lambda, \kappa^{s_{-i},\lambda}\right)\right)
\end{aligned}
$$

Define type $\tilde{t}_i$ as

$$
\begin{aligned}
\tau_i(\tilde{t}_i) &= marg_{\Theta_0^*\times L_{-i}^{k-1}\times S_{-i}}\mu^{s_i}(\cdot|\phi)\circ\varphi^{-1}\\
&= \mu^{s_i}(\cdot|\phi)\circ proj^{-1}_{\Theta_0^*\times L_{-i}^{k-1}\times S_{-i}}\circ\varphi^{-1}
\end{aligned}
$$

By construction (for the inductive hypothesis), the first k orders of beliefs are the same for t_i and $\tilde{t}_i$ (which is point 1 in the lemma)

$$
\begin{aligned}
\tilde{\pi}_i^k(\tilde{t}_i) &= marg_{\Theta^*\times L_{-i}^{k-1}}\tau_{\tilde{t}_i}\\
&= \mu^{s_i}(\cdot|\phi)\circ proj^{-1}_{\Theta^*\times L_{-i}^{k-1}\times S_{-i}}\circ\varphi^{-1}\circ proj^{-1}_{\Theta^*\times L_{-i}^{k-1}}\\
&= \mu^{s_i}(\cdot|\phi)\circ proj^{-1}_{\Theta^*\times L_{-i}^{k-1}}\\
&= \left(\mu^{s_i}(\cdot|\phi)\circ proj^{-1}_{\Theta^*\times T_{-i}^*}\right)\circ proj^{-1}_{\Theta^*\times L_{-i}^{k-1}}\\
&= marg_{\Theta^*\times L_{-i}^{k-1}}\tau_{t_i}\\
&= \tilde{\pi}_i^k(t_i)
\end{aligned}
$$

Where the first equality is by definition, the second is from construction of $\tau_i\left(\tilde{t}_i\right)$ above; the third is from the definition of φ, for which

$$proj^{-1}_{\Theta_0^*\times L_{-i}^{k-1}\times S_{-i}}\circ\varphi^{-1}\circ proj^{-1}_{\Theta_0^*\times L_{-i}^{k-1}}=proj^{-1}_{\Theta_0^*\times L_{-i}^{k-1}}$$

The fourth and fifth are simply notational, and the last one by definition. We need to show that $\mathcal{ISR}_i^{k+1}\left(\tilde{t}_i\right)=\{s_i\}$. To this end, notice that each $(\theta_0,t_{-i})\in supp\left(\tau_i\left(\tilde{t}_i\right)\right)$ is of the form $(\theta_0,t_{-i})=\left(\theta_0,\left(\lambda,\kappa^{s_{-i},\lambda}\right)\right)$, and it is s.t. $\mathcal{ISR}_{-i}^k\left(\left(\lambda,\kappa^{s_{-i},\lambda}\right)\right)=\{s_{-i}\}$. Hence, the array of conditional conjectures (i.e. CPS) that are consistent with $\tilde{t}_i$ and with the restrictions $\mathcal{ISR}_{-i}^k$ are $\tilde{\mu}^i$ s.t.:

$$\tau_i\left(\hat{t}_i\right)=marg_{\Theta_0^*\times T_{-i}^*}\tilde{\mu}^i\left(\cdot|\phi\right)$$

and

$$\tilde{\mu}^i\left(\Theta_0^*\times\left\{(t_{-i},s_{-i}):\mathcal{ISR}_{-i}^k\left(t_{-i}\right)=\{s_{-i}\}\right\}|\phi\right)=1$$

hence, the conditional conjectures are uniquely determined for all $h\in\mathcal{H}\left(s_{-i}\right)$ for some $s_{-i}:\{s_{-i}\}=\mathcal{ISR}_{-i}^k\left(t_{-i}\right)$ and $t_{-i}\in supp\left(marg_{T_{-i}}\tau_i\left(\hat{t}_i\right)\right)$. But since, from hypothesis, $\bar{T}_{-i}\left(s_i\right)\subseteq supp\left(marg_{T_{-i}}\mu^{s_i}\right)$, and that by definition of $\bar{T}_{-i}\left(s_i\right)$ we trivially have that $\bigcup_{t_{-i}\in\bar{T}_{-i}(s_i)}\mathcal{ISR}_{-i}^k\left(t_{-i}\right)=S_{-i}$, the conditional conjectures are uniquely determined for all $h\in\mathcal{H}$. These conjectures are given by $\tilde{\mu}^i\left(\cdot|\phi\right)=\tau_i\left(\hat{t}_i\right)\circ\phi^{-1}$, with ϕ defined as

$$\phi\left(\theta_0,\left(\lambda,\kappa^{s_{-i},\lambda}\right)\right)=\left(\theta_0,\left(\lambda,\kappa^{s_{-i},\lambda}\right),s_{-i}\right)$$

Furthermore, for each h:

$$marg_{\Theta^*\times S_{-i}}\tilde{\mu}^i\left(\cdot|h\right)=marg_{\Theta^*\times S_{-i}}\mu^{s_i}\left(\cdot|h\right)$$

To see this, given the observation that $\mu^{s_i}\left(\cdot|\phi\right)$ is completely mixed (i.e. all histories are reachable with positive probability), it suffices to show that $marg_{\Theta^*\times S_{-i}}\tilde{\mu}^i\left(\cdot|\phi\right)=marg_{\Theta^*\times S_{-i}}\mu^{s_i}\left(\cdot|\phi\right)$. But this is immediate, given that from the definition of ϕ and φ, we have:

$$proj^{-1}_{\Theta_0^*\times L_{-i}^{k-1}\times S_{-i}}\circ\phi\circ\varphi=I.$$

(I is the identity map). Hence, $\tilde{\mu}^i$ is uniquely determined for all h, and it is equal to μ^{s_i}, to which s_i is the unique best response. Hence $\mathcal{ISR}_i^{k+1}\left(\tilde{t}_i\right)=\{s_i\}$.

The proof of statement (3) in the lemma is identical to WY's: Define

$$\tilde{T}_i^{\tilde{t}_i}=\left\{\tilde{t}_i\right\}\cup\left(\bigcup_{\left(\theta,t_{-i}^{s_{-i}}\right)\in supp\left(\tau_i\left(\hat{t}_i\right)\right)}T_i^{t_{-i}^{s_{-i}}}\right),$$

$$\tilde{T}_j^{\tilde{t}_i}=\bigcup_{\left(\theta,t_{-i}^{s_{-i}}\right)\in supp\left(\tau_i\left(\hat{t}_i\right)\right)}T_{-i}^{t_{-i}^{s_{-i}}}\qquad(\text{for } j\neq i)$$

■

9 Appendix C: Epistemic Characterization.

For any standard Borel space X, write $\Delta(X)$ for the set of all probability measures on X, endowed with the topology of weak convergence and the corresponding Borel σ-algebra. Given any other Borel space X', every measurable mapping $f : X \to X'$ has an associated *pushforward mapping* $\hat{f} : \Delta(X) \to \Delta(X')$ s.t. $\hat{f}(\mu)[E'] = \mu[f^{-1}(E')]$ for each measurable $E' \subseteq X'$.

Fix a space of primitive uncertainty of the form

$$Y = \Theta_0 \times Y_1 \times ... \times Y_n$$

Let Y*-based type spaces* (tuples $\mathcal{T}_Y = \langle Y, (T_{i,Y}, \mathbf{y}_i, \tau_i)_{i\in N}\rangle$, Y*-based hierarchies* (elements of $\times_{k\geq 1}\Delta\left(Z^{k-1}_{Y,i}\right))$) and the Y*-based universal type space* ($\mathcal{T}^*_Y$) be defined as in section 3.[38] Similarly, let $\hat{\pi}^*_{Y,i} : T_i \to T^*_{Y,1}$ denote the *canonical belief morphism* from $\mathcal{T}_Y$ to $\mathcal{T}^*_Y$. We'll be concerned with two cases: $Y_i = \Theta_i$ for all $i \in I$, or $Y_i = \Theta_i \times S_i$ for all $i \in I$. In these two cases we have $Y = \Theta$ or $Y \approx \Theta \times S$, respectively.[39]

Type Space Marginalization. Each *explicit* $(\Theta \times S)$*-based type* naturally induces a Θ-*based type*, forgetting all information other than that about Θ, thus deriving Θ-based beliefs via recursive marginalization: let $m^0_i : Z^0_{\Theta\times S,i} \to Z^0_{\Theta,i}$ be the natural projection and $\widehat{m}^0_i : \Delta\left(Z^0_{\Theta\times S,i}\right) \to \Delta\left(Z^0_{\Theta,i}\right)$ the associated pushforward mapping; recursively, define $m^k_i : Z^k_{\Theta\times S,i} \to Z^k_{\Theta,i}$ as $\left(z^{k-1}_{\Theta\times S,i}, \pi^k_{\Theta\times S,-i}\right) \mapsto \left(m^{k-1}_i\left(z^{k-1}_{\Theta\times S,i}\right), \widehat{m}^{k-1}_{-i}\left(\pi^k_{\Theta\times S,-i}\right)\right)$ and let $\widehat{m}^k_i : \Delta\left(Z^k_{\Theta\times S,i}\right) \to \Delta\left(Z^k_{\Theta,i}\right)$ be the associated pushforward mapping; finally, define $\mathbf{m}^k_i : T^*_{\Theta\times S,i} \to H^k_{\Theta,i}$ and $\mathbf{m}^*_i : T^*_{\Theta\times S,i} \to T^*_{\Theta,i}$ respectively as

$$\left(\theta_i, s_i, \pi^1_{\Theta\times S,i}, \pi^2_{\Theta\times S,i}, \ldots\right) \mapsto \left(\theta_i, \widehat{m}^0_i(\pi^1_{\Theta\times S,i}), \ldots, \widehat{m}^{k-1}_i(\pi^k_{\Theta\times S,i})\right)$$
$$\text{and } \left(\theta_i, s_i, \pi^1_{\Theta\times S,i}, \pi^2_{\Theta\times S,i}, \ldots\right) \mapsto \left(\theta_i, \widehat{m}^0_i(\pi^1_{\Theta\times S,i}), \widehat{m}^1{}_i(\pi^2_{\Theta\times S,i}), \ldots\right).$$

Epistemic Model. Set $X^0_i = \Theta_{-i} \times S_{-i}$ and for all $n \geq 1$, let $X^n_i = X^{n-1}_i \times \Delta^{\mathcal{H}}\left(X^{n-1}_{-i}\right)$. Elements of $\Delta^{\mathcal{H}}\left(X^{k-1}_i\right)$ are i's *k-order CPSs.* Elements of $\Psi_i = \prod^{\infty}_{k=0}\Delta^{\mathcal{H}}_i\left(X^k\right)$ are *hierarchies of CPSs.* For each $m \geq 0$, $n \geq 1$ and $A \subseteq X^m_i$, let $\mathcal{C}^n(A)$ denote the subset of X^{n+m}_i corresponding to A: for $A \subseteq X^m$, $\mathcal{C}^n(A) = A \times \prod^{n+m-1}_{k=m}\left[\Delta^{\mathcal{H}}\left(X^k_{-i}\right)\right] \subseteq X^{m+n}_i$. An infinite hierarchy of CPSs $\left(\psi^1_i, \psi^2_i, ..., \psi^n_i, ...\right) \in \Psi_i$ is *coherent* if, for all $h \in \mathcal{H}$, and for all $n = 1, 2, ...$: $marg_{X^{n-1}_i}\psi^{n+1}_i(\cdot|\mathcal{C}^n([h]_i)) = \psi^n_i(\cdot|\mathcal{C}^{n-1}([h]_i))$, where $[h]_i \equiv \Theta_{-i} \times S_{-i}(h)$. Let Ψ^*_i denote the set of i's *collectively coherent CPS-hierarchies.* Battigalli and Siniscalchi (1999) show

[38] Set $Z^0_{Y,i} = \Theta_0 \times Y_{-i}$ and for $k \geq 1$, $Z^k_{Y,i} = Z^{k-1}_{Y,i} \times \Delta\left(Z^{k-1}_{Y,-i}\right)$. The set $T^*_{Y,i}$ of collectively coherent Y-hierarchies is defined analogously to footnote 25, with $H^k_{Y,i}$ denoting the set of player i's *k-order coherent hierarchies.*

[39] $Y = \Theta_0 \times (\Theta_1 \times S_1) \times ... \times (\Theta_n \times S_n)$ is isomorphic to $\Theta \times S$, written $Y \approx \Theta \times S$.

that there is a homeomorphism $g_i^* : \Psi_i^* \to \Delta^{\mathcal{H}}\left(\Theta_{-i} \times S_{-i} \times \Psi^*_{-i}\right)$. Furthermore, under the maintained assumptions, Ψ_i^* is compact.

Definition 9 *The* (universal) Epistemic Model *is defined as* $\left\langle \Omega_0, (\Omega_i, \Psi_i^*, g_i^*)_{i\in N}\right\rangle$ *s.t.:*

1. $\Omega_0 = \Theta_0$

2. *for* $i \in N : \Omega_i = \Theta_i \times S_i \times \Psi_i^*$ *and* $g_i^* : \Psi_i^* \to \Delta^{\mathcal{H}}(\Theta_0 \times \Omega_{-i})$.

For any $i \in N$*, the elements of* Ψ_i^* *are referred to as* i*'s* epistemic types.

Let $\Omega = \times_{j=0}^{n} \Omega_j$ denote the set of *states of the world*, and $\Omega_{-i} = \times_{j\in N\setminus\{i\}}\Omega_j$. Any $\omega = \left(\theta_0, (\theta_i, s_i, \psi_i)_{i\in N}\right) \in \Omega$ specifies a *state of nature* $(\theta_0, \theta_1, ..., \theta_n) \in \Theta$, each player i's *disposition to act* (i.e. his strategy s_i) and his *disposition to believe* (his system of conditional beliefs $g_i^*(\psi_i) = \left(g_{i,h}^*(\psi_i)\right)_{h\in\mathcal{H}}$).

For every $i \in N$, let $f_i : \Psi_i^* \to \Delta^{\mathcal{H}}(\Theta_0 \times \Theta_{-i} \times S_{-i})$ denote his *first-order belief* mapping: for all $\psi_i \in \Psi_i^*$ and $h \in \mathcal{H}$, $f_{i,h}(\psi_i) = marg_{\Theta_0\times\Theta_{-i}\times S_{-i}} g_{i,h}^*(\psi_i)$. Define the event "*player* i *is sequentially rational*" as $SR_i \subseteq \Omega$ s.t.:

$$SR_i = \left\{\omega = \left(\hat{\theta}, \hat{s}, \psi\right) \in \Omega : \hat{s}_i \in r_i\left(f_i(\psi_i)\,|\hat{\theta}_i\right)\right\}, \text{ where}$$

$$r_i\left(f_i(\psi_i)\,|\hat{\theta}_i\right) = \left\{ s_i \in S_i : \begin{array}{c} \text{for all } h \in \mathcal{H}(s_i), \text{ for all } s_i' \in S_i(h) \\ \int_{\Theta_0\times\Theta_{-i}\times S_{-i}} u_i\left(O(s_i, s_{-i}), \theta_0, \theta_{-i}, \hat{\theta}_i\right)\cdot df_{i,h}(\psi_i) \\ \geq \int_{\Theta_0\times\Theta_{-i}\times S_{-i}} u_i\left(O(s_i', s_{-i}), \theta_0, \theta_{-i}, \hat{\theta}_i\right)\cdot df_{i,h}(\psi_i) \end{array} \right\}$$

Let $SR = \bigcap_{i\in N} SR_i$. For each $i \in N$, $h \in \mathcal{H}$ and event $E \in \mathcal{A}$, $B_{i,h}(E) = \{(\theta, s, \psi) \in \Omega : g_{i,h}^*(\theta_i, s_i, \psi_i)[E]$ corresponds to the event "*i is certain of E at history h*". Define $B_h(E) = \bigcap_{i\in N} B_{i,h}(E)$, and let $B_h^0(E) = E$, and $B_h^k(E) = B_h\left(B_h^{k-1}(E)\right)$ for $k = 1, 2, ...$ Then the events $CB_\phi(E) := \bigcap_{k=0}^{\infty} B_\phi^k(E)$, and $MB_\phi^k(E) = \bigcap_{l=0}^{k} B_\phi^l(E)$ (for each $k \geq 0$) correspond, respectively, to "*E and initial common certainty of E*" and to "*E and k-order (initial) mutual certainty of E*". Denote respectively by $CB_{i,\phi}(E)$ and $MB_{i,\phi}^k(E)$ the projections of these events on Ω_i. Notice that $MB_i^0(E) = E_i$ and $MB_{i,\phi}^k(E) = E_i \cap B_{i,\phi}\left(MB_{i,\phi}^{k-1}(E)\right)$; moreover, $CB(E) = \bigcap_{k=0}^{\infty} MB_\phi^k(E)$.

Let $\rho_i^0 : X_i^0 \to Z_{\Theta\times S,i}^0$ be the identity map, and $\left(\widehat{\rho}_{i,h}^0\right)_{h\in\mathcal{H}} : \Delta^{\mathcal{H}}(X_i^0) \to \times_{h\in\mathcal{H}}\Delta\left(Z_{\Theta\times S,i}^0\right)$ the collection of pushforward mappings. In particular, $\hat{\rho}_{i,\phi}^0 \in \Delta\left(Z_{\Theta\times S,i}^0\right)$ is the pushforward of the initial (first order) beliefs. Then, recursively, define

$$\rho_{i,\phi}^k : X_i^k \to Z_{\Theta\times S,i}^k \qquad \text{as} \qquad \left(x_i^{k-1}, \psi_{-i}^k\right) \mapsto \left(\rho_{i,\phi}^{k-1}\left(x_i^{k-1}\right), \widehat{\rho}_{-i,\phi}^{k-1}\left(\psi_{-i}^k\right)\right)$$

where $\widehat{\rho}_{i,\phi}^{k-1} : \Delta\left(X_i^{k-1}\right) \to \Delta\left(Z_{\Theta\times S,i}^{k-1}\right)$ is the pushforward of $\rho_{i,\phi}^{k-1}$. Finally, let

$$\rho_{i,\phi}^{*} : \Omega_i \to T_{\Theta\times S,i}^{*} \quad \text{as} \quad \left(\theta_i, s_i, \psi_i^1, \psi_i^2, \ldots\right) \mapsto \left(\theta_i, s_i, \widehat{\rho}_{i,\phi}^{0}(\psi_i^1), \widehat{\rho}_{i,\phi}^{1}(\psi_i^2), \ldots\right).$$

That is, each *individual state* $\omega_i \in \Omega_i$ induces a $\Theta \times S$-hierarchy at the initial history. Consider the mapping $\mathbf{m}_i : \Omega_i \to T_{\Theta,i}^{*}$ defined as $\boldsymbol{v}_i^{*} = \mathbf{m}_i^{*} \circ \rho_{i,\phi}^{*}$ and for each k, $\boldsymbol{v}_i^{k} = \mathbf{m}_i^{k} \circ \rho_{i,\phi}^{*}$: $\boldsymbol{v}_i$ assigns to each epistemic type a corresponding Θ-hierarchy on the state of nature.

Proposition 6 *Fix a Θ-based type space $\mathcal{T} = \left\langle \Theta_0, (T_i, \tau_i)_{i\in N}\right\rangle$. For each i and each $k \geq 1$,*

$$\mathcal{ISR}_i^{\mathcal{T},k} = \left\{ \left(\hat{t}_i, \hat{s}_i\right) \in T_i \times S_i : \begin{array}{c} \exists \omega_i = (\theta_i, \hat{s}_i, \psi_i) \in MB_{i,\phi}^{k-1}(SR) \\ s.t.\ \boldsymbol{v}_i^{k}(\omega_i) = \left(\hat{\pi}_i^{l}\left(\hat{t}_i\right)\right)_{l=0}^{k} \end{array} \right\}$$

$$\mathcal{ISR}_i^{\mathcal{T}} = \left\{ \left(\hat{t}_i, \hat{s}_i\right) \in T_i \times S_i : \begin{array}{c} \exists \omega_i = (\theta_i, \hat{s}_i, \psi_i) \in CB_{i,\phi}(SR) \\ s.t.\ \boldsymbol{v}_i^{*}(\omega_i) = \hat{\pi}_i^{*}\left(\hat{t}_i\right) \end{array} \right\}$$

The proof exploits the following lemma:

Lemma 5 *Given a Θ-based type space $\mathcal{T}$, fix a CPS $\mu^i \in \Delta^{\mathcal{H}}(\Theta_0 \times T_{-i} \times S_{-i})$ and $\gamma_{-i} : T_{-i} \times S_{-i} \to \Psi_{-i}^{*}$ arbitrarily. Then: $\exists \psi_i \in \Psi_i^{*} : \forall h \in \mathcal{H}$, $g_{i,h}^{*}(\psi_i) \in \Delta\left(\Theta_0 \times \Theta_{-i} \times S_{-i} \times \Psi_{-i}^{*}\right)$ has finite support, and such that*

$$\begin{aligned} & g_{i,h}^{*}(\psi_i)\left[\theta_0, \hat{\pi}_{-i}^{0}(t_{-i}), s_{-i}, \gamma_{-i}(t_{-i}, s_{-i})\right] \\ = & \int_{\tilde{t}_{-i}:\theta_{-i}=\hat{\pi}_{-i}^{0}\left(\tilde{t}_{-i}\right)} \mu^i\left(\theta_0, \tilde{t}_{-i}, s_{-i}|h\right) d\tilde{t}_{-i} \end{aligned}$$

Proof: Let $\left(\delta^i(h)\right)_{h\in\mathcal{H}} \in \times_{h\in\mathcal{H}} \Delta\left(\Theta_0 \times \Theta_{-i} \times S_{-i} \times \Psi_{-i}^{*}\right)$ be defined as

$$\begin{aligned} & \delta^i(h)\left[\theta_0, (\theta_j)_{j\neq i,0}, s_{-i}, \gamma_{-i}(t_{-i}, s_{-i})\right] \\ = & \begin{cases} \displaystyle\int_{\tilde{t}_{-i}:(\theta_j)_{j\neq i,0}=\hat{\pi}_{-i}^{0}\left(\tilde{t}_{-i}\right)} \mu^i\left(\theta_0, \tilde{t}_{-i}, s_{-i}|h\right) d\tilde{t}_{-i} & \\ \qquad\qquad \text{if } (\theta_j)_{j\neq i,0} = \hat{\pi}_{-i}^{0}(t_{-i}) & \\ 0 \qquad \text{otherwise} & \end{cases} \end{aligned}$$

Notice that the map $(\theta_0, t_{-i}, s_{-i}) \mapsto \left(\theta_0, \hat{\pi}_{-i}^{0}(t_{-i}), s_{-i}, \gamma_{-i}(\theta_0, t_{-i}, s_{-i})\right)$ determines an embedding of $\bigcup_{h\in\mathcal{H}} supp\left(marg_{\Theta_0\times\Theta_{-i}\times S_{-i}}\mu^i(\cdot|h)\right)$ (a finite set) in $\Theta_0 \times \Theta_{-i} \times S_{-i} \times \Psi_{-i}^{*}$. Hence $\left(\delta^i(h)\right)_{h\in\mathcal{H}} \in \Delta^{\mathcal{H}}\left(\Theta_0 \times \Theta_{-i} \times S_{-i} \times \Psi_{-i}^{*}\right)$ (i.e. that it is a CPS).[40] By construction, $supp\left(\delta^i(h)\right)$

[40] Cf. Battigalli and Siniscalchi, 2007.

is finite for each $h \in \mathcal{H}$. Finally, since $g_i^* : \Psi_i^* \to \Delta^{\mathcal{H}}\left(\Theta_0 \times \Theta_{-i} \times S_{-i} \times \Psi_{-i}^*\right)$ is onto, there exists $\psi_i \in \Psi_i^* : \forall h \in \mathcal{H}$, $g_{i,h}^*(\psi_i) = \delta^i(h)$.■

proof of proposition 6:

The proof is by induction:

(Part 1: Initial Step:)

Part 1.1 ($\subseteq$)**:** if $\left(\hat{t}_i, \hat{s}_i\right) \in \mathcal{ISR}_i^{\mathcal{T},1}$, $\exists \hat{\mu}^i \in \Delta^{\mathcal{H}}(\Theta_0 \times T_{-i} \times S_{-i})$ s.t.

$$\begin{aligned}
&(1).\ \hat{s}_i \in r_i\left(\hat{\mu}^i|\hat{t}_i\right)\\
&(2).\ \tau_i\left(\hat{t}_i\right) = marg_{\Theta_0 \times T_{-i}}\hat{\mu}^i(\cdot|\phi)\\
&(3).\ \textstyle\int_{\Theta_0 \times T_{-i} \times S_{-i}} \hat{\mu}^i(\cdot|\phi) = 1
\end{aligned}$$

Since $MB_{i,\phi}^0(SR) \equiv SR_i$, we just need to prove that $\exists\omega_i = \left(\hat{\pi}_i^0\left(\hat{t}_i\right), \hat{s}_i, \psi_i\right) \in SR_i$ s.t. $\boldsymbol{v}_i^1(\omega_i) = \hat{\pi}_i^1\left(\hat{t}_i\right)$.

Let $\gamma_{-i}^0 : T_{-i} \times S_{-i} \to \Psi_{-i}^*$ be arbitrarily specified. From lemma 5, $\exists\hat{\psi}_i \in \Psi_i^* : \forall h \in \mathcal{H}$, $g_{i,h}^*\left(\hat{\psi}_i\right) \in \Delta\left(\Theta_0 \times \Theta_{-i} \times S_{-i} \times \Psi_{-i}^*\right)$ has finite support, and such that

$$\begin{aligned}
&g_{i,h}^*\left(\hat{\psi}_i\right)\left[\theta_0, \hat{\pi}_{-i}^0(t_{-i}), s_{-i}, \gamma_{-i}(t_{-i}, s_{-i})\right]\\
= &\int_{\tilde{t}_{-i}:\theta_{-i} = \hat{\pi}_{-i}^0\left(\tilde{t}_{-i}\right)} \mu^i\left(\theta_0, \tilde{t}_{-i}, s_{-i}|h\right) d\tilde{t}_{-i}
\end{aligned}$$

Clearly, $\hat{\omega}_i = \left(\hat{\pi}_i^0\left(\hat{t}_i\right), \hat{s}_i, \psi_i\right)$ is such that $\boldsymbol{v}_i^1(\hat{\omega}_i) = \hat{\pi}_i^1\left(\hat{t}_i\right)$, and $\hat{s}_i \in r_i\left(f_i\left(\hat{\psi}_i\right)|\hat{\pi}_i^0\left(\hat{t}_i\right)\right)$: hence $\hat{\omega}_i \in SR_i$.

Part 1.2 ($\supseteq$)**:** let $\omega_i = \left(\hat{\theta}_i, \hat{s}_i, \psi_i\right) \in SR_i$ s.t. $\boldsymbol{v}_i^1(\omega_i) = \hat{\pi}_i^1\left(\hat{t}_i\right)$, then $\exists\hat{\mu}^i \in \Delta^{\mathcal{H}}(\Theta_0 \times T_{-i} \times S_{-i})$ such that

$$\begin{aligned}
\tau_i\left(\hat{t}_i\right) = marg_{\Theta_0 \times T_{-i}}\hat{\mu}^i(\cdot|\phi) \text{ and } \forall h \in \mathcal{H},\\
marg_{\Theta_0 \times \Theta_{-i} \times S_{-i}}\hat{\mu}^i(\cdot|h) = marg_{\Theta_0 \times \Theta_{-i} \times S_{-i}} g_{i,h}^*(\psi_i)
\end{aligned}$$

Clearly, $\hat{s}_i \in r_i\left(\hat{\mu}^i|\hat{t}_i\right)$, hence $\left(\hat{t}_i, \hat{s}_i\right) \in \mathcal{ISR}_i^1$.

(Inductive Step:) For each $k \geq 1$,

$$\textbf{(I.H.)}\ \mathcal{ISR}_i^{\mathcal{T},m} = \left\{\left(\hat{t}_i, \hat{s}_i\right) \in T_i \times S_i : \begin{array}{c}\exists\omega_i = \left(\hat{\pi}_i^0\left(\hat{t}_i\right), \hat{s}_i, \psi_i\right) \in MB_{i,\phi}^{m-1}(SR)\\ \text{s.t. } \boldsymbol{v}_i^m(\omega_i) = \left(\hat{\pi}_i^l\left(\hat{t}_i\right)\right)_{l=0}^m\end{array}\right\}$$

for all $i \in N$ and $m \leq k$

$$\Rightarrow$$

$$\mathcal{ISR}_i^{\mathcal{T},k+1} = \left\{\left(\hat{t}_i, \hat{s}_i\right) \in T_i \times S_i : \begin{array}{c}\exists\omega_i = \left(\hat{\pi}_i^0\left(\hat{t}_i\right), \hat{s}_i, \psi_i\right) \in MB_{i,\phi}^k(SR)\\ \text{s.t. } \boldsymbol{v}_i^{k+1}(\omega_i) = \hat{\pi}_i^{k+1}\left(\hat{t}_i\right)\end{array}\right\}$$

(Proof of the Inductive Step:)

Part 2.1: ($\subseteq$) let $\left(\hat{t}_i, \hat{s}_i\right) \in \mathcal{ISR}_i^{\mathcal{T},k+1}$. Hence, $\exists \hat{\mu}^i \in \Delta^{\mathcal{H}}\left(\Theta_0 \times T_{-i} \times S_{-i}\right)$ s.t.

$$\begin{array}{l} (1).\ \hat{s}_i \in r_i\left(\hat{\mu}^i | \hat{t}_i\right) \\ (2).\ \tau_i\left(\hat{t}_i\right) = marg_{\Theta_0 \times T_{-i}} \hat{\mu}^i\left(\cdot|\phi\right) \\ (3).\ supp\left(\mu^i\left(\cdot|\phi\right)\right) \subseteq \Theta_0 \times \mathcal{ISR}_{-i}^{\mathcal{T},k} \end{array}$$

From the inductive hypothesis, we can construct $\gamma_{-i} : T_{-i} \times S_{-i} \rightarrow \Psi_{-i}^*$ such that for all $\left(t_{-i}, s_{-i}\right) \in \mathcal{ISR}_{-i}^{\mathcal{T},k}$, $\omega_{-i} \equiv \left(\hat{\pi}_{-i}^0\left(t_{-i}\right), s_{-i}, \gamma_{-i}\left(t_{-i}, s_{-i}\right)\right) \in MB_{-i,\phi}^{k-1}\left(SR\right)$ and it is s.t. $\boldsymbol{v}_{-i}^k\left(\omega_{-i}\right) = \left(\hat{\pi}_{-i}^l\left(t_{-i}\right)\right)_{l=0}^k$. From lemma 5, $\exists \hat{\psi}_i \in \Psi_i^* : \forall h \in \mathcal{H}$, $g_{i,h}^*\left(\hat{\psi}_i\right) \in \Delta\left(\Theta_0 \times \Theta_{-i} \times S_{-i} \times \Psi_{-i}^*\right)$ has finite support, and such that

$$\begin{aligned} & g_{i,h}^*\left(\hat{\psi}_i\right)\left[\theta_0, \hat{\pi}_{-i}^0\left(t_{-i}\right), s_{-i}, \gamma_{-i}\left(t_{-i}, s_{-i}\right)\right] \\ = & \int\limits_{\tilde{t}_{-i}:\theta_{-i}=\hat{\pi}_{-i}^0\left(\tilde{t}_{-i}\right)} \hat{\mu}^i\left(\theta_0, \tilde{t}_{-i}, s_{-i}|h\right) d\tilde{t}_{-i} \end{aligned}$$

Clearly, by construction, for all $t_i \in T_i$, $s_i \in r_i\left(f_i\left(\hat{\psi}_i\right)|\hat{\pi}_i^0\left(t_i\right)\right)$ if and only if $s_i \in r_i\left(\hat{\mu}^i|t_i\right)$. Furthermore, from the definition of γ_{-i} and $supp(\mu^i\left(\cdot|\phi\right)) \subseteq \Theta_0 \times \mathcal{ISR}_{-i}^{\mathcal{T},k}$, the epistemic type thus constructed $\hat{\psi}_i$ is such that $\left(\hat{\pi}_i^0\left(\hat{t}_i\right), \hat{s}_i, \hat{\psi}_i\right) \in MB_{i,\phi}^k\left(SR\right)$ and $\boldsymbol{v}_i^{k+1}\left(\left(\hat{\pi}_i^0\left(\hat{t}_i\right), \hat{s}_i, \hat{\psi}_i\right)\right) = \hat{\pi}_i^{k+1}\left(\hat{t}_i\right)$.

Part 2.2: ($\supseteq$) let $\left(\hat{t}_i, \hat{s}_i\right) \in T_i \times S_i$ be such that $\exists \hat{\omega}_i = \left(\theta_i, \hat{s}_i, \hat{\psi}_i\right) \in MB_{i,\phi}^k\left(SR\right)$ s.t. $\boldsymbol{v}_i^{k+1}\left(\hat{\omega}_i\right) = \hat{\pi}_i^{k+1}\left(\hat{t}_i\right)$. Hence,

$$supp\left(g_{i,\phi}^*\left(\hat{\psi}_i\right)\right) \subseteq \left[MB_{-i,\phi}^{k-1}\left(SR\right) \cap \left\{\omega_{-i} : \exists t_{-i} \in T_{-i} \text{ s.t. } \boldsymbol{v}_{-i}^k\left(\omega_{-i}\right) = \left(\hat{\pi}_{-i}^l\left(t_{-i}\right)\right)_{l=0}^k\right\}\right]$$

Define, for each $\left(t'_{-i}, s'_{-i}\right) \in T_{-i} \times S_{-i}$ and $k \geq 0$

$$\Omega_{-i}^k\left(t'_{-i}, s'_{-i}\right) := \left\{\omega_{-i} \in \Omega_{-i} : \begin{array}{c} \omega_{-i} = \left(\hat{\pi}_{-i}^0\left(t'_{-i}\right), s'_{-i}, \psi_{-i}\right) \\ \text{and } \boldsymbol{v}_{-i}^k\left(\omega_{-i}\right) = \left(\hat{\pi}_{-i}^l\left(t_{-i}\right)\right)_{l=0}^k \end{array}\right\}$$

Notice that $\boldsymbol{v}_i^{k+1}\left(\hat{\omega}_i\right) = \hat{\pi}_i^{k+1}\left(\hat{t}_i\right)$implies that

$$\begin{aligned} & \int_{S_{-i}}\left(\int_{\Omega_{-i}^k\left(t_{-i}, s'_{-i}\right)} g_{i,\phi}^*\left(\hat{\psi}_i\right)\left[\theta_0, \left(\theta_j\right)_{j\neq i,0}, s'_{-i}, \psi_{-i}\right] d\omega_{-i}\right) ds'_{-i} \\ = & \int\limits_{\substack{t'_{-i} \in T_{-i}: \\ \left(\hat{\pi}_{-i}^l\left(t'_{-i}\right)\right)_{l=0}^k = \left(\hat{\pi}_{-i}^l\left(t_{-i}\right)\right)_{l=0}^k}} \tau_i\left(\hat{t}_i\right)\left[\theta_0, t'_{-i}\right] dt'_{-i} \\ \equiv & \tau_i\left(\hat{t}_i\right)\left[\left\{\theta_0\right\} \times H_{-i}^k\left(t_{-i}\right)\right] \end{aligned}$$

Let $\hat{\mu}^i \in \Delta^{\mathcal{H}}(\Theta_0 \times T_{-i} \times S_{-i})$ be such that

$$\hat{\mu}^i(\theta_0, t_{-i}, s_{-i}) = \int_{\Omega^k_{-i}(t_{-i}, s'_{-i})} g^*_{i,\phi}\left(\hat{\psi}_i\right)\left[\theta_0, (\theta_j)_{j \neq i,0}, s'_{-i}, \psi_{-i}\right] d\omega_{-i}$$
$$\cdot \frac{\tau_i\left(\hat{t}_i\right)[\theta_0, t_{-i}]}{\tau_i\left(\hat{t}_i\right)\left[\{\theta_0\} \times H^k_{-i}(t_{-i})\right]}$$
$$\text{and for all } h \in \mathcal{H}\text{: } marg_{\Theta_0 \times \Theta_{-i} \times S_{-i}} \hat{\mu}^i(\cdot|h) = marg_{\Theta_0 \times \Theta_{-i} \times S_{-i}} g^*_{i,h}(\psi_i)$$

Clearly, the latter conditions implies that $r_i\left(\hat{\mu}^i|\hat{t}_i\right)$; $\tau_i\left(\hat{t}_i\right) = marg_{\Theta_0 \times T_{-i}} \hat{\mu}^i(\cdot|\phi)$ is satisfied, simply integrating over S_{-i} both sides of the equation, substituting for $\tau_i\left(\hat{t}_i\right)\left[\{\theta_0\} \times H^k_{-i}(t_{-i})\right]$ and simplifying; finally, under the inductive hypothesis, also $supp(\mu^i(\cdot|\phi)) \subseteq \Theta_0 \times \mathcal{ISR}^{\mathcal{T},k-1}_{-i}$ holds. Thus: $\left(\hat{t}_i, \hat{s}_i\right) \in \mathcal{ISR}^{\mathcal{T},k+1}_i$.

10 Appendix D: ISR in the normal form

The analysis in this section applies to finite models in *PV-environments*, we thus focus on Bayesian games $\Gamma^{\mathcal{T}} = \left\langle N, \bar{\mathcal{H}}, (T_i, \tau_i, \hat{u}_i)_{i \in N} \right\rangle$ s.t. $|T| < \infty$ and $\hat{u}_i : T_i \times \mathcal{Z} \to \mathbb{R}$ for each $i \in N$.[41]

Define the ex-ante payoffs as: For each $(t_i, s_i) \in T_i \times S_i$ and $\sigma_{-i} \in \Delta(T_{-i} \times S_{-i})$,

$$U_i(s_i, \sigma_{-i}, t_i) = \int_{T_{-i} \times S_{-i}} u_i(O(s_i, s_{-i}), \boldsymbol{\theta}_i(t_i))\, d\sigma_{-i}(t_{-i}, s_{-i})$$

Definition 10 *A strategy $s_i \in S_i$ is* weakly dominated *for t_i, if for all $\sigma_{-i} \in \Delta(T_{-i} \times S_{-i})$ s.t.* $\text{marg}_{T_{-i}} \sigma_{-i} = \tau_i(t_i)$ *and s.t.* $\sigma_{-i}[s_{-i}] > 0$ *for each* $s_{-i} \in S_{-i}$

$$s_i \notin \arg\max_{s'_i \in S_i} U_i(s'_i, \sigma_{-i}, t_i)$$

Say that $\Gamma^{\mathcal{T}}$ is in generic position if for every $t_i \in T_i$, $z \neq z'$ implies that $u_i(z, \boldsymbol{\theta}_i(t_i)) \neq u_i(z', \boldsymbol{\theta}_i(t_i))$. Notice that in *PV-environments*, t_i's beliefs $\tau_i(t_i) \in \Delta(T_{-i})$ are payoff irrelevant. The following is thus a well-known fact (e.g., lemma 1.2 in Ben-Porath, 1997).

Lemma 6 *If $\Gamma^{\mathcal{T}}$ is in generic position, s_i is* not weakly dominated *for t_i if and only if s_i is* sequentially rational *for t_i.*[42]

The next definition introduces Dekel and Fudenberg's (1990) $S^\infty W$-procedure for the interim normal form of $\Gamma^{\mathcal{T}}$.

[41] As in section 3 $\hat{u}_i(z, t_i) = u_i(z, \boldsymbol{\theta}_i(t_i))$.

[42] Sequentially rational strategies were defined in definition 5.

Definition 11 *For each* $t_i \in T_i$*, let* $S^0 W_i^{\mathcal{T}}(t_i) \equiv r_i(t_i)$*.For each* $k = 0, 1, 2, ...,$ *and* $t_i \in T_i$*, let* $S^k W_i^{\mathcal{T}}(t_i) = \left\{s_i \in S_i : (t_i, s_i) \in S^k W_i^{\mathcal{T}}\right\}$*,* $S^k W^{\mathcal{T}} = \times_{i=1,...,n} S^k W_i^{\mathcal{T}}$ *and* $S^k W_{-i}^{\mathcal{T}} = \times_{j \neq i, 0} S^k W_j^{\mathcal{T}}$*. Recursively, for* $k = 1, 2, ...,$ *and* $t_i \in T_i$

$$S^k W_i^{\mathcal{T}}(t_i) = \left\{ \hat{s}_i \in S^{k-1} W_i^{\mathcal{T}}(t_i) : \begin{array}{l} \exists \sigma_{-i} \in \Delta\left(S^{k-1} W_{-i}^{\mathcal{T}}\right) \text{ s.t.} \\ (1).\ \tau_i(t_i) = \text{marg}_{T_{-i}} \sigma_{-i} \\ (2).\ \hat{s}_i \in \arg\max_{s_i' \in S_i} U_i(s_i', \sigma_{-i}, t_i) \end{array} \right\}$$

Finally: $S^\infty W_i^{\mathcal{T}}(t_i) = \bigcap_{k \geq 0} S^k W_i^{\mathcal{T}}(t_i)$

Proposition 7 *If* $\Gamma^{\mathcal{T}}$ *is in generic position, for each* $i \in N$*,* $t_i \in T_i$ *and* $k \geq 1$*,*$\mathcal{ISR}_i^{\mathcal{T},k}(t_i) = S^{k-1} W_{-i}^{\mathcal{T}}(t_i)$*. Hence* $\mathcal{ISR}^{\mathcal{T}} = S^\infty W^{\mathcal{T}}$*.*

Proof: From Lemma 6, $\mathcal{ISR}_i^{\mathcal{T},1}(t_i) = S^0 W_i(t_i)$. In the following it will be shown that for each $k \geq 1$ and $t_i \in T_i$, $\mathcal{ISR}_i^{\mathcal{T},k}(t_i) = S^{k-1} W_i^{\mathcal{T}}(t_i)$ implies $\mathcal{ISR}_i^{\mathcal{T},k+1}(t_i) = S^k W_i^{\mathcal{T}}(t_i)$.

Step 1($\subseteq$)**:** Let $\hat{s}_i \in \mathcal{ISR}_i^{\mathcal{T},k+1}(t_i)$ and $\mu^i \in \Phi_i(t_i)$ be s.t. $supp(marg_{S_{-i}} \mu^i(\cdot|\phi)) \subseteq \mathcal{ISR}_{-i}^{\mathcal{T},k}$ and $\hat{s}_i \in r_i(\mu^i|t_i)$. Set $\sigma_{-i} = \mu^i(\cdot|\phi)$. Under the inductive hypothesis, $\sigma_{-i} \in \Delta\left(S^{k-1} W_{-i}^{\mathcal{T}}\right)$ and trivially by construction: $\tau_i(t_i) = marg_{T_{-i}} \sigma_{-i}$ and $\hat{s}_i \in \arg \max_{s_i \in S^{k-1} W_i^{\mathcal{T}}(t_i)} U_i(s_i, \sigma_{-i}, t_i)$. Hence, $\hat{s}_i \in S^k W_i^{\mathcal{T}}(t_i)$.

Step 2($\supseteq$)**:** Let $\hat{s}_i \in S^k W_i^{\mathcal{T}}(t_i)$ and $\hat{\sigma}_{-i} \in \Delta\left(S^{k-1} W_{-i}^{\mathcal{T}}\right)$ s.t. $\tau_i(t_i) = marg_{T_{-i}} \hat{\sigma}_{-i}$ and $\hat{s}_i \in \arg \max_{s_i \in S^{k-1} W_i^{\mathcal{T}}(t_i)} U_i(s_i, \hat{\sigma}_{-i}, t_i)$. From the inductive hypothesis, $\hat{s}_i \in S^k W_i^{\mathcal{T}}(t_i) \subseteq S^{k-1} W_i^{\mathcal{T}}(t_i) = \mathcal{ISR}_i^{\mathcal{T},k}(t_i)$, hence $\exists \mu^i \in \Phi_i(t_i)$ s.t. $supp(marg_{S_{-i}} \mu^i(\cdot|\phi)) \subseteq \mathcal{ISR}_{-i}^{\mathcal{T},k-1}$ and $\hat{s}_i \in r_i(\mu^i|t_i)$. Let $\hat{\mu}^i$ such that $\hat{\mu}^i(\cdot|\phi) = \hat{\sigma}_{-i}$, and for each h at which $\hat{\mu}^i(\cdot|h)$ is not specified by Bayes' Rule, set $\hat{\mu}^i(\cdot|h) = \mu^i(\cdot|h)$. Since $supp(\hat{\sigma}_{-i}) \subseteq S^{k-1} W_{-i}^{\mathcal{T}} = \mathcal{ISR}_{-i}^{\mathcal{T},k}(t_i)$ under the inductive hypothesis, $\hat{\mu}^i \in \Phi_i(t_i)$ and is concentrated on $\mathcal{ISR}_{-i}^{\mathcal{T},k}$. That also $\hat{s}_i \in r_i(\mu^i|t_i)$ holds is immediate, as $\hat{\mu}^i(\cdot|\phi)$ agree with $\hat{\sigma}_{-i}$, and the conditional conjectures at unexpected histories agree with μ^i.■

References

1. Angeletos, G.M., C. Hellwig, and A. Pavan (2006), "Signaling in a Global Game: Coordination and Policy Traps," *Journal of Political Economy,* **114**, 452-484.

2. Angeletos, G.M., C. Hellwig, and A. Pavan (2007), "Dynamic Global Games of Regime Change: Learning, Multiplicity, and the Timing of Attacks," *Econometrica,* **75**, 711-756.

3. Angeletos, G.M., I. Werning (2006), "Crises and Prices: Information Aggregation, Multiplicity and Volatility," *American Economic Review,* **96**, 1721-1737.

4. Battigalli, P. and M. Siniscalchi (1999), Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games, Journal of Economic Theory **88**, 188-230.

5. Battigalli, P. and M. Siniscalchi (2003), "Rationalization and Incomplete Information", *Advances in Theoretical Economics*, **3**, Article 3. [http://www.bepress.com/bejte/advances/

6. Battigalli, P. and M. Siniscalchi (2008), "Interactive Epistemology in Games with Payoff Uncertainty" *Research in Economics,* forthcoming.

7. Battigalli, P. A. Di Tillio, E. Grillo and A. Penta (2008), "Interactive Epistemology and Solution Concepts for Games with Asymmetric Information" *mimeo,* Bocconi University.

8. Ben-Porath, E. (1997), "Rationality, Nash Equilibrium and Backwards Induction in Perfect Information Games," *Review of Economic Studies*, **64**, 23-46.

9. Brandenburger, A. and E. Dekel (1993), "Hierarchies of Beliefs and Common Knowledge," *Journal of Economic Theory,* **59**, 189-198.

10. Carlsson, H. and E. Van Damme (1993), "Global Games and Equilibrium Selection," *Econometrica,* **61**, 989-1018.

11. Chamley, C. (1999), "Coordinating Regime Switches" *Quarterly Journal of Economics*, **114**, 869-905.

12. Chen, Y.C. (2008), "A Structure Theorem for Rationalizability in Extensive Games" *mimeo*, Northwestern University.

13. Dekel, E. and D. Fudenberg (1990) "Rational Behavior with Payoff Uncertainty", *Journal of Economic Theory*, **52**, 243-267.

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

15. Dekel, E. and F. Gul (1997), "Rationality and Knowledge in Game Theory, *Advances in Economics and Econometrics: Theory and Applications: 7th World Congress of the Econometric Society* Vol I , 1997, ed. by David M. Kreps and Kenneth F. Wallis.

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

17. Frankel, D. and A. Pauzner (2000), "Resolving Indeterminacy in Dynamic Settings: the Role of Shocks" *Quarterly Journal of Economics*, **115**, 283-304.

18. Frankel, D., S. Morris and A. Pauzner (2003), "Equilibrium Selection in Global Games with Strategic Complementarities" *Journal of Economic Theory*, **108**, 1-44.

19. Friedenberg, A. and M. Meyer (2007), "The Context of the Game" *mimeo.*

20. Fudenberg, D. and J. Tirole (1991), *Game Theory*, Cambridge MA: MIT Press.

21. Gul, F.(1998) "A Comment on Aumann's Bayesian View", *Econometrica*, **66**, 923-927.

22. Harsanyi, J. (1967-68), "Games of Incomplete Information Played by Bayesian Players. Parts I, II, III," *Management Science,* **14**, 159-182, 320-334, 486-502.

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

24. Morris S. and H.S. Shin (2003), "Global Games: Theory and Applications," in *Advances in Economics and Econometrics, 8th World Congress of the Econometric Society,* ed. by M. Dewatripont, L. Hansen ans S. Turnvosky. Cambridge, UK: Cambridge University Press, 56-114.

25. Pearce, D. (1984), "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica,* **52**, 1029-1050.

26. Penta, A. (2009), "Relaxing (some) Common Knowledge Assumptions in Static Games", *mimeo*, UPenn.

27. Rubinstein, A. (1989), "The Electronic Mail Game: Strategic Behavior under 'Almost Common Knowledge'", *American Economic Review.*, 79, 385-391.

28. Rubinstein, A. (1991), "Comments on the Interpretation of Game Theory", *Econometrica,* **59**.

29. Weinstein, J. and M. Yildiz (2007): "A Structure Theorem for Rationalizability with Application to Robust Predictions of Refinements," *Econometrica,* **75**, 365-400.