Robust Predictions in Dynamic Games with Incomplete Information*

Ziwei Wang†

March 2, 2026

Abstract

This paper studies predictions that are robust to higher-order payoff uncertainty in dynamic games. Players' knowledge about payoffs is represented by an information structure, while their initial beliefs are modeled through a type space. We focus on the robustness of an interim solution concept, *weak rationalizability*, which captures rationality and common initial belief in rationality in dynamic environments. Employing a collection-based approach, we derive conditions that fully characterize strategies that can be uniquely selected and refinements that are locally robust around finite types. We then provide necessary and sufficient conditions under which (i) a Structure Theorem for weak rationalizability holds and (ii) its predictions are generically unique. We further show that strong rationalizability, a key refinement incorporating forward-induction reasoning, is a robust refinement. Finally, we apply our framework to two economic applications, studying robust refinements under higher-order uncertainty regarding the privacy of information and the observability of actions.

KEYWORDS: Dynamic games, sequential rationality, rationalizability, incomplete information, robustness, higher-order beliefs, type spaces.

*I thank Emiliano Catonini, Lester Chan, Yi-Chun Chen, Antonio Penta, Dan Quint, Marzena Rostek, Bill Sandholm, Lones Smith, Marek Weretka, Peio Zuazo-Garin, and participants at 2023 Asia Meeting of the Econometric Society (Tsinghua) for helpful comments and discussions. Financial support from the National Natural Science Foundation of China (No. 72403188) is gratefully acknowledged. The usual disclaimer applies.

†Guanghua School of Management, Peking University. Email: `zwang.econ@gmail.com`.

1 Introduction

Models of strategic interactions often impose strong common knowledge assumptions on payoffs, yet they typically admit a large set of rationalizable or equilibrium strategies. A long line of research has shown that slightly perturbing common knowledge assumptions can sharply refine predictions and even select a unique outcome (see, for example, Rubinstein, 1989; Carlsson and van Damme, 1993; Kajii and Morris, 1997; Morris and Shin, 1998, 2000). Weinstein and Yildiz (2007) substantially generalize this insight in static environments under the so-called *richness condition*, which relaxes all common knowledge assumptions on payoffs. In particular, they show that (i) (the *Structure Theorem*) any interim correlated rationalizable (ICR) strategy of a type can be uniquely selected by perturbing higher-order beliefs, and (ii) (*generic uniqueness*) the prediction delivered by ICR is generically unique in the space that contains all belief hierarchies (i.e., the universal type space). Taken together, these results suggest that multiplicity in standard solution concepts arises from the knife-edge assumptions on beliefs; on the other hand, any attempt to refine the predictions fails to be locally robust, because it rules out some uniquely rationalizable strategy for a nearby type.

Chen (2012) and Penta (2012) extend these results to dynamic games. The richness condition of Weinstein and Yildiz (2007) effectively rules out all genuinely dynamic settings because it requires every strategy to be strictly dominant in some state—an impossibility when certain information sets may not be reached. To address this issue, Chen (2012) studies the normal-form representation of dynamic games while maintaining the solution concept ICR, establishing analogous results under a weakened version of extensive-form richness. Penta (2012) instead studies a more general dynamic setting where players can receive private information about payoffs. Based on sequential rationality, he proposes the solution concept *weak rationalizability* (or *interim sequential rationalizability*), which we adopt in this paper, and proves both the Structure Theorem and generic uniqueness under an appropriate richness condition for dynamic environments

In this paper, we study dynamic environments in which common knowledge about payoffs is represented by an *information structure*. The information structure captures players' *persistent* beliefs that govern belief updating throughout the play. Departing from the existing literature, we dispense with any richness condition and instead analyze robust refinements of weak rationalizability under an arbitrary information structure. This perspective is both natural and economically important. First, the richness condition may be too demanding and can render key features of dynamic games moot. In particular, it implies that every opponent strategy can be made sequentially rational at some state, leaving off-path beliefs essentially unrestricted. As a result, the notion of strong belief, which plays an important role in dynamic reasoning, is no longer relevant. Indeed, we show in Section 4.1 that strong rationalizability constitutes a robust refinement of weak rationalizability, but loses its refinement power under richness. Second, when researchers analyzing a dynamic interaction are confident in a particular information structure that does not satisfy richness, existing results provide little guidance for obtaining robust refinements. Our framework instead offers tools for identifying predictions that remain robust to misspecifications of higher-order beliefs, subject to the informational constraints embodied in the information structure. The applications in Sections 4.2 and 4.3 illustrate how this approach can sharpen predictions in economically relevant settings.

A solution concept is said to be *robust* if it satisfies upper hemicontinuity when viewed as a correspondence defined on the universal type space (Fudenberg et al., 1988; Dekel and Fudenberg, 1990; Weinstein and Yildiz, 2007). Intuitively, robustness requires that the prediction for a given type does not exclude any strategy that is admissible for sufficiently "close" types.[1] Penta (2010) shows that weak rationalizability, denoted by $\mathbf{W}$ in this paper, which captures players' rationality and common initial belief in rationality, is upper hemicontinuous on the universal type space. This raises a natural question: given a finite type space, can weak rationalizability be refined without sacrificing robustness? We focus on finite types because they are ubiquitous in applied game theory. Formally, a prediction for a finite type space specifies a set of strategies for each type, and we say the prediction is a *robust refinement* of weak rationalizability if it is consistent with an upper hemicontinuous sub-correspondence of weak rationalizability defined on the universal type space (Definition 5).

To answer this question, we employ the curb collection approach developed by Chen et al. (2014, 2022).[2] In particular, we define and compute the *upper* and *lower* $\mathbf{W}$ *collections* for each payoff type. The upper (lower) $\mathbf{W}$ collection of a payoff type consists of all sets of strategies that contain (are contained by, respectively) the set of weakly rationalizable strategies for some type in the universal type space consistent with that payoff type. We then define the *local upper* $\mathbf{W}$ *collection* for each type in a given finite type space, which traces out all minimal sets of weakly rationalizable strategies arising in a neighborhood of that type within the universal type space. These local collections are characterized via an iterative procedure. By definition, a strategy is uniquely selected for a finite type if and only if the singleton set containing that strategy belongs to the local upper $\mathbf{W}$ collection of that type (Proposition 2). Moreover, for a given finite type space, a prediction is a robust refinement of weak rationalizability if and only if, for each type, the predicted strategy set intersects every

[1] We endow the universal type space with the product topology, so two types are considered close if their first n orders of beliefs are arbitrarily similar, with n sufficiently large.

[2] The *curb collection* generalizes the notion of a *curb set* proposed by Basu and Weibull (1991). Chen et al. (2014) first introduce this notion to study *robust selections* of ICR; Chen et al. (2022) then use it to study robust refinements of ICR. Both papers focus on static games.

element of that type's local upper **W** collection (Proposition 3).

Building on the characterization of local upper **W** collections, we derive necessary and sufficient conditions under which the Structure Theorem and the generic uniqueness of weak rationalizability hold. Specifically,

- (Proposition 5) A Structure Theorem of **W** holds if and only if every rationalizable strategy of a type in the universal type space is uniquely rationalizable for some type with the same information. Under this condition, **W** constitutes the strongest robust refinement of itself.
- (Proposition 6) Generic uniqueness of EFR holds if and only if every set of rationalizable strategies of a type in the universal type space contains at least one strategy that is uniquely rationalizable for some type with the same information. In this case, an iterative procedure initialized with uniquely rationalizable strategies yields the strongest robust refinement of **W**.

We emphasize that, although the statements above concern the infinite-dimensional universal type space, verifying the conditions only requires solving finite-dimensional optimization problems.

The characterization of local upper **W** collections also provide a tool for identifying robust refinements of weak rationalizability when the Structure Theorem fails to hold. In dynamic environments, the role of strong belief has received considerable attention. Informally, strong belief in rationality requires that even when a player is surprised at an off-path information set, he should maintain the assumption that the opponents are rational as long as this remains consistent with the observed past play. This reasoning embodies a form of forward induction and leads to a refinement of weak rationalizability known as *strong rationalizability* (see Pearce, 1984; Battigalli, 1996, 1997). In Section 4.1, we define strong rationalizability for an arbitrary type space, where the type space is only a model of initial beliefs which may be abandoned once a surprise occurs. We show that this formulation of strong rationalizability is a robust refinement of weak rationalizability. This finding echoes earlier results of Piermont and Zuazo-Garin (2026), who show that strong rationalizability, defined on the universal type space, is an upper hemicontinuous correspondence.

In two economic applications, we further illustrate how our results can be used to refine weak rationalizability without upsetting robustness. The general approach begins with a benchmark setting that is typically assumed to be common knowledge. We then relax the common knowledge assumption in a minimal way that reflects a specific economic concern, while keeping the remaining aspects of the model unchanged. This exercise differs fundamentally from the one based on the richness condition that relaxes all common knowledge assumptions. In the first application, we study two-player signaling games and introduce uncertainty about the privacy of the information received by one party (i.e., the sender). Specifically, we expand the information structure to allow for the possibility that the receiver *knows* the type of the sender. The original benchmark now corresponds to a type space that assumes initial common belief in the privacy of sender's information. Applying our results, we show that, under conditions weaker than richness, weak rationalizability is generically unique on the universal type space, and the strongest robust refinement can sometimes yield strictly sharper predictions (Proposition 8); we illustrate the result in the Beer-Quiche game of Cho and Kreps (1987).

In the second application, we introduce uncertainty about the observability of actions in a two-player static game and fully generalize the analysis in Penta and Zuazo-Garin (2022). In particular, we allow for arbitrary partial observation of the opponent's action. This requires a slight extension of our framework, in which a collection of extensive forms is associated with different states of nature. Each player knows only whether the opponent's action is observable, but not the precise structure governing observability. We show that, regardless of the uncertainty about observability, the prediction consisting solely of the Stackelberg action and the best response to the opponent's Stackelberg action always constitutes a robust refinement of weak rationalizability. Moreover, provided that for each player there exists a state in which the Stackelberg action can be identified, this prediction becomes the *strongest* robust refinement. We further illustrate the effect of asynchronous play: if it is common knowledge that one player cannot be the second mover, then that player's unique Stackelberg outcome becomes the strongest robust refinement of rationalizability, yielding a particularly sharp prediction.

Related Literature. This paper contributes to a game-theoretic literature that studies predictions robust to higher-order payoff uncertainty. Several papers investigate this question in dynamic games, as we do here. In addition to the work by Chen (2012) and Penta (2012) discussed above, Weinstein and Yildiz (2013) extend the insight of Weinstein and Yildiz (2007) to infinite-horizon repeated games and establish an "unrefinable" Folk Theorem in that setting. Piermont and Zuazo-Garin (2026) introduce persistent model misspecifications concerning players' knowledge about payoffs, and study the implications of perturbations to both initial and persistent higher-order beliefs in dynamic environments. While all of these papers rely on a richness condition,[3] the present paper dispenses with this assumption.

This paper is inspired by a branch of the literature that studies robust predictions without imposing richness. Penta (2013) provides sufficient conditions under which the results of Weinstein and Yildiz (2007) continue to hold. Building on this insight, Chen et al. (2022) fully

[3]The richness condition in Piermont and Zuazo-Garin (2026) is implicitly assumed, since the payoff states in their universal type space coincide with the set of all utility functions defined on terminal nodes.

characterize the Structure Theorem and generic uniqueness of ICR under arbitrary payoff uncertainty. Weinstein and Yildiz (2011) analyze the sensitivity of *Bayes Nash equilibrium* to perturbations of higher-order beliefs without relying on any richness assumption. Ely and Peski (2011) show that *regular* types (i.e., types that exhibit strategic continuity) are generic in the universal type space without assuming richness. In a mechanism design setting, Oury and Tercieux (2012) characterize social choice functions that can be partially implemented in a *continuous* manner, so that the desired outcome can be induced for all nearby types of the initial model. While all these papers relax the richness condition in static environments, the present paper considers dynamic games and employs a solution concept tailored to genuinely dynamic settings.

Most of the work mentioned above delivers a negative result: despite the abundance of rationalizable or equilibrium strategies in commonly used models, robust refinements of predictions are generally impossible. While our results identify conditions under which this conclusion indeed applies, they also offer a more positive perspective when researchers are willing to relax only specific aspects of common knowledge assumptions (see the applications in Section 4). In such cases, robustness to higher-order uncertainty may actually *help* refine the multiplicity of rationalizability. This perspective is aligned with the insight of Penta and Zuazo-Garin (2022). In a different direction, Heifetz and Kets (2018) and Germano et al. (2020) weaken the solution concept ICR by introducing (higher-order) uncertainty about limited reasoning ability or limited rationality, respectively. They show that when the assumption of common belief in rationality is perturbed, robust proper refinements become possible.

The remainder of this paper is organized as follows. Section 2 introduces the game-theoretic framework and the solution concept. Section 3 presents the main characterization results. Applications of these results are provided in Section 4. Section 5 concludes.

2 Preliminaries

2.1 Game-Theoretic Model

We consider finite dynamic games with incomplete and possibly imperfect information, perfect recall, and no chance moves (Kuhn, 1953; Battigalli, 1993).[4] A dynamic game $\Gamma = \{\mathcal{E}, \mathcal{I}, u\}$ consists of an extensive form $\mathcal{E}$ that defines the rule of the game, an information structure $\mathcal{I}$ that describes players' knowledge about their payoffs, and a profile u specifying the utilities of players.

[4] Our notation is adapted to encompass multistage games with observable actions (Fudenberg and Tirole, 1991; Osborne and Rubinstein, 1994) as special cases.

Formally, an *extensive form* is given by a tuple

$$\mathcal{E} = \{I, (\mathcal{X}_i)_{i\in I}, \mathcal{Z}, \mathcal{L}, (\mathcal{H}_i)_{i\in I}, (A_i)_{i\in I}\}.$$

There is a finite set of players I. For each player $i \in I$, let $\mathcal{X}_i$ be the set of decision nodes at which player i moves, and $\mathcal{H}_i$ denotes the collection of his *information sets*, which form a partition of $\mathcal{X}_i$. The set of *terminal nodes* is denoted by $\mathcal{Z}$. Each player i has a finite action set A_i; for each information set $h \in \mathcal{H}_i$, let $A_i(h) \subseteq A_i$ denote the actions available to player i at h. Nodes are connected by edges in $\mathcal{L}$, each of which corresponds to an action profile chosen simultaneously by the players active at the node from which the edge emanates. Note that players are allowed to share the same decision node or information set.[5]

To facilitate analysis, write ϕ for the *initial information set* that contains the root node of the game, and assume $\phi \in \mathcal{H}_i$ for all $i \in I$ without loss of generality. Under perfect recall, each player's information sets admit a partial order induced by the precedence relation on decision nodes. For any $h, h' \in \mathcal{H}_i$, we say h *follows* h' if every decision node in h follows one in h'. A strategy of player i specifies an action in $A_i(h)$ for each $h \in \mathcal{H}_i$. We identify two strategies that only differ at precluded information sets, and let S_i denote the set of player i's *reduced form pure strategies* (henceforth *strategies*, for brevity). Let $\mathcal{S}_i = 2^{S_i}\backslash\{\emptyset\}$ denote the collection of nonempty subsets of S_i. The set of player i's opponents' strategies is denoted by $S_{-i} = \times_{j\neq i} S_j$. For every information set $h \in \bigcup_{i\in I} \mathcal{H}_i$, we write $S(h)$ for the set of strategy profiles $s \in \times_{i\in I} S_i$ that reach h. By perfect recall, if $h \in \mathcal{H}_i$, we can decompose $S(h) = S_i(h)\times S_{-i}(h)$, where $S_i(h)$ and $S_{-i}(h)$ represent player i's strategies and his opponents' strategy profiles that do not preclude h, respectively. Moreover, let $\mathcal{H}_i(s_i)$ be the set of player i's information sets not precluded by $s_i \in S_i$; that is, $\mathcal{H}_i(s_i) = \{h \in \mathcal{H}_i : s_i \in S_i(h)\}$. Finally, let $z(s) \in \mathcal{Z}$ denote the terminal node induced by a strategy profile s.

An *information structure* is a tuple

$$\mathcal{I} = \{\Omega, (\Theta_i)_{i\in I}\},$$

where Ω is a finite set of *the states of nature*, and each Θ_i is a finite partition of Ω, which is interpreted as the set of player i's *payoff types*. When $\omega \in \Omega$ is realized, player i knows that the true state lies in $\theta_i(\omega) \subseteq \Omega$. Therefore, each payoff type $\theta_i \in \Theta_i$ represents a piece of player i's hard information about payoffs and is known by player i before the game starts. Let $\theta_{-i}(\omega) = \{\theta_{-i} \in \Theta_{-i} : \omega \in \bigcap_{j\neq i} \theta_j\}$ be the set of player i's opponents' information consistent with the state ω. Note that this partition model of information may not have a product

[5]This formulation allows us to formally distinguish between truly simultaneous moves at a node and situations in which some players move earlier and others follow without observing those actions—a distinction that is conceptually important for our application in Section 4.3.

structure. Finally, for each player $i \in I$, denote by $u_i : \Omega \times \mathcal{Z} \to \mathbb{R}$ his utility function, which depends on the outcome of the game and the state of nature.

We now describe players' belief hierarchies based on an information structure $\mathcal{I}$ (Mertens and Zamir, 1985; Brandenburger and Dekel, 1993). The construction is standard as in the literature except for the informational restrictions on beliefs; see a related construction in Penta and Zuazo-Garin (2022). For each $i \in I$ and $\theta_i \in \Theta_i$, let $Z_i^1(\theta_i) = \Delta(\theta_i)$ be the set of player i's first order beliefs.[6] Iteratively, for $n \geq 1$, let

$$Z_{-i}^n(\theta_i) = \left\{ (\omega, \zeta_{-i}^n) \in \theta_i \times \bigtimes_{j \neq i} Z_j^n(\theta_j) : \omega \in \bigcap_{i' \in I} \theta_{i'} \right\},$$

and define

$$Z_i^{n+1}(\theta_i) = \left\{ \zeta_i^{n+1} = \left(\tau_i^1, \ldots, \tau_i^{n+1}\right) \in Z_i^n(\theta_i) \times \Delta\left(Z_{-i}^n(\theta_i)\right) : \underset{Z_{-i}^{n-1}(\theta_i)}{\operatorname{marg}} \tau_i^{n+1} = \tau_i^n \right\},$$

where $\Delta\left(Z_{-i}^n(\theta_i)\right)$ is the set of $(n+1)$-th order belief of player i who has information θ_i. The set of player i's collectively coherent belief hierarchies with information θ_i is

$$H_i(\theta_i) = \left\{ \zeta_i = \left(\tau_i^1, \tau_i^2, \ldots\right) \in \Delta(\theta_i) \times \bigtimes_{n \geq 1} \Delta\left(Z_{-i}^n(\theta_i)\right) : \left(\tau_i^1, \ldots, \tau_i^n\right) \in Z_i^n(\theta_i) \text{ for all } n \geq 1 \right\}.$$

When analyzing a game with incomplete information, we usually associate it with a type space, which provides a compact description of the players' belief hierarchies at the beginning of the game.

Definition 1 (Type Spaces)**.** A *type space* is a tuple $\mathcal{T} = \{(T_i, \vartheta_i, \kappa_i)_{i \in I}\}$, where T_i is a compact metrizable space that contains types of player i, $\vartheta_i : T_i \to \Theta_i$ is a continuous function that specifies a payoff type for each $t_i \in T_i$, and $\kappa_i : t_i \mapsto \kappa_i(t_i) \in \Delta(\vartheta_i(t_i) \times T_{-i})$ such that $\kappa_i(t_i)\big[\big\{(\omega, t_{-i}) : \omega \in \bigcap_{j \neq i} \vartheta_j(t_j)\big\}\big] = 1$ is a continuous function that describes type t_i's belief about the state of nature and his opponents' types.

A type space is called *finite* if T_i is finite for every i. Each type t_i induces a belief hierarchy of player i, $(\tau_i^1(t_i), \tau_i^2(t_i), \ldots) \in H_i(\vartheta_i(t_i))$, as usual.[7] Generalizing the analysis from Mertens

[6]In this paper, for any metrizable space X, we write $\Delta(X)$ for the space of probability measures defined on the Borel σ-algebra of X. We endow $\Delta(X)$ with the weak* topology, a product space with the product topology, and a finite space with the discrete topology.

[7]To be specific, the first order belief of type t_i is defined by

$$\tau_i^1(t_i)[E] = \kappa_i(t_i)\left[\{(\omega, t_{-i}) : \omega \in E\}\right]$$

and Zamir (1985) and Brandenburger and Dekel (1993), it can be shown that when $H_i(\theta_i)$ is endowed with the product topology, there exists a belief-preserving homeomorphism

$$\beta_i(\theta_i) : H_i(\theta_i) \to \Delta\left(\bigcup_{\omega\in\theta_i} \{\omega\} \times H_{-i}(\omega)\right),$$

where

$$H_{-i}(\omega) = \left\{\zeta_{-i} \in \bigtimes_{j\neq i} H_j(\theta_j) : \omega \in \bigcap_{j\neq i} \theta_j\right\}.$$

We now define a tuple $\mathcal{T}^* = \{(T_i^*, \vartheta_i^*, \kappa_i^*)_{i\in I}\}$, where $T_i^* = \{(\theta_i, \zeta_i) : \theta_i \in \Theta_i \text{ and } \zeta_i \in H_i(\theta_i)\}$, and for each $t_i = (\theta_i, \zeta_i) \in T_i^*$, (i) $\vartheta_i^*(t_i) = \theta_i$ and (ii) $\kappa_i^*(t_i) = \beta_i(\theta_i)(\zeta_i)$. It is easy to check that this tuple $\mathcal{T}^*$ satisfies Definition 1, and is therefore referred to as the *universal type space*.[8] For a given type space $\mathcal{T} = \{(T_i, \vartheta_i, \kappa_i)_{i\in I}\}$, and $t_i \in T_i$, we write $\varphi_i^* : t_i \mapsto (\vartheta_i(t_i), \tau_i^1(t_i), \tau_i^2(t_i), ...) \in T_i^*$ for the mapping that maps each type into the universal type space. Since $\beta_i(\theta_i)$ is belief-preserving, the mapping φ_i^* satisfies

$$\kappa_i^*(\varphi_i^*(t_i))[E] = \kappa_i(t_i)\left[(\omega, t_{-i}) : (\omega, \varphi_{-i}^*(t_{-i})) \in E\right]$$

for all measurable $E \subseteq \vartheta_i(t_i) \times T_{-i}^*$. A type $t_i \in T_i^*$ is called a *finite type* if it can be induced by a type in a finite type space.

2.2 Solution Concept

Fix a player $i \in I$ with payoff type $\theta_i \in \Theta_i$. For every information set $h \in \mathcal{H}_i$, let $[h] \subseteq \theta_i \times S_{-i}$ denote the event that h is not precluded by player i's opponents' strategies, i.e.,

$$[h] = \theta_i \times S_{-i}(h).$$

Definition 2 (Conditional Probability Systems; Rényi (1955), Myerson (1986))**.** A collection $\pi_i = (\pi_i(h))_{h\in\mathcal{H}_i}$ of probability distributions $\pi_i(h) \in \Delta(\theta_i \times S_{-i})$ is called a *conditional probability system* (CPS) over $\theta_i \times S_{-i}$ if the following two conditions are satisfied:

(i) For each information set $h \in \mathcal{H}_i$, $\pi_i(h)[h] = 1$;

(ii) For every $E \subseteq [h] \subseteq [h']$, we have $\pi_i(h)[E] \cdot \pi_i(h')[h] = \pi_i(h')[E]$.

---

for every measurable $E \subseteq \vartheta_i(t_i)$. Moreover, for every measurable $E \subseteq Z_{-i}^1(\vartheta_i(t_i))$,

$$\tau_i^2(t_i)[E] = \kappa_i(t_i)\left[\{(\omega, t_{-i}) : \left(\omega, \tau_{-i}^1(t_{-i})\right) \in E\}\right]$$

defines the second order belief of type t_i. We can therefore recursively compute the entire belief hierarchy $\left(\tau_i^1(t_i), \tau_i^2(t_i), \ldots\right) \in H_i(\vartheta_i(t_i))$ induced by type t_i.

[8] With a slight abuse of terminology, we sometimes also refer to $T^* = \bigtimes_{i\in I} T_i^*$ as the universal type space.

Let $\Delta^{\mathcal{H}_i}(\theta_i \times S_{-i})$ denote the set of CPS over $\theta_i \times S_{-i}$.

A *conjecture* of player i with payoff type θ_i is a CPS π_i over $\theta_i \times S_{-i}$ that describes player i's subjective beliefs about the state of nature and opponents' behavior. Given a conjecture $\pi_i \in \Delta^{\mathcal{H}_i}(\theta_i \times S_{-i})$, we write $r_i(\pi_i \,|\, \theta_i)$ for the set of *sequentially best responses* against π_i, in the sense of Reny (1992); that is, $s_i \in r_i(\pi_i \,|\, \theta_i)$ if and only if, for every information set $h \in \mathcal{H}_i(s_i)$,

$$s_i \in \arg\max_{s'_i \in S_i(h)} \sum_{\omega, s_{-i}} u_i(\omega, z(s'_i, s_{-i}))\pi_i(h)[\omega, s_{-i}].$$

The solution concept we employ in this dynamic environment is *weak rationalizability* applied at the interim stage (also referred to as *interim sequential rationalizability*), as proposed by Penta (2012). This concept generalizes earlier notions developed for complete information settings by Ben-Porath (1997) and Battigalli and Siniscalchi (1999).[9] Intuitively, weak rationalizability captures players' behavior under the assumption of sequential rationality and common belief in sequential rationality at the beginning of the game.

Definition 3 (Weak Rationalizability)**.** Fix a type space $\{(T_i, \vartheta_i, \kappa_i)_{i \in I}\}$. For every $i \in I$ and $t_i \in T_i$, let $\mathbf{W}_i^0(t_i) = S_i$. For $n \geq 1$, write $\mathbf{W}_{-i}^{n-1}(t_{-i}) = \times_{j \neq i} \mathbf{W}_j^{n-1}(t_j)$ and define

$$\mathbf{W}_i^n(t_i) = \left\{ s_i \in S_i : \begin{array}{l} \exists(\pi_i, \mu_i) \in \Delta^{\mathcal{H}_i}(\vartheta_i(t_i) \times S_{-i}) \times \Delta(\vartheta_i(t_i) \times T_{-i} \times S_{-i}) \text{ s.t.} \\ \text{(i) } \operatorname{marg}_{\vartheta_i(t_i) \times T_{-i}} \mu_i = \kappa_i(t_i); \\ \text{(ii) } \mu_i\left[\left\{(\omega, t_{-i}, s_{-i}) : (t_{-i}, s_{-i}) \in \operatorname{graph}\left(\mathbf{W}_{-i}^{n-1}\right)\right\}\right] = 1; \\ \text{(iii) } \pi_i(\phi) = \operatorname{marg}_{\vartheta_i(t_i) \times S_{-i}} \mu_i; \\ \text{(iv) } s_i \in r_i(\pi_i \,|\, \vartheta_i(t_i)) \end{array} \right\}.$$

Finally, let $\mathbf{W}_i(t_i) = \bigcap_{n \geq 0} \mathbf{W}_i^n(t_i)$. Also, denote $\mathbf{W} = \times_{i \in I} \mathbf{W}_i$.

In each round, a player of type t_i entertains an initial belief μ_i consistent with the belief mapping κ_i and the previous round of reasoning. A strategy survives this round of elimination if it is a sequentially best response against some CPS π_i that coincides with the initial belief μ_i at the beginning of the game. The following type-space invariance property of weak rationalizability is important for our analysis.

Lemma 1 (Penta (2010))**.** *Fix a type space* $\{(T_i, \vartheta_i, \kappa_i)_{i \in I}\}$*. For every* $t_i \in T_i$*,* $\mathbf{W}_i(t_i) = \mathbf{W}_i^*(\varphi_i^*(t_i))$*.*

This lemma implies that the predictions delivered by $\mathbf{W}$ depend only on the belief hierarchies induced by types. Therefore, they are invariant to the particular representation of

[9] Battigalli and Siniscalchi (2007) introduce a related notion of *weak* Δ*-rationalizability*, which imposes initial beliefs restrictions under payoff uncertainty.

the type space. Formally, if we use superscripts to indicate the type space on which $\mathbf{W}$ is defined, then for any two type spaces $\mathcal{T}$ and $\mathcal{T}'$, $\mathbf{W}_i^{\mathcal{T}}(t_i) = \mathbf{W}_i^{\mathcal{T}'}(t_i')$ if $\varphi_i^*(t_i) = \varphi_i^*(t_i')$. Due to this invariance property, we can study $\mathbf{W}$ directly as a correspondence defined on the universal type space T^*.

Upper hemicontinuity captures the local robustness of a solution concept—defined on the universal type space—to perturbations in higher-order beliefs. A solution concept that fails this property may exclude strategies played by an arbitrarily "close" type under the product topology on the universal type space. In such cases, its predictions depend sensitively on the precise specification of the entire infinite hierarchy of higher-order beliefs.[10] It turns out that $\mathbf{W}$ is robust in this sense.

Lemma 2 (Penta (2010))**.** *For every $n \geq 0$, $\mathbf{W}_i^n(\cdot)$ is upper hemicontinuous on T_i^*. Therefore, $\mathbf{W}_i(\cdot)$ is upper hemicontinuous on T_i^*; that is, for each $t_i \in T_i^*$ and any sequence $\{t_{i,m}\}_{m\in\mathbb{N}} \subseteq T_i^*$ such that $t_{i,m} \to t_i$, if $s_i \in \mathbf{W}_i(t_{i,m})$ for all m then $s_i \in \mathbf{W}_i(t_i)$.*

The convergence $t_{i,m} \to t_i$ (in the product topology) means that $\vartheta_i^*(t_{i,m}) = \vartheta_i^*(t_i)$ for large enough m, and $\tau_i^n(t_{i,m}) \to \tau_i^n(t_i)$ (in the weak* topology) as $m \to \infty$ for every n.[11] We next define a notion of unique selection that plays a crucial role.

Definition 4 (Unique Selections)**.** Given a type space $\{(T_i, \vartheta_i, \kappa_i)_{i\in I}\}$, we say that strategy $s_i \in S_i$ can be *uniquely selected* for type $t_i \in T_i$ if there exists a sequence $\{t_{i,m}\}_{m\in\mathbb{N}} \subseteq T_i^*$ such that $t_{i,m} \to t_i$, and $\{s_i\} = \mathbf{W}_i(t_{i,m})$ for all m.

We envision a type $t_i \in T_i$ as an element of T_i^* and identify it with $\varphi^*(t_i)$, i.e., the conjunction of its payoff type and belief hierarchy. If a strategy s_i can be uniquely selected for a type t_i, then s_i is the *only* weakly rationalizable strategy for a type arbitrarily close to t_i. Therefore, a researcher cannot reject the strategy s_i being played by type t_i when she cannot precisely pin down the infinite sequence of belief hierarchies.

In many economic applications, we focus on a finite type space $\{(T_i, \vartheta_i, \kappa_i)_{i\in I}\}$ as a concise model of players beliefs. In this case, the analyst makes a *prediction* $P = \times_{i\in I} P_i$ for the types in the type space. Each $P_i : T_i \to \mathcal{S}_i$ is a mapping that assigns to every type of player

[10]The idea of robust prediction captured by upper hemicontinuity originates from Fudenberg et al. (1988) and Dekel and Fudenberg (1990).

[11]For a detailed discussion of the interpretation of robustness with respect to the product topology, see Weinstein and Yildiz (2007). While the product topology is natural and intuitive, other topologies that impose convergence requirements on behavior or beliefs have also been studied. Dekel et al. (2006) introduce the *strategic topology*, which preserves the continuity of behavior, and examine its implications for the denseness of finite types. Chen et al. (2010) further consider the finer *uniform-weak topology* and *uniform-strategic topology*. They show that finite types are nowhere dense under these two topologies, although both are equivalent to the strategic topology around finite types.

i a nonempty subset of strategies that may be played. Following Chen et al. (2022), we define a notion of robust refinement which requires the prediction $P_i(t_i)$ of t_i to coincide with an *upper hemicontinuous* refinement of $\mathbf{W}$ on T_i^*. Intuitively, robustness demands that the prediction $P_i(t_i)$ include at least some weakly rationalizable strategies of the "true" type, when the analyst's model of the belief hierarchy can be slightly misspecified.

Definition 5 (Robust Refinements)**.** Given a finite type space $\{(T_i, \vartheta_i, \kappa_i)_{i\in I}\}$, we say P is a *robust refinement* of weak rationalizability if there exists an upper hemicontinuous sub-correspondence P^* of $\mathbf{W}$ defined on T^* such that $P_i(t_i) = P_i^*(\varphi_i^*(t_i))$ for every $i \in I$ and $t_i \in T_i$.

Note that by Lemmas 1 and 2, the prediction $\mathbf{W}$ is by default a robust refinement of itself. In the next section, we provide a tight condition for a prediction P to be a robust refinement of $\mathbf{W}$. We further use this condition to characterize the Structure Theorem and generic uniqueness of $\mathbf{W}$ (Weinstein and Yildiz, 2007).

3 Characterizations

For our characterization results, we employ a collection-based approach introduced by Chen et al. (2014, 2022), which generalizes the notion of curb sets in Basu and Weibull (1991). In particular, we study collections of subsets of strategies and the corresponding sequential best responses against conjectures restricted by these collections. Although our characterization results bear a formal resemblance to those in Chen et al. (2022), our analysis differs in two important respects. First, rather than focusing on static games, we consider dynamic environments in which sequential rationality is imposed throughout play. Second, players receive information from a commonly known information structure before the game starts, so the relevant collections must be defined based on this information structure. Consequently, these collections are indexed not only by players but also by their payoff types.

3.1 The Upper and Lower W Collections

Our first goal is to compute the following two families of collections: The family of player i's *upper* $\mathbf{W}$ *collections* $\{\overline{\mathcal{R}}_i^{\uparrow}(\theta_i)\}_{\theta_i\in\Theta_i}$, where

$$\overline{\mathcal{R}}_i^{\uparrow}(\theta_i) = \{R_i \in \mathcal{S}_i : \exists t_i \in T_i^* \text{ s.t. } \vartheta_i^*(t_i) = \theta_i \text{ and } R_i \supseteq \mathbf{W}_i(t_i)\},$$

and the family of player i's *lower* $\mathbf{W}$ *collections* $\{\overline{\mathcal{R}}_i^{\downarrow}(\theta_i)\}_{\theta_i\in\Theta_i}$, where

$$\overline{\mathcal{R}}_i^{\downarrow}(\theta_i) = \{R_i \in \mathcal{S}_i : \exists t_i \in T_i^* \text{ s.t. } \vartheta_i^*(t_i) = \theta_i \text{ and } R_i \subseteq \mathbf{W}_i(t_i)\}.$$

The notation with a bar indicates that these collections have an ex ante interpretation: they do not depend on any specific type space and are only defined for each payoff type. Recall that $\mathcal{S}_i = 2^{S_i}\backslash\{\emptyset\}$ and write $\mathcal{S}_{-i} = \times_{j\neq i}\mathcal{S}_j$.

Definition 6. Fix a payoff type $\theta_i \in \Theta_i$. For a given $\overline{\nu}_i \in \Delta(\theta_i \times \mathcal{S}_{-i})$, we say a distribution $\lambda_i \in \Delta(\theta_i \times S_{-i})$ is *consistent with* $\overline{\nu}_i$ if there exists a function $\bar{f}_i : \theta_i \times \mathcal{S}_{-i} \to \Delta(S_{-i})$ such that

(i) $\bar{f}_i(\omega, R_{-i})[R_{-i}] = 1$ for every $R_{-i} \in \mathcal{S}_{-i}$, and

(ii) $\lambda_i[\omega, s_{-i}] = \sum_{R_{-i}\in\mathcal{S}_{-i}} \overline{\nu}_i[\omega, R_{-i}]\bar{f}_i(\omega, R_{-i})[s_{-i}]$.

Moreover, let $\mathrm{C}_i(\overline{\nu}_i \,|\, \theta_i) = \{\lambda_i \in \Delta(\theta_i \times S_{-i}) : \lambda_i \text{ is consistent with } \overline{\nu}_i\}$ be the set of distributions over $\theta_i \times S_{-i}$ that are consistent with $\overline{\nu}_i$.

To compute the upper **W** collections, let $\overline{\mathcal{R}}_i^{\uparrow,0}(\theta_i) = \{S_i\}$ for each $\theta_i \in \Theta_i$ and $i \in I$. For $n \geq 1$, write $\overline{\mathcal{R}}_{-i}^{\uparrow,n-1}(\theta_{-i}) = \times_{j\neq i}\overline{\mathcal{R}}_j^{\uparrow,n-1}(\theta_j)$ and define recursively

$$\overline{\mathcal{R}}_i^{\uparrow,n}(\theta_i) = \left\{ R_i \in \mathcal{S}_i : \begin{array}{l} \exists\overline{\nu}_i \in \Delta(\theta_i \times \mathcal{S}_{-i}) \text{ s.t.} \\ \text{(i) } \overline{\nu}_i\left[\left\{(\omega, R_{-i}) : (\theta_{-i}(\omega), R_{-i}) \in \operatorname{graph}\left(\overline{\mathcal{R}}_{-i}^{\uparrow,n-1}\right)\right\}\right] = 1; \\ \text{(ii) } R_i \supseteq \bigcup_{\pi_i\in\Delta^{\mathcal{H}_i}(\theta_i\times S_{-i}) : \pi_i(\phi)\in\mathrm{C}_i(\overline{\nu}_i \,|\, \theta_i)} r_i(\pi_i \,|\, \theta_i) \end{array} \right\}. \tag{1}$$

Note that the sequence $\overline{\mathcal{R}}_i^{\uparrow,n}(\theta_i)$ is increasing in n and converges in finitely many rounds for each θ_i of player i. For computational purposes, observe that only the *minimal* elements (under set inclusion) of $\overline{\mathcal{R}}_{-i}^{\uparrow,n-1}(\theta_{-i})$ are relevant for computing $\overline{\mathcal{R}}_i^{\uparrow,n}(\theta_i)$. Moreover, the entire collection $\overline{\mathcal{R}}_i^{\uparrow,n}(\theta_i)$ can be determined by identifying its minimal elements. Consequently, in each round of the process described by (1), it suffices to keep track only of these minimal sets.

Similarly, to compute the lower **W** collections, we let $\overline{\mathcal{R}}_i^{\downarrow,0}(\theta_i) = \mathcal{S}_i$ for each $\theta_i \in \Theta_i$ and $i \in I$. For $n \geq 1$, write $\overline{\mathcal{R}}_{-i}^{\downarrow,n-1}(\theta_{-i}) = \times_{j\neq i}\overline{\mathcal{R}}_j^{\downarrow,n-1}(\theta_j)$ and define recursively

$$\overline{\mathcal{R}}_i^{\downarrow,n}(\theta_i) = \left\{ R_i \in \mathcal{S}_i : \begin{array}{l} \exists\overline{\nu}_i \in \Delta(\theta_i \times \mathcal{S}_{-i}) \text{ s.t.} \\ \text{(i) } \overline{\nu}_i\left[\left\{(\omega, R_{-i}) : (\theta_{-i}(\omega), R_{-i}) \in \operatorname{graph}\left(\overline{\mathcal{R}}_{-i}^{\downarrow,n-1}\right)\right\}\right] = 1; \\ \text{(ii) } R_i \subseteq \bigcup_{\pi_i\in\Delta^{\mathcal{H}_i}(\theta_i\times S_{-i}) : \pi_i(\phi)\in\mathrm{C}_i(\overline{\nu}_i \,|\, \theta_i)} r_i(\pi_i \,|\, \theta_i) \end{array} \right\}. \tag{2}$$

In this case, the sequence $\overline{\mathcal{R}}_i^{\downarrow,n}(\theta_i)$ is decreasing in n and converges in finitely many rounds for each type θ_i of player i. This time, only the *maximal* elements (under set inclusion) need to be tracked throughout this procedure. The limits of the two procedures above exactly characterize the upper and lower **W** collections.

Proposition 1. *For every $i \in I$ and $\theta_i \in \Theta_i$, we have $\overline{\mathcal{R}}_i^{\uparrow}(\theta_i) = \bigcup_{n\geq 0}\overline{\mathcal{R}}_i^{\uparrow,n}(\theta_i)$, and $\overline{\mathcal{R}}_i^{\downarrow}(\theta_i) = \bigcap_{n\geq 0}\overline{\mathcal{R}}_i^{\downarrow,n}(\theta_i)$.*

All proofs are relegated to the appendix.

3.2 Unique Selections and Robust Predictions

We now focus on an arbitrary *finite* type space $\mathcal{T} = \{(T_i, \vartheta_i, \kappa_i)_{i\in I}\}$. We begin by slightly modifying Definition 6 to adapt it to the analysis of types at the interim stage.

Definition 7. Fix a type $t_i \in T_i$. For a given $\nu_i \in \Delta(\vartheta_i(t_i) \times T_{-i} \times \mathcal{S}_{-i})$ such that $\text{marg}_{\vartheta_i(t_i)\times T_{-i}} \nu_i = \kappa_i(t_i)$, we say a distribution $\lambda_i \in \Delta(\vartheta_i(t_i) \times S_{-i})$ is *consistent with* ν_i if there exists a function $f_i : \vartheta_i(t_i) \times T_{-i} \times \mathcal{S}_{-i} \to \Delta(S_{-i})$ such that

(i) $f_i(\omega, t_{-i}, R_{-i})[R_{-i}] = 1$,

(ii) $\lambda_i[\omega, s_{-i}] = \sum_{t_{-i}\in T_{-i}} \sum_{R_{-i}\in\mathcal{S}_{-i}} \nu_i[\omega, t_{-i}, R_{-i}] f_i(\omega, t_{-i}, R_{-i})[s_{-i}]$.

As before, let $\text{C}_i(\nu_i \,|\, t_i) = \{\lambda_i \in \Delta(\vartheta_i(t_i) \times S_{-i}) : \lambda_i \text{ is consistent with } \nu_i\}$ be the set of distributions over $\vartheta_i(t_i) \times S_{-i}$ that are consistent with ν_i.

We now define, for each finite type $t_i \in T_i$, the *local upper* $\mathbf{W}$ *collection* $\mathcal{R}_i^{\text{loc}}(t_i)$, which consists of all strategy sets that constrain the rationalizable behavior of types in a neighborhood of t_i. Formally, for each type $t_i \in T_i$, let

$$\mathcal{R}_i^{\text{loc}}(t_i) = \{R_i \in \mathcal{S}_i : \exists \{t_{i,m}\}_{m\in\mathbb{N}} \subseteq T_i^* \text{ s.t. } t_{i,m} \to t_i \text{ and } R_i \supseteq \mathbf{W}_i(t_{i,m}) \; \forall m\}\,.$$

These collections allow us to identify strategies that can be uniquely selected, as well as predictions that are robust refinements of $\mathbf{W}$. For example, by definition, any strategy uniquely selected for a type corresponds to a singleton set in the local upper $\mathbf{W}$ collection for that type.

Proposition 2 (Unique Selections)**.** *A strategy s_i can be uniquely selected for a finite type t_i if and only if $\{s_i\} \in \mathcal{R}_i^{\text{loc}}(t_i)$.*

Moreover, a prediction is a robust refinement of weak rationalizability precisely when, for each type, it assigns a set of strategies that intersects every element of that type's local upper $\mathbf{W}$ collection.[12]

Proposition 3 (Robust Refinements)**.** *Given a finite type space $\{(T_i, \vartheta_i, \kappa_i)_{i\in I}\}$. A prediction P is a robust refinement of weak rationalizability if and only if for every $i \in I$ and $t_i \in T_i$, we have $P_i(t_i) \cap R_i \neq \varnothing$ for all $R_i \in \mathcal{R}_i^{\text{loc}}(t_i)$.*

[12]Therefore, robust refinements are analogous to the notion of robust equilibrium sets introduced by Morris and Ui (2005) for complete-information games. In particular, an equilibrium set is robust if every perturbed incomplete-information game in which payoffs coincide with those of the complete-information game with high probability admits an equilibrium whose induced behavior is close to some equilibrium in the set.

Our next objective is to describe a procedure for computing the local upper **W** collections associated with a given finite type space. Fixing $(\nu_i, \overline{\nu}_i) \in \Delta(\vartheta_i(t_i) \times T_{-i} \times \mathcal{S}_{-i}) \times \Delta(\vartheta_i(t_i) \times \mathcal{S}_{-i})$ and an arbitrary $\varepsilon \in (0,1]$, we define

$$\mathrm{C}_i(\nu_i, \overline{\nu}_i, \varepsilon \,|\, t_i) = \left\{ \lambda_i \in \Delta(\vartheta_i(t_i) \times S_{-i}) : \begin{array}{l} \exists (\lambda_i', \lambda_i'') \in \mathrm{C}_i(\nu_i \,|\, t_i) \times \mathrm{C}_i(\overline{\nu}_i \,|\, \vartheta_i(t_i)) \text{ s.t.} \\ \lambda_i = (1-\varepsilon)\lambda_i' + \varepsilon \lambda_i'' \end{array} \right\}.$$

We now introduce an iterative procedure to compute $\mathcal{R}_i^{\mathrm{loc}}(t_i)$. For each $i \in I$ and $t_i \in T_i$, set $\mathcal{R}_i^{\mathrm{loc},0}(t_i) = \overline{\mathcal{R}}_i^{\uparrow}(\vartheta_i(t_i))$. For $n \geq 1$, let $\mathcal{R}_{-i}^{\mathrm{loc},n-1}(t_{-i}) = \bigtimes_{j \neq i} \mathcal{R}_j^{\mathrm{loc},n-1}(t_j)$ and define

$$\mathcal{R}_i^{\mathrm{loc},n}(t_i) = \left\{ R_i \in \overline{\mathcal{R}}_i^{\uparrow}(\vartheta_i(t_i)) : \begin{array}{l} \forall \varepsilon \in (0,1], \\ \exists (\nu_i, \overline{\nu}_i) \in \Delta(\vartheta_i(t_i) \times T_{-i} \times \mathcal{S}_{-i}) \times \Delta(\vartheta_i(t_i) \times \mathcal{S}_{-i}) \text{ s.t.} \\ \text{(i) } \mathrm{marg}_{\vartheta_i(t_i) \times T_{-i}} \nu_i = \kappa_i(t_i); \\ \text{(ii) } \nu_i \left[\left\{ (\omega, t_{-i}, R_{-i}) : (t_{-i}, R_{-i}) \in \mathrm{graph}\left(\mathcal{R}_{-i}^{\mathrm{loc},n-1} \right) \right\} \right] = 1; \\ \text{(iii) } \overline{\nu}_i \left[\left\{ (\omega, R_{-i}) : (\theta_{-i}(\omega), R_{-i}) \in \mathrm{graph}\left(\overline{\mathcal{R}}_{-i}^{\uparrow} \right) \right\} \right] = 1; \\ \text{(iv) } R_i \supseteq \bigcup_{\pi_i \in \Delta^{\mathcal{H}_i}(\vartheta_i(t_i) \times S_{-i}) : \pi_i(\phi) \in \mathrm{C}_i(\nu_i, \overline{\nu}_i, \varepsilon \,|\, t_i)} r_i(\pi_i \,|\, \vartheta_i(t_i)). \end{array} \right\} \tag{3}$$

Each sequence $\left\{ \mathcal{R}_i^{\mathrm{loc},n}(t_i) \right\}_{n \geq 0}$ is decreasing and converges in finitely many rounds. Intuitively, at each round n, the distribution ν_i represents a small perturbation of type t_i's belief hierarchy from the second order up to order n, while $\overline{\nu}_i$ captures an arbitrary but rationalizable first-order perturbation weighted by a small probability ε. Taken together, these distributions describe how player i behaves under slight perturbations of the first n levels of the belief hierarchy, which is then summarized by $\mathcal{R}_i^{\mathrm{loc},n}(t_i)$. Therefore, the limit of this procedure characterizes the local upper **W** collections.

Proposition 4. *Fix a finite type space $\{(T_i, \vartheta_i, \kappa_i)_{i \in I}\}$. For each player $i \in I$ and type $t_i \in T_i$, we have $\mathcal{R}_i^{\mathrm{loc}}(t_i) = \bigcap_{n \geq 0} \mathcal{R}_i^{\mathrm{loc},n}(t_i)$.*

3.3 The Structure Theorem and Generic Uniqueness

In this section, we build on our previous analysis to identify exact conditions on the information structure under which the Structure Theorem and generic uniqueness of weak rationalizability hold, respectively.

Proposition 5 (The Structure Theorem)**.** *For a given information structure, the following statements are equivalent:*

(1) For every finite type $t_i \in T_i^*$, any strategy $s_i \in \mathbf{W}_i(t_i)$ can be uniquely selected for t_i;

(2) For every $i \in I$, $\theta_i \in \Theta_i$, and $R_i \in \overline{\mathcal{R}}_i^{\downarrow}(\theta_i)$, we have $\{s_i\} \in \overline{\mathcal{R}}_i^{\uparrow}(\theta_i)$ for all $s_i \in R_i$.

This characterization provides a practical tool to determine whether a Structure Theorem for weak rationalizability holds: it suffices to compute the families of upper and lower $\mathbf{W}$ collections for each player i. Although the proof of Proposition 5 relies on the characterization of the local upper $\mathbf{W}$ collections $\mathcal{R}_i^{\text{loc}}(t_i)$, computing them is not necessary when *applying* Proposition 5 (the same observation applies to Proposition 6 below). Our condition is both sufficient and necessary: the Structure Theorem holds if and only if every strategy that is rationalizable for a type is uniquely rationalizable for some type with the same payoff type. This parallels the characterizations obtained by Chen (2012) and Chen et al. (2022) for static games.

Remark 1. The Structure Theorem for weak rationalizability generalizes the one for interim correlated rationalizability (ICR) in static games. Here, we emphasize a caveat in the interpretation of this generalization. In static games, enlarging the space of payoff uncertainty permits richer perturbations of higher-order beliefs but does not alter the ICR predictions for a given type space. Therefore, when working with a finite type space, a researcher can conclude that, provided richness holds, the computed ICR strategies deliver the strongest robust prediction. However, no analogous statement can be made for weak rationalizability $\mathbf{W}$ (or its refinements) in dynamic games. This is because $\mathbf{W}$ is sensitive to the common knowledge assumption embedded in the information structure, so enriching the candidate information structure may change the set of weakly rationalizable strategies associated with a given type space.[13]

Proposition 6 (Generic Uniqueness)**.** *For a given information structure, the following statements are equivalent:*

(1) For every $i \in I$, the set $\mathcal{U}_i = \{t_i \in T_i^* : |\mathbf{W}_i(t_i)| = 1\}$ is open and dense in T_i^*;

(2) For every $i \in I$, $\theta_i \in \Theta_i$, and $R_i \in \overline{\mathcal{R}}_i^{\uparrow}(\theta_i)$, we have $\{s_i\} \in \overline{\mathcal{R}}_i^{\uparrow}(\theta_i)$ for some $s_i \in R_i$.

This result shows that the $\mathbf{W}$ correspondence on T^* is generically single-valued if and only if every set of rationalizable strategies for a type contains at least one strategy that is uniquely rationalizable for some type with the same payoff type. This condition is strictly weaker than the one characterizing the Structure Theorem. Note that when the richness condition of Penta (2012) holds, we have $\{s_i\} \in \overline{\mathcal{R}}_i^{\uparrow}(\theta_i)$ for every $s_i \in S_i$, so condition (2) of both Propositions 5 and 6 is satisfied. Hence, these two propositions together imply Theorem 1 of Penta (2012).

[13]In the terminology of Penta (2012), while ICR is *information-invariant* in static settings, weak rationalizability need not be information-invariant in general.

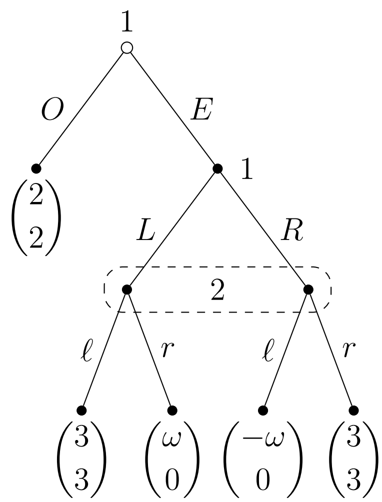


Figure 1: Example 1—A two-player game.

Remark 2. The proof of Proposition 6 yields an induced solution concept that is of interest in its own right. Define

$$\overline{\mathbf{W}}_i^{\mathrm{u}}(\theta_i) = \left\{ s_i \in S_i : \{s_i\} \in \overline{\mathcal{R}}_i^{\uparrow}(\theta_i) \right\}$$

as the set of strategies that are uniquely rationalizable for payoff type θ_i.[14] For each $t_i \in T_i$, let $\mathbf{W}_i^{\mathrm{u},0}(t_i) = \overline{\mathbf{W}}_i^{\mathrm{u}}(\vartheta_i(t_i))$. For $n \geq 1$, let $\mathbf{W}_{-i}^{\mathrm{u},n-1}(t_{-i}) = \times_{j \neq i} \mathbf{W}_j^{\mathrm{u},n-1}(t_j)$, and define

$$\mathbf{W}_i^{\mathrm{u},n}(t_i) = \left\{ s_i \in \overline{\mathbf{W}}_i^{\mathrm{u}}(\vartheta_i(t_i)) : \begin{array}{l} \exists(\pi_i, \mu_i) \in \Delta^{\mathcal{H}_i}(\vartheta_i(t_i) \times S_{-i}) \times \Delta(\vartheta_i(t_i) \times T_{-i} \times S_{-i}) \text{ s.t.} \\ \text{(i) } \operatorname{marg}_{\vartheta_i(t_i) \times T_{-i}} \mu_i = \kappa_i(t_i); \\ \text{(ii) } \mu_i \left[\left\{ (\omega, t_{-i}, s_{-i}) : (t_{-i}, s_{-i}) \in \operatorname{graph}\left(\mathbf{W}_{-i}^{\mathrm{u},n-1} \right) \right\} \right] = 1; \\ \text{(iii) } \pi_i(\phi) = \operatorname{marg}_{\vartheta_i(t_i) \times S_{-i}} \mu_i; \\ \text{(iv) } s_i \in r_i(\pi_i \,|\, \vartheta_i(t_i)) \end{array} \right\}.$$

Finally, let $\mathbf{W}_i^{\mathrm{u}}(t_i) = \bigcap_{n \geq 0} \mathbf{W}_i^{\mathrm{u},n}(t_i)$. This procedure defines a version of weak rationalizability restricted to strategies that are uniquely rationalizable. The arguments in Appendix A.1.6 imply that, when condition (2) of Proposition 6 is satisfied, $\mathbf{W}^{\mathrm{u}}$ is nonempty and delivers *the* strongest robust refinement of $\mathbf{W}$. That is, $\mathbf{W}^{\mathrm{u}}$ refines any upper hemicontinuous sub-correspondence of $\mathbf{W}$, and any proper refinement of $\mathbf{W}^{\mathrm{u}}$ fails to be upper hemicontinuous. Of course, when condition (2) of Proposition 5 holds, we have $\mathbf{W}^{\mathrm{u}} = \mathbf{W}$.

We now use an example to illustrate how to apply our characterization results.

[14] This set plays a role analogous to $\mathcal{A}_i^{\infty}$ in Penta (2013) and $\mathcal{B}_i$ in Penta and Zuazo-Garin (2022).

Example 1. Consider the two-player game depicted in Figure 1. Player 1 moves first and chooses either to opt out (O) or to enter the second stage (E). If player 1 enters, he then chooses between left (L) and right (R), after which player 2 chooses between left (ℓ) and right (r) without observing player 1's choice. Payoff uncertainty concerns only player 1's payoff when the second stage is reached, and it is common knowledge that player 1 privately knows his own payoff. Therefore, Θ_1 is the discrete partition of Ω, Θ_2 is the trivial partition, and u_2 is constant on Ω. We examine three possible specifications of Ω.

Case 1. First, suppose the analyst models this strategic interaction as a complete information game with $\Omega = \{0\}$; i.e., it is common *knowledge* that player 1 obtains 0 after miscoordination in the second stage. The analyst makes a prediction using weak rationalizability. In this case, the universal type space is degenerate and only contains one type for each player: for $i = 1, 2$,

$$\vartheta_i^*\big(t_i^{\text{CB}}\big) = \{0\},\ T_i^* = \big\{t_i^{\text{CB}}\big\} \text{ and } \kappa_i\left(t_i^{\text{CB}}\right)\left[t_{-i}^{\text{CB}}\right] = 1.$$

It is easy to see that all strategies are rationalizable. For player 1, EL is the (sequentially) best response to ℓ, ER is the best response to r, and O is the best response to a conjecture that attaches equal probabilities to ℓ and r; For player 2, ℓ is the best response to EL, and r to ER. Therefore, $\mathbf{W}_1\big(t_1^{\text{CB}}\big) = \{O, EL, ER\}$ and $\mathbf{W}_2\big(t_2^{\text{CB}}\big) = \{\ell, r\}$. Moreover, since the universal type space is degenerate, any refinement of $\mathbf{W}$ is robust by definition.

Case 2. Now suppose the analyst relaxes the common knowledge assumption and acknowledges the *possibility* that $\omega = 4$. However, for model simplicity, she maintains the assumption of common initial belief in $\omega = 0$. In other words, the analyst continues to work with a type space that only contains t_1^{CB} and t_2^{CB}, but the universal type space becomes larger due to the richer payoff uncertainty $\Omega = \{0, 4\}$. Using the argument in Case 1, we have $\mathbf{W}_1\big(t_1^{\text{CB}}\big) = \{O, IL, IR\}$ and $\mathbf{W}_2\big(t_2^{\text{CB}}\big) = \{\ell, r\}$.[15] The upper $\mathbf{W}$ collections can be computed according to definition (1), and we only need to record those minimal elements in the iterative procedure, as shown in the following table

n	0	1	2	3	$\cdots$
$\overline{\mathcal{R}}_1^{\uparrow,n}(\theta_1 = \{0\})$	S_1	S_1	S_1	$\{EL\}$	$\{EL\}$
$\overline{\mathcal{R}}_1^{\uparrow,n}(\theta_1 = \{4\})$	S_1	$\{EL\}$	$\{EL\}$	$\{EL\}$	$\{EL\}$
$\overline{\mathcal{R}}_2^{\uparrow,n}(\theta_2 = \Omega)$	S_2	S_2	$\{\ell\}$	$\{\ell\}$	$\{\ell\}$

There are two key steps in the computation above. In the second round, we can pick a distribution $\overline{\nu}_2 \in \Delta(\Omega \times \mathcal{S}_1)$ for player 2 such that $\overline{\nu}_2[(4, \{EL\})] = 1$; therefore, $\{\ell\}$ becomes an element of $\overline{\mathcal{R}}_2^{\uparrow,2}(\theta_2)$. Then in the third round, we can pick a distribution $\overline{\nu}_1 \in \Delta(\{0\} \times \mathcal{S}_2)$

[15] Note that such invariance of weakly rationalizable strategies with respect to the information structure is not a general property. See the discussion following Proposition 5 and the analysis of *information invariance* in Penta (2012).

for player 1 such that $\bar{\nu}_1[(0,\{\ell\})] = 1$, which makes $\{EL\}$ an element of $\overline{\mathcal{R}}_1^{\uparrow,3}(\theta_1 = \{0\})$. The process converges after four rounds.

Now applying Proposition 6, we know the prediction of weak rationalizability is generically unique on the universal type space. In particular, only EL is uniquely rationalizable for player 1 regardless of his private information, and ℓ for player 2. Therefore, $\mathbf{W}_1^{\mathrm{u}}\left(t_1^{\mathrm{CB}}\right) = \{EL\}$ and $\mathbf{W}_2^{\mathrm{u}}\left(t_2^{\mathrm{CB}}\right) = \{\ell\}$ is the strongest robust refinement of $\mathbf{W}$ for the complete-information types that the analyst uses to model players' beliefs.

Case 3. The analyst maintains the assumption of common initial belief in $\omega = 0$, but now relaxed the common knowledge assumption to $\Omega = \{0, 4, -4\}$. As before, we have $\mathbf{W}_1\left(t_1^{\mathrm{CB}}\right) = \{O, EL, ER\}$ and $\mathbf{W}_2\left(t_2^{\mathrm{CB}}\right) = \{\ell, r\}$. We now summarize the iterative procedure of computing upper $\mathbf{W}$ collections by recording the minimal elements in each round

n	0	1	2	3	$\cdots$
$\overline{\mathcal{R}}_1^{\uparrow,n}(\theta_1 = \{0\})$	S_1	S_1	S_1	$\{O\}, \{EL\}, \{ER\}$	$\{O\}, \{EL\}, \{ER\}$
$\overline{\mathcal{R}}_1^{\uparrow,n}(\theta_1 = \{4\})$	S_1	$\{EL\}$	$\{EL\}$	$\{EL\}$	$\{EL\}$
$\overline{\mathcal{R}}_1^{\uparrow,n}(\theta_1 = \{-4\})$	S_1	$\{ER\}$	$\{ER\}$	$\{ER\}$	$\{ER\}$
$\overline{\mathcal{R}}_2^{\uparrow,n}(\theta_2 = \Omega)$	S_2	S_2	$\{\ell\}, \{r\}$	$\{\ell\}, \{r\}$	$\{\ell\}, \{r\}$

Note that $\{O\}$ enters $\overline{\mathcal{R}}_1^{\uparrow,n}(\theta_1 = \{0\})$ starting from the third round, because it can be justified by a distribution $\bar{\nu}_1 \in \Delta(\{0\} \times \mathcal{S}_2)$ such that $\bar{\nu}_1[(0,\{\ell\})] = \bar{\nu}_1[(0,\{r\})] = \frac{1}{2}$. Intuitively, even though player 1 knows that $\omega = 0$, he may assign probability $\frac{1}{2}$ to a player 2 who chooses ℓ while believing that the state is $\omega = 4$ and player 1 plays EL, and complementary probability $\frac{1}{2}$ to a player 2 who chooses r while believing that the state is $\omega = -4$ and player 1 plays ER. Under such a conjecture, playing O becomes the unique best response.

On the other hand, the lower $\mathbf{W}$ collections can be easily seen from the facts that $\mathbf{W}_1\left(t_1^{\mathrm{CB}}\right) = S_1$, $\mathbf{W}_2\left(t_2^{\mathrm{CB}}\right) = S_2$, and player 1 has a dominant strategy when $\theta_1 = \{4\}$ or $\{-4\}$. Specifically,

$$\overline{\mathcal{R}}_1^{\downarrow}(\theta_1 = \{0\}) = \mathcal{S}_1,\ \overline{\mathcal{R}}_1^{\downarrow}(\theta_1 = \{4\}) = \{EL\},\ \overline{\mathcal{R}}_1^{\downarrow}(\theta_1 = \{-4\}) = \{ER\},$$

and

$$\overline{\mathcal{R}}_2^{\downarrow}(\theta_2 = \Omega) = \mathcal{S}_2.$$

Applying Proposition 5, we conclude that the Structure Theorem holds for weak rationalizability; in other words, every rationalizable strategy can be uniquely selected for t_1^{CB} and t_2^{CB}. Therefore, any proper refinement of $\mathbf{W}$ is not a robust one.[16] ◇

[16] The information structure in Case 3 also provides an example where the richness condition of Penta (2012) is not satisfied, but the Structure Theorem does hold by our characterization result.

4 Applications

In this section, we demonstrate how our characterization results can be used in three applications. We first show that a refinement of weak rationalizability, termed *strong rationalizability* that is of particular relevance in dynamic games, is robust. In the remaining two applications, we begin with modeling frameworks commonly used in applied game theory, such as finite type spaces or complete-information environments, and then relax common knowledge assumptions in ways tailored to specific modeling concerns. The idea is that we, as researchers, may have confidence in certain aspects of a model while being uncertain about others. We show that robustness to higher-order uncertainty can, in some cases, help us sharpen our predictions.

4.1 Strong Rationalizability: A Robust Refinement

The notion of *strong rationalizability*, sometimes referred to as *extensive form rationalizability*, was first studied by Pearce (1984) and Battigalli (1996, 1997).[17] Beyond assuming that opponents are rational, this concept embodies a "best rationalization principle": players should always believe that their opponents are employing one of the "most rational" strategy profiles that are consistent with their private information and observed behavior, attributing any inconsistency to the opponents' highest level of strategic sophistication. This logic leads to a form of forward induction reasoning, whereby observed behavior shapes a player's conjectures about opponents' information and their future play.

In the literature, strong rationalizability is typically defined in a belief-free manner, without specifying a type space that models players beliefs and higher-order beliefs (see, for example, Battigalli and Siniscalchi, 2002). We first present this definition and subsequently introduce an interim version based on it.

Definition 8 (Belief-Free Strong Rationalizability). For every $i \in I$ and $\theta_i \in \Theta_i$, let $\overline{\mathbf{S}}_i^0(\theta_i) = S_i$. For $n \geq 1$, write $\overline{\mathbf{S}}_{-i}^{n-1}(\theta_{-i}) = \times_{j \neq i} \overline{\mathbf{S}}_j^{n-1}(\theta_j)$ and define

$$\overline{\mathbf{S}}_i^n(\theta_i) = \left\{ s_i \in \overline{\mathbf{S}}_i^{n-1}(\theta_i) : \begin{array}{l} \exists \pi_i \in \Delta^{\mathcal{H}_i}(\theta_i \times S_{-i}) \text{ s.t.} \\ \text{(i) } \pi_i(\phi)\left[\left\{(\omega, s_{-i}) : (\theta_{-i}(\omega), s_{-i}) \in \text{graph}\left(\overline{\mathbf{S}}_{-i}^{n-1}\right)\right\}\right] = 1; \\ \text{(ii) } \forall h \in \mathcal{H}_i, \left(\bigcup_{\omega \in \theta_i} \theta_{-i}(\omega) \times S_{-i}(h)\right) \cap \text{graph}\left(\overline{\mathbf{S}}_{-i}^{n-1}\right) \neq \varnothing \\ \text{implies } \pi_i(h)\left[\left\{(\omega, s_{-i}) : (\theta_{-i}(\omega), s_{-i}) \in \text{graph}\left(\overline{\mathbf{S}}_{-i}^{n-1}\right)\right\}\right] = 1; \\ \text{(iii) } s_i \in r_i(\pi_i \,|\, \theta_i) \end{array} \right\}.$$

[17] Battigalli and Siniscalchi (2002, 2007) extend strong rationalizability to dynamic games with incomplete information and provide epistemic characterizations. Piermont and Zuazo-Garin (2026) analyze an interim version under potential model misspecifications and examine its robustness in this more general setting.

Finally, let $\overline{\mathbf{S}}_i(\theta_i) = \bigcap_{n\geq 0} \overline{\mathbf{S}}_i^n(\theta_i)$.

This is an iterative elimination procedure that imposes restrictions on belief revision at unexpected information sets. In particular, at each round n, if an unexpected information set is consistent with some combination of opponents' payoff types and strategy profiles surviving the previous round $\overline{\mathbf{S}}_{-i}^{n-1}$, then player i attaches probability one to $\overline{\mathbf{S}}_{-i}^{n-1}$ (i.e., continues to believe that opponents are rational). This procedure formally captures players' rationality and common *strong* belief in rationality (Battigalli and Siniscalchi, 2002).

In many economic applications, however, we are interested in predictions of strong rationalizability when a small type space is used to represent players' initial beliefs (e.g., a type space with finitely many payoff types). This motivates an interim notion of strong rationalizability for a given type space. The key idea is that players' conjectures should be consistent with the type space at the initial information set, but upon being "surprised," initial beliefs assigned by the type space may be abandoned. Nevertheless, knowledge of the information structure and strong belief in rationality should be maintained.[18] This implies that belief revision must be constrained by the belief-free notion of strong rationalizability introduced above, which records the strategies that are strongly rationalizable for *some* type in the universal type space associated with each payoff type.[19] This leads naturally to the following interim notion of strong rationalizability, denoted by $\mathbf{S}$.

Definition 9 (Strong Rationalizability)**.** Fix a type space $\{(T_i, \vartheta_i, \kappa_i)_{i\in I}\}$. For every $i \in I$ and $t_i \in T_i$, let $\mathbf{S}_i^0(t_i) = \overline{\mathbf{S}}(\vartheta_i(t_i))$. For $n \geq 1$, write $\mathbf{S}_{-i}^{n-1}(t_{-i}) = \times_{j\neq i} \mathbf{S}_j^{n-1}(t_j)$, and define

$$\mathbf{S}_i^n(t_i) = \left\{ s_i \in \overline{\mathbf{S}}_i(\vartheta_i(t_i)) : \begin{array}{l} \exists(\pi_i, \mu_i) \in \Delta^{\mathcal{H}_i}(\vartheta_i(t_i) \times S_{-i}) \times \Delta(\vartheta_i(t_i) \times T_{-i} \times S_{-i}) \text{ s.t.} \\ \text{(i) } \operatorname{marg}_{\vartheta_i(t_i)\times T_{-i}} \mu_i = \kappa_i(t_i); \\ \text{(ii) } \mu_i \left[\left\{(\omega, t_{-i}, s_{-i}) : (t_{-i}, s_{-i}) \in \operatorname{graph}\left(\mathbf{S}_{-i}^{n-1}\right)\right\}\right] = 1; \\ \text{(iii) } \pi_i(\phi) = \operatorname{marg}_{\vartheta_i(t_i)\times S_{-i}} \mu_i; \\ \text{(iv) } s_i \in r_i(\pi_i \,|\, \vartheta_i(t_i)) \end{array} \right\}.$$

Finally, let $\mathbf{S}_i(t_i) = \bigcap_{n\geq 0} \mathbf{S}_i^n(t_i)$. Also, denote $\mathbf{S} = \times_{i\in I} \mathbf{S}_i$.

The definition of $\mathbf{S}$ highlights a conceptual distinction between the information structure and the type space: the type space only serves as a modeling device for the analyst to represent players' initial beliefs, whereas the information structure records hard information that players hold throughout the play. Definitions 8 and 9 reflect precisely this hierarchy between belief and information.

[18] This interpretation is analogous to a setting studied in Battigalli and Siniscalchi (2007), where there is common certainty of a state conditional on the initial information set ϕ alone.

[19] We do not provide a formal proof of this statement here. A related result for belief-free rationalizability in static games is established by Ziegler (2022).

The formulation of $\mathbf{S}$ differs from the one in Piermont and Zuazo-Garin (2026), where the type space encodes not only players' initial beliefs but also their *persistent* beliefs that are maintained throughout. Although our definition of $\mathbf{S}$ involves two steps, it follows a consistent pattern. Much like the notion of $\mathbf{W}^{\mathrm{u}}$ in Remark 2, the refinement is computed in the exact same manner as $\mathbf{W}$ in Definition 3, starting from a particular subset of strategies. Moreover, the procedure is computationally tractable in economic applications, especially when the type space is finite: the iterative procedure requires solving only finitely many optimization problems and avoids working with the infinite-dimensional universal type space. A natural question is whether this refinement of weak rationalizability, which incorporates strong belief, is robust. The following result provides a positive answer.

Proposition 7. *For any given finite type space,* $\mathbf{S}$ *is a robust refinement of* $\mathbf{W}$*.*

A closely related result was first established by Piermont and Zuazo-Garin (2026), who show the local robustness of strong rationalizability to misspecifications of higher-order beliefs. In their paper, strong rationalizability is shown to be an upper hemicontinuous correspondence on the universal type space. Although our proof of Proposition 7 builds on Propositions 3 and 4, one could alternatively argue as follows: first, show that our definition of $\mathbf{S}$ is equivalent to that in Piermont and Zuazo-Garin (2026) when restricted to the universal type space; then, show that $\mathbf{S}$ is type-space invariant (i.e., $\mathbf{S}_i(t_i) = \mathbf{S}_i(\varphi_i^*(t_i))$ for any finite type t_i); and finally, invoke the upper hemicontinuity of $\mathbf{S}$ on the universal type space.

Our proof, however, sheds some light on *why* common strong belief in rationality preserves local robustness. When perturbing beliefs at each order (equation (3)), one can always construct a conjecture consistent with the perturbed hierarchy that also respects common strong belief in rationality. This implies that, for every finite type t_i and every element of the local upper $\mathbf{W}$ collection $\mathcal{R}_i^{\mathrm{loc}}(t_i)$, *some* strategy in that set, as a best response to the constructed conjecture, must also lie in $\mathbf{S}_i(t_i)$. Piermont and Zuazo-Garin (2026) further note that when the information structure is sufficiently rich, strong rationalizability loses its refinement power. Our results echo this observation: when condition (2) of Proposition 5 holds, for every finite type t_i, any proper refinement of $\mathbf{W}_i(t_i)$ fails to be robust; therefore, we must have $\mathbf{S}_i(t_i) = \mathbf{W}_i(t_i)$.[20]

4.2 Privacy of Information

When modeling strategic interactions with incomplete information, we sometimes assume that a party receives a hard piece of information, while the other party is uninformed. In

[20]This equivalence provides a generalization of the insight in Battigalli and Siniscalchi (2007), which focuses on complete-information types in environments with rich payoff uncertainty.

other words, the information is privately learned and exclusive to the owner. For example, in an auction with private values, the seller does not possess any information about bidders' valuations (except the prior distribution); in a signaling game, the characteristic of a sender is private information and unavailable to the receiver. In this subsection, we consider predictions that are robust when we perturb common knowledge about the *privacy* of such information.

Example 2. We use the well-known Beer-Quiche game (Cho and Kreps, 1987) to illustrate how the common knowledge assumption regarding the privacy of information can be relaxed. A sender (player 1) and a receiver (player 2) move sequentially in a two-stage game. The sender's characteristic can be either high (h) or low (ℓ), with ex ante probabilities $\rho \in (0,1)$ and $1-\rho$, respectively. In the original formulation, the sender knows his own characteristic but the receiver does not, and this is common knowledge.

We extend the information structure by allowing for the possibility that the sender's characteristic is not private—that is, the receiver may also observe it. The state space is $\Omega = \{\omega_o^h, \omega_p^h, \omega_o^\ell, \omega_p^\ell\}$, where the superscripts indicate the characteristic of the sender and the subscripts indicate whether the characteristic is observed (o) or private (p). The players' information partitions are given by

$$\Theta_1 = \{\theta_1^h, \theta_1^\ell\}, \text{ where } \theta_1^h = \{\omega_o^h, \omega_p^h\} \text{ and } \theta_1^\ell = \{\omega_o^\ell, \omega_p^\ell\},$$

and

$$\Theta_2 = \{\theta_2^h, \theta_2^\ell, \theta_2^p\}, \text{ where } \theta_2^h = \{\omega_o^h\}, \ \theta_2^\ell = \{\omega_o^\ell\} \text{ and } \theta_2^p = \{\omega_p^h, \omega_p^\ell\}.$$

In words, player 1 observes his own characteristic but does not know whether player 2 observes it, while player 2 either observes player 1's characteristic or remains uninformed. The resulting information structure is illustrated in Figure 2, where player 1's information partition is depicted using solid boundaries and player 2's using dashed ones.

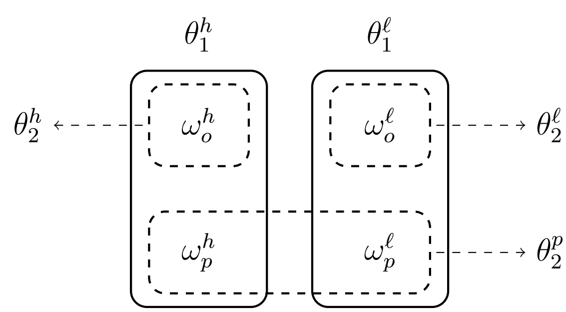


Figure 2: Information structure in Example 2.

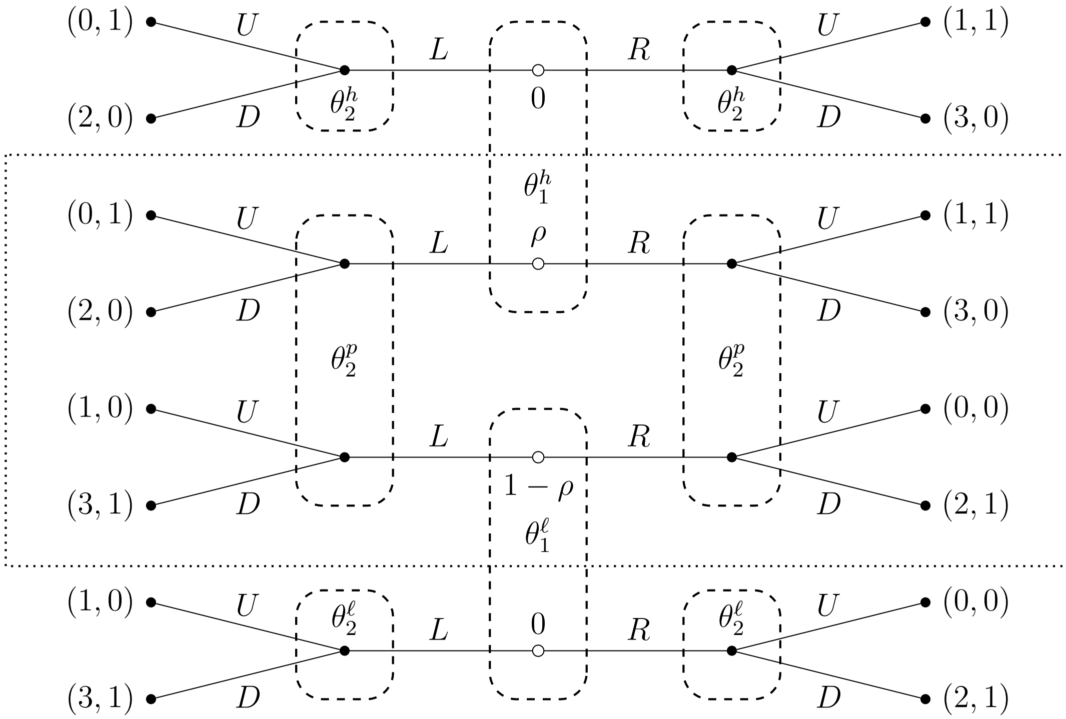


Figure 3: Beer-Quiche game with uncertainty about information privacy.

We can now envision the original formulation as a type space $\hat{\mathcal{T}}$ defined as

$$T_1 = \left\{t_1^h, t_1^\ell\right\}, \ T_2 = \left\{t_2^p\right\},$$
$$\vartheta_1\left(t_1^h\right) = \theta_1^h, \ \vartheta_1\left(t_1^\ell\right) = \theta_1^\ell, \ \vartheta_2\left(t_2^p\right) = \theta_2^p,$$
$$\kappa_1\left(t_1^h\right)\left[t_2^p\right] = 1, \ \kappa_1\left(t_1^\ell\right)\left[t_2^p\right] = 1, \text{ and } \kappa_2\left(t_2^p\right)\left[t_1^h\right] = 1 - \kappa_2\left(t_2^p\right)\left[t_1^\ell\right] = \rho.$$

By employing this type space, the analyst assumes common initial belief in the privacy of sender's information. However, higher-order uncertainty about information privacy is present when types in $\hat{\mathcal{T}}$ are perturbed in the universal type space.

Actions and payoffs are depicted in Figure 3.[21] Intuitively, a sender of characteristic h (ℓ) prefers the message R (L). The receiver's action U reduces the sender's payoff regardless of the sender's characteristic, while the receiver prefers action U only when facing a sender of characteristic ℓ. The dotted rectangle in the center corresponds to the original Beer–Quiche game; the additional branches represent states in which the sender's characteristic is no longer private information.

[21]The figure provides an *ex ante* description of the signaling game, whereas our solution concept is defined at the interim stage. In particular, the dashed rectangles do not represent any information sets $h \in \bigcup_{i\in I} \mathcal{H}_i$ in the extensive form $\mathcal{E}$; rather, they are induced by the information partitions Θ_i in the information structure $\mathcal{I}$.

n	$\overline{\mathcal{R}}_1^{\uparrow,n}(\theta_1^h)$	$\overline{\mathcal{R}}_1^{\uparrow,n}(\theta_1^\ell)$	$\overline{\mathcal{R}}_2^{\uparrow,n}(\theta_2^p)$	$\overline{\mathcal{R}}_2^{\uparrow,n}(\theta_2^h)$	$\overline{\mathcal{R}}_2^{\uparrow,n}(\theta_2^\ell)$
0	S_1	S_1	S_2	S_2	S_2
1	S_1	S_1	$S_2\backslash\{UU\}$, $S_2\backslash\{DD\}$	$\{UU\}$	$\{DD\}$
2	$\{R\}$	$\{L\}$	$S_2\backslash\{UU\}$, $S_2\backslash\{DD\}$	$\{UU\}$	$\{DD\}$
3	$\{R\}$	$\{L\}$	$\{DU\}$	$\{UU\}$	$\{DD\}$
4	$\{L\},\{R\}$	$\{L\}$	$\{DU\}$	$\{UU\}$	$\{DD\}$
5	$\{L\},\{R\}$	$\{L\}$	$\{DU\},\{UU\}$	$\{UU\}$	$\{DD\}$
$\cdots$	$\{L\},\{R\}$	$\{L\}$	$\{DU\},\{UU\}$	$\{UU\}$	$\{DD\}$

Table 1: Upper **W** collections. Only minimal elements are reported in each round.

Weak rationalizability has weak predictive power for types in $\hat{\mathcal{T}}$. Write the strategy of the receiver as a concatenation of actions when facing messages L and R. Only UU (DD) is removed in the first round when $\rho < \frac{1}{2}$ ($\rho > \frac{1}{2}$, respectively), and the iterative procedure converges in the second round. In particular, we have

$$\mathbf{W}_1(t_1^h) = \mathbf{W}_1(t_1^\ell) = \{L, R\}, \text{ and } \mathbf{W}_2(t_2^p) = \begin{cases} \{DU, UD, DD\} \text{ if } \rho < \frac{1}{2} \\ \{DU, UD, DD, UU\} \text{ if } \rho = \frac{1}{2} \\ \{DU, UD, UU\} \text{ if } \rho > \frac{1}{2}. \end{cases}$$

We next show that **W** admits a robust refinement. To this end, we compute the upper **W** collections summarized in Table 1. As before, only minimal elements are listed in each round. By Proposition 6, we know that **W** is generically unique on the universal type space. Moreover, by Remark 2, the strongest robust refinement of **W**, $\mathbf{W}^{\mathrm{u}}$, can be computed by an iterative procedure restricted to uniquely rationalizable strategies, which leads to the following predictions:

- When $\rho < \frac{1}{2}$, $\mathbf{W}_1^{\mathrm{u}}(t_1^h) = \{L\}$, $\mathbf{W}_1^{\mathrm{u}}(t_1^\ell) = \{L\}$, and $\mathbf{W}_2^{\mathrm{u}}(t_2^p) = \{DU\}$;
- When $\rho \geq \frac{1}{2}$, $\mathbf{W}_1^{\mathrm{u}}(t_1^h) = \{L, R\}$, $\mathbf{W}_1^{\mathrm{u}}(t_1^\ell) = \{L\}$, and $\mathbf{W}_2^{\mathrm{u}}(t_2^p) = \{DU, UU\}$.

For any value of ρ, $\mathbf{W}^{\mathrm{u}}$ delivers strictly sharper predictions than **W**. When the ex ante likelihood of the sender having characteristic h is low ($\rho < \frac{1}{2}$), the strongest robust refinement of weak rationalizability predicts that the sender of either characteristic plays L in the first stage, and that the receiver plays DU as a response. Any other behavior would rely critically on the knife-edge assumption that the privacy of the sender's characteristic is a common initial belief. Notably, this prediction coincides with the only sequential equilibrium outcome that satisfies the Intuitive Criterion (Cho and Kreps, 1987). However, our argument is concerned with relaxing common knowledge of information privacy and higher-order uncertainty due to such relaxation, but does not rely on the common belief in a fixed equilibrium outcome (see Cho and Kreps, 1987; Battigalli and Siniscalchi, 2002, 2003).

On the other hand, when the ex ante likelihood of h is high ($\rho \geq \frac{1}{2}$), both messages are rationalizable for the sender of characteristic h, and any proper refinement of this prediction fails to be robust. Nevertheless, the strongest robust refinement still yields sharp predictions for other types: it implies that the sender of characteristic ℓ plays L, and that the receiver responds with U upon receiving the message R. ♢

In Example 2, weak rationalizability is generically unique on the universal type space and therefore admits a strongest robust refinement. In what follows, we establish this result more generally under certain conditions. Consider a finite signaling game. A sender (player 1) chooses a message $m \in M$ from a finite set M in the first stage, and a receiver (player 2) responds with an action $a \in A$ upon observing the message. In our language, the strategy sets of the two players are $S_1 = M$ and $S_2 = A^M$, respectively.

There are $K \in \mathbb{N}$ possible characteristics of the sender. The state space Ω satisfies $|\Omega| = 2K$, and we write

$$\Omega = \{\omega_o^1, \omega_p^1, \omega_o^2, \omega_p^2, \ldots, \omega_o^K, \omega_p^K\}.$$

As before, the superscript o implies that the quality of the sender is observed, while p means the quality is private. The information partitions of the two players are

$$\Theta_1 = \{\theta_1^1, \theta_1^2, \ldots, \theta_1^K\}, \text{ where } \theta_1^k = \{\omega_o^k, \omega_p^k\} \text{ for each } k = 1, 2, \ldots, K,$$

and

$$\Theta_2 = \{\theta_2^1, \theta_2^2, \ldots, \theta_2^K, \theta_2^p\}, \text{ where } \theta_2^k = \{\omega_o^k\} \text{ for each } k = 1, 2, \ldots, K, \text{ and } \theta_2^p = \{\omega_p^1, \omega_p^2, \ldots, \omega_p^K\}.$$

The players' utilities depend on the characteristic of the sender but not on whether this characteristic is observed by the receiver; that is, for each $i = 1, 2$, $u_i(\omega_o^k, m, a) = u_i(\omega_p^k, m, a)$ for all $k = 1, 2, \ldots, K$.

We make the following assumption regarding payoffs in the signaling game.

Assumption 1. (i) For each $\omega \in \Omega$, $m \in M$, and $a, a' \in A$ such that $a \neq a'$, $u_2(\omega, m, a) \neq u_2(\omega, m, a')$. Moreover, for each $\omega \in \Omega$, $m, m' \in M$ such that $m \neq m'$, and $a, a' \in A$, $u_1(\omega, m, a) \neq u_1(\omega, m', a')$.

(ii) Let $a(\omega, m)$ denote the response such that $\{a(\omega, m)\} = \arg\max_{a \in A} u_2(\omega, m, a)$. For each $m \in M$, there is a state $\omega^m \in \Omega$ in which $\arg\max_{m' \in M} u_1(\omega^m, m', a(\omega^m, m'))) = \{m\}$.

Part (i) of Assumption 1 requires that the game has no relevant ties. This condition is slightly weaker than the assumption that payoffs in the signaling game are in "generic position" (see, for example, Battigalli, 1997). Part (ii) is weaker than the standard "richness condition" in the literature (Chen, 2012; Penta, 2012). In particular, it does not require that each message be a dominant strategy for the sender in some state; instead, it only requires that the message be the unique one selected by backward induction under complete information in some state. Clearly, Example 2 satisfies Assumption 1.

Proposition 8. *Suppose the signaling game satisfies Assumption 1. The set of types that have a unique weakly rationalizable strategy is open and dense in* T_i^* *for* $i = 1, 2$.

For any type space that imposes common initial belief in the privacy of information, an analyst can then employ $\mathbf{W}^{\mathrm{u}}$ to deliver a robust prediction. Naturally, the specific messages and responses that stand out under this refinement of weak rationalizability depend on the ex ante likelihood that players assign to different characteristics of the sender (e.g., the parameter ρ in Example 2).

4.3 Observability of Actions

In our analysis so far, we have assumed that players interact according to a commonly known extensive form $\mathcal{E}$. Nonetheless, it is both natural and important to investigate the consequences of perturbing this common knowledge assumption. For example, there may be higher-order uncertainty about whether moves are truly simultaneous or whether a player can distinguish among nodes within a given information set. In other words, the game itself may be subject to *extensive-form uncertainty.* Penta and Zuazo-Garin (2022) formulate and study this local robustness problem in static games.[22] They first characterize the predictions that captures Rationality and Common Belief in Rationality (RCBR) and then introduce a robust refinement that typically delivers sharp predictions when higher-order beliefs about observability of actions are perturbed.

In this section, we illustrate how our results can be applied to generalize their analysis and shed light on the underlying forces driving their findings. In particular, we focus on observability of actions in static games as in Penta and Zuazo-Garin (2022). However, unlike their setting, where actions are either perfectly observed or not observed at all, we allow for arbitrary *partial* observability of earlier moves. To this end, we slightly extend our framework to incorporate extensive-form uncertainty, as described below.

Consider a two-player static game with complete information $\hat{\Gamma}$, referred to as the *base game.* Each player $i = 1, 2$ has a finite set of actions A_i. When $(a_1, a_2) \in A_1 \times A_2$ is played, each player i obtains payoff $\hat{u}_i(a_1, a_2)$. We impose an assumption on payoffs in the base game so that there are no relevant ties.

[22] Earlier work on extensive-form uncertainty includes Robson (1994), Reny and Robson (2004), Kalai (2004), Solan and Yariv (2004), and Zuazo-Garin (2017). In an information design setting, Doval and Ely (2020) and Makris and Renou (2023) study extensive-form robustness from a global perspective.

Assumption 2 (Penta and Zuazo-Garin (2022))**.** For each $i \in \{1, 2\}$, $j \neq i$, and $a_i \in A_i$, there exists a unique $a_j^*(a_i)$ such that $\arg\max_{a_j \in A_j} \hat{u}_j(a_j, a_i) = \{a_j^*(a_i)\}$. Moreover, for each $a_i, a_i' \in A_i$ such that $a_i \neq a_i'$, $\hat{u}_i(a_i, a_j^*(a_i)) \neq \hat{u}_i(a_i', a_j^*(a_i'))$.

Under this assumption, if it is common knowledge that one player moves after perfectly observing the action chosen by the other player, the game admits a unique outcome obtained by backward induction. We denote by a_i^i the unique maximizer of $\hat{u}_i(a_i, a_j^*(a_i))$ and define $a_j^i = a_j^*(a_i^i)$. That is, a_i^i is the *Stackelberg action* chosen by player i when moving first, and a_j^i is player j's best response upon observing a_i^i.

We now introduce a derived game with incomplete information that can accommodate uncertainty about the observability of actions. For each $i = 1, 2$, let Π_i be a (potentially empty) collection of *nontrivial* partitions of the action set A_i, and define the state space as

$$\Omega = \{\omega_0, (\omega_1^{\mathcal{P}_1})_{\mathcal{P}_1 \in \Pi_1}, (\omega_2^{\mathcal{P}_2})_{\mathcal{P}_2 \in \Pi_2}\}.$$

In state ω^0, the game is a genuine simultaneous-move game; in state $\omega_i^{\mathcal{P}_i}$, player i moves first, while player j observes player i's action only up to the partition $\mathcal{P}_i$. The information partition of each player $i = 1, 2$ is

$$\Theta_i = \{\theta_i^0, (\theta_i^{\mathcal{P}_j})_{\mathcal{P}_j \in \Pi_j}\}, \text{ where } \theta_i^0 = \{\omega_0, (\omega_i^{\mathcal{P}_i})_{\mathcal{P}_i \in \Pi_i}\} \text{ and } \theta_i^{\mathcal{P}_j} = \{\omega_j^{\mathcal{P}_j}\} \text{ for all } \mathcal{P}_j \in \Pi_j.$$

In words, player i either is uncertain about whether the opponent can observe his own action, or knows that he can observe the opponent's action according to some partition $\mathcal{P}_j$. The resulting information structure $\mathcal{I} = \{\Omega, (\Theta_1, \Theta_2)\}$ is illustrated in Figure 4.

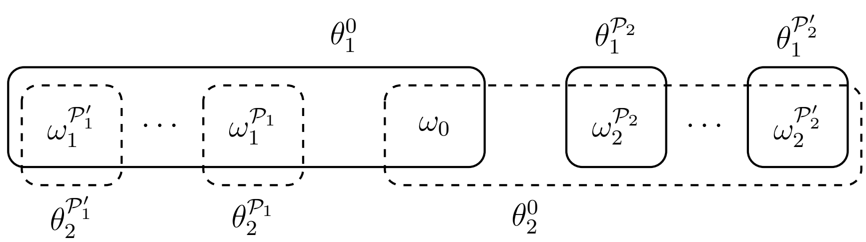


Figure 4: Information partitions with extensive-form uncertainty.

The only departure from our earlier framework is that every state is now associated with a different extensive form. In state ω_0, the extensive form coincides with the base game $\Gamma_0 = \hat{\Gamma}$ where two players move simultaneously. For each $i = 1, 2$ and action partition $\mathcal{P}_i \in \Pi_i$, the extensive form $\Gamma_i^{\mathcal{P}_i}$ corresponding to state $\omega_i^{\mathcal{P}_i}$ is a two-stage dynamic game. In this game, player i moves first by chooses an action in A_i, after which player j selects an action in A_j conditional on the cell of the partition $\mathcal{P}_i$ containing player i's action. That is, each cell of $\mathcal{P}_i$

forms an information set for player j at the second stage. The set of strategies available to player i depends on his payoff type and is defined as

$$S_i(\theta_i) = \begin{cases} A_i \text{ if } \theta_i = \theta_i^0, \\ A_i^{\mathcal{P}_j} \text{ if } \theta_i = \theta_i^{\mathcal{P}_j}, \end{cases}$$

where $A_i^{\mathcal{P}_j}$ denotes the set of mappings from $\mathcal{P}_j$ to A_i. Thus, when player i cannot observe the opponent's action, he simply chooses an action in A_i, whereas when he moves second, his reaction must be measurable with respect to his observation. Let $\mathcal{P}_i(a_i)$ denote the cell in $\mathcal{P}_i$ that contains action a_i. The utility function of player $i = 1, 2$ in the derived game is then given by

$$u_i(\omega, z(s_i, s_j)) = \begin{cases} \hat{u}_i(s_i, s_j) \text{ if } \omega = \omega_0, \\ \hat{u}_i(s_i, s_j(\mathcal{P}_i(s_i))) \text{ if } \omega = \omega_i^{\mathcal{P}_i}, \\ \hat{u}_i(s_i(\mathcal{P}_j(s_j)), s_j) \text{ if } \omega = \omega_j^{\mathcal{P}_j}. \end{cases}$$

Although the extensive form varies with the state of nature, we emphasize that each player *knows* his own set of available strategies. Moreover, throughout the play, reaching any particular information set does not reveal additional information beyond what has been encoded in the information structure. Precisely for this reason, all of our analysis carries over once the strategy sets S_i are replaced by $S_i(\theta_i)$ (or $S_i(\vartheta_i(t_i))$) for a player i with payoff type θ_i (or type t_i, respectively).[23] Therefore, all results apply verbatim in this environment and can be used directly in what follows.

Of particular interest is the case in which the game is truly simultaneous (i.e., $\omega = \omega_0$) and there is common (initial) belief therein. This situation can be modeled by a type space $\mathcal{T}_0 = \{(T_i, \vartheta_i, \kappa_i)_{i=1,2}\}$ defined as

$$T_i = \{t_i^0\},\ \vartheta_i(t_i^0) = \theta_i^0,\ \text{and}\ \kappa_i(t_i^0)[t_{-i}^0] = 1,\ \text{for}\ i = 1, 2.$$

Viewing this type space as embedded in the universal type space generated by $\mathcal{I}$, we may ask which predictions are robust when higher-order beliefs of each t_i^0 about the observability of actions are perturbed.

We focus on a class of games considered in Penta and Zuazo-Garin (2022), which includes, for example, the unanimity games in Harsanyi (1981) and Kalai and Samet (1984).

Proposition 9. *Suppose the base game $\hat{\Gamma}$ satisfies Assumption 2, that both players strictly prefer any Nash equilibrium of $\hat{\Gamma}$ to any non-equilibrium profile, and that they are indifferent over non-equilibrium profiles. Then, given the type space $\mathcal{T}_0$,*

[23]Accordingly, $\theta_i \times S_j$ should be replaced by $\bigcup_{\omega \in \theta_i} \{\omega\} \times S_j(\theta_j(\omega))$, and $\theta_i \times \mathcal{S}_j$ by $\bigcup_{\omega \in \theta_i} \{\omega\} \times \mathcal{S}_j(\theta_j(\omega))$ where $\mathcal{S}_j(\theta_j)$ is the collection of nonempty subsets of $S_j(\theta_j)$.

(i) The prediction P defined by $P(t_i^0) = \{a_i^i, a_i^j\}$ for each $i = 1, 2$ is a robust refinement of weak rationalizability;

(ii) If there exists a partition $\mathcal{P}_i \in \Pi_i$ such that $\{a_i^i\} \in \mathcal{P}_i$ for both $i = 1, 2$, then the prediction P with $P(t_i^0) = \{a_i^i, a_i^j\}$ for each i is the strongest robust refinement of weak rationalizability;

(iii) If there exists a partition $\mathcal{P}_i \in \Pi_i$ such that $\{a_i^i\} \in \mathcal{P}_i$ and $\Pi_j = \emptyset$, then the prediction defined by $P(t_i^0) = \{a_i^i\}$ and $P(t_j^0) = \{a_j^i\}$ is the strongest robust refinement of weak rationalizability.

Part (i) of Proposition 9 shows that, under common belief that the game is static, the prediction that selects, for each player, only the Stackelberg action and the best response to the opponent's Stackelberg action always constitutes a robust refinement of weak rationalizability, *regardless of* the information structure that governs action observability. This result is particularly powerful, as the robust refinement contains only two actions for each player, although the set of weakly rationalizable actions may be arbitrarily large. Part (ii) extends Proposition 2 of Penta and Zuazo-Garin (2022) by demonstrating that the "Stackelberg selections" result holds in a more general environment. Rather than assuming actions are either perfectly observable or not at all, it suffices that the Stackelberg action of each player can be identified. Part (iii) parallels their finding for one-sided uncertainty and shows that, when it is common knowledge that only one player can be the first mover, that player has a *de facto* first-mover advantage: the Stackelberg outcome is uniquely selected and constitutes the strongest robust refinement of $\mathbf{W}$.[24]

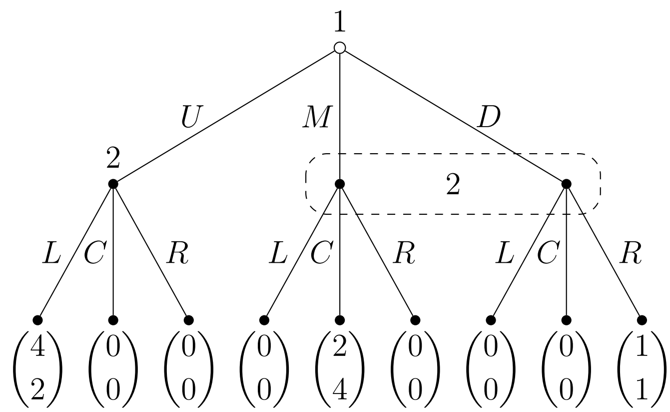
Remark 3. Unlike in Penta and Zuazo-Garin (2022), the results in Proposition 9 do not come from the fact that weak rationalizability is generically unique on the universal type space. This is because, when the follower cannot perfectly observe the action chosen by the first mover, multiple best responses may arise, depending on the beliefs formed under imperfect observation. Nevertheless, these responses still generate the appropriate "seeds" for first-mover behavior, allowing the argument to proceed by invoking Proposition 3.

Example 3. Consider the augmented Battle of the Sexes game depicted in Figure 5a. Clearly, this base game satisfies the assumptions in Proposition 9. By Proposition 9(i), the prediction $\{U, M\}$ for player 1 and $\{L, C\}$ for player 2 constitutes a robust refinement of rationalizability when there is common belief in the base game, regardless of the information structure describing potential observability of actions.

[24] See, for instance, Cooper et al. (1993) for experimental evidence documenting this focal-point effect generated by asynchronous moves.

	L	C	R
U	$4,2$	$0,0$	$0,0$
M	$0,0$	$2,4$	$0,0$
D	$0,0$	$0,0$	$1,1$

(a) The base game



(b) The extensive form in state $\omega_1^{\mathcal{P}_1}$

Figure 5: The two extensive forms in Example 3

Now suppose there are two possible states $\Omega = \{\omega_0, \omega_1^{\mathcal{P}_1}\}$, where $\mathcal{P}_1 = \{\{U\}, \{M, D\}\}$. Thus, either the game is truly simultaneous, or player 1 moves first and player 2 observes player 1's action only up to the partition $\mathcal{P}_1$: she can identify U but cannot distinguish between M and D. The extensive form associated with state $\omega_1^{\mathcal{P}_1}$ is depicted in Figure 5b. Player 1 has a single payoff type, $\theta_1^0 = \{\omega_0, \omega_1^{\mathcal{P}_1}\}$, and hence cannot distinguish between the two extensive forms in Figure 5. Player 2 has two payoff types, $\theta_2^0 = \{\omega_0\}$ and $\theta_2^{\mathcal{P}_1} = \{\omega_1^{\mathcal{P}_1}\}$: the former knows that the interaction is the simultaneous-move base game, whereas the latter knows the game is dynamic with partial observation. Applying Proposition 9(iii), the prediction that player 1 plays U and player 2 responds with L delivers the strongest robust refinement of rationalizability under the assumption of common belief in the base game. $\diamondsuit$

5 Conclusion

In this paper, we study weak rationalizability and its refinements in dynamic games with general payoff uncertainty. A central feature of our analysis is the distinction between the information structure, which captures commonly known features of payoffs and the extensive form, and the type space, which serves as a modeling device to represent players' initial beliefs. We investigate the local robustness of predictions to perturbations of higher-order beliefs in the universal type space generated by a given information structure, thereby restricting the admissible directions of belief perturbations.

Using a collection-based approach, we characterize the strategies that can be uniquely selected and identify refinements of weak rationalizability that remain robust to misspecifications of higher-order beliefs around finite types. We further use these results to provide necessary and sufficient conditions on the information structure under which the Structure

Theorem and generic uniqueness hold for weak rationalizability.

We apply our framework to show that strong rationalizability—a solution concept that incorporates forward-induction reasoning that has received considerable attention in the literature—constitutes a robust refinement of weak rationalizability. We also demonstrate how to apply our results in two economic applications in which common knowledge assumptions about either the information structure or the extensive form are perturbed. In contrast to the work in the literature where *all* common knowledge assumptions are relaxed by a richness condition, our analysis shows that relaxing common knowledge along economically meaningful dimensions can *sharpen* predictions through local robustness.

We view the applications studied here primarily as illustrations of a broader methodological toolbox. The framework developed in this paper opens the door to studying local robustness in richer environments, such as settings in which players can acquire new information during play or face uncertainty about the observability of actions in genuinely dynamic interactions. The relevance of these questions ultimately depends on the empirical importance of relaxing common knowledge assumptions in specific economic environments, and exploring these directions remains an important avenue for future research.

A Appendix: Omitted Proofs

A.1 Proofs for Section 3: Characterizations

A.1.1 Proof of Proposition 1

We first prove a lemma.

Lemma 3. *For any $n \geq 0$, we have*

$$\overline{\mathcal{R}}_i^{\uparrow,n}(\theta_i) = \left\{R_i \in 2^{S_i} : \exists t_i \in T_i^* \text{ s.t. } \vartheta_i^*(t_i) = \theta_i \text{ and } R_i \supseteq \mathbf{W}_i^n(t_i)\right\},$$

and

$$\overline{\mathcal{R}}_i^{\downarrow,n}(\theta_i) = \left\{R_i \in 2^{S_i} : \exists t_i \in T_i^* \text{ s.t. } \vartheta_i^*(t_i) = \theta_i \text{ and } R_i \subseteq \mathbf{W}_i^n(t_i)\right\},$$

Proof. We focus on proving the first equality. The proof for the second one is analogous, and hence we omit. We prove by induction. Since $\mathbf{W}_i^0(\cdot) = S_i$ and $\overline{\mathcal{R}}_i^{\uparrow,0}(\cdot) = \{S_i\}$ by definition, the case for $n = 0$ is obvious. Suppose the equality holds for $n-1$.

For the "$\subseteq$" direction, take any $R_i \in \overline{\mathcal{R}}_i^{\uparrow,n}(\theta_i)$. Therefore, there exists $\overline{\nu}_i \in \Delta(\theta_i \times 2^{S_{-i}})$ such that $\overline{\nu}_i\left[\left\{(\omega, R_{-i}) : (\theta_{-i}(\omega), R_{-i}) \in \text{graph}\left(\overline{\mathcal{R}}_{-i}^{\uparrow,n-1}\right)\right\}\right] = 1$ and $R_i \supseteq \bigcup_{\pi_i(\phi) \in \mathrm{C}_i(\overline{\nu}_i \,|\, \theta_i)} r_i(\pi_i \,|\, \theta_i)$. By the induction hypothesis, for every $R_{-i} \in \overline{\mathcal{R}}_{-i}^{\uparrow,n-1}(\theta_{-i})$, there exists $t_{-i}^{(\theta_{-i},R_{-i})} \in T_{-i}^*$ such that $\vartheta^*_{-i}\big(t_{-i}^{(\theta_{-i},R_{-i})}\big) = \theta_{-i}$ and $R_{-i} \supseteq \mathbf{W}^{n-1}_{-i}\big(t_{-i}^{(\theta_{-i},R_{-i})}\big)$. Define a type $t_i \in T_i^*$ by letting $\vartheta_i^*(t_i) = \theta_i$ and

$$\kappa_i^*(t_i)[\omega, t_{-i}] = \overline{\nu}_i\left[\left\{(\omega, R_{-i}) : t_{-i}^{(\theta_{-i}(\omega),R_{-i})} = t_{-i}\right\}\right].$$

This type is well-defined because $\kappa_i^*(t_i)\big[\{(\omega, t_{-i}) : \omega \in \bigcap_{j\neq i}\vartheta_j(t_j)\}\big] = 1$. We next show that $R_i \supseteq \mathbf{W}_i^n(t_i)$. Take any $s_i \in \mathbf{W}_i^n(t_i)$. By Definition 3, this implies that for some pair $(\pi_i, \mu_i) \in \Delta^{\mathcal{H}_i}(\vartheta_i(t_i) \times S_{-i}) \times \Delta(\vartheta_i(t_i) \times T^*_{-i} \times S_{-i})$, we have

(i) $\operatorname{marg}_{\vartheta_i(t_i)\times T^*_{-i}} \mu_i = \kappa_i^*(t_i)$;

(ii) $\mu_i\left[\{(\omega, t_{-i}, s_{-i}) : (t_{-i}, s_{-i}) \in \operatorname{graph}\left(\mathbf{W}^{n-1}_{-i}\right)\}\right] = 1$;

(iii) $\pi_i(\phi) = \operatorname{marg}_{\vartheta_i(t_i)\times S_{-i}} \mu_i$,

(iv) $s_i \in r_i(\pi_i \,|\, \vartheta_i(t_i))$.

We now show that $\pi_i(\phi) \in \mathrm{C}_i(\overline{\nu}_i \,|\, \theta_i)$. Define f_i as the conditional probability of μ_i on each $\big(\omega, t_{-i}^{(\theta_{-i},R_{-i})}\big)$, i.e.

$$f_i(\omega, R_{-i})[s_{-i}] = \mu_i\left[s_{-i} \,|\, \omega, t_{-i}^{(\theta_{-i},R_{-i})}\right].$$

Condition (i) in Definition 6 is satisfied because $\mu_i\left[s_{-i} \in \mathbf{W}^{n-1}_{-i}(t_{-i})\right] = 1$ and $R_{-i} \supseteq \mathbf{W}^{n-1}_{-i}\big(t_{-i}^{(\theta_{-i},R_{-i})}\big)$ for every $R_{-i} \in \overline{\mathcal{R}}^{\uparrow,n-1}_{-i}(\theta_{-i})$ by the induction hypothesis. For condition (ii), the initial conjecture $\pi_i(\phi)$ is consistent with $\overline{\nu}_i$ because

$$\begin{aligned}
\pi_i(\phi)[\omega, s_{-i}] &= \sum_{t_{-i}:\omega\in\vartheta^*_{-i}(t_{-i})} \kappa_i^*(t_i)[\omega, t_{-i}]\mu_i[s_{-i} \,|\, \omega, t_{-i}] \\
&= \sum_{R_{-i}\in 2^{S_{-i}}} \kappa_i^*(t_i)\left[\omega, t_{-i}^{(\theta_{-i},R_{-i})}\right] \mu_i\left[s_{-i} \,|\, \omega, t_{-i}^{(\theta_{-i},R_{-i})}\right] \\
&= \sum_{R_{-i}\in 2^{S_{-i}}} \nu_i[\omega, R_{-i}] f_i(\omega, R_{-i})[s_{-i}].
\end{aligned}$$

Thus, we have $\pi_i(\phi) \in \mathrm{C}_i(\overline{\nu}_i \,|\, \theta_i)$, which implies $s_i \in r_i(\pi_i \,|\, \theta_i) \subseteq R_i$. We can conclude that $R_i \supseteq \mathbf{W}_i^n(t_i)$.

For the converse direction "$\supseteq$", take $R_i \in 2^{S_i}$ such that there is a type $t_i \in T_i^*$ for which $\vartheta_i^*(t_i) = \theta_i$ and $R_i \supseteq \mathbf{W}_i^n(t_i)$. Define $\overline{\nu}_i \in \Delta(\theta_i \times 2^{S_{-i}})$ as

$$\overline{\nu}_i[\omega, R_{-i}] = \kappa_i^*(t_i)\left[(\omega, t_{-i}) : \mathbf{W}^{n-1}_{-i}(t_{-i}) = R_{-i}\right].$$

By the induction hypothesis, we have $\overline{\nu}_i\big[(\omega, R_{-i}) : (\theta_{-i}(\omega), R_{-i}) \in \operatorname{graph}\big(\overline{\mathcal{R}}^{\uparrow,n-1}_{-i}\big)\big] = 1$. It is left to show that for any $\pi_i(\phi) \in \mathrm{C}_i(\overline{\nu}_i \,|\, \theta_i)$, we have $R_i \supseteq r_i(\pi_i \,|\, \theta_i)$. Suppose $\pi_i(\phi) \in \mathrm{C}_i(\overline{\nu}_i \,|\, \theta_i)$, so there exists a function $f_i : \theta_i \times 2^{S_{-i}} \to \Delta(S_{-i})$ such that

(i) $f_i(\omega, R_{-i})[R_{-i}] = 1$,

(ii) $\pi_i(\phi)[\omega, s_{-i}] = \sum_{R_{-i}\in 2^{S_{-i}}} \overline{\nu}_i[\omega, R_{-i}] f_i(\omega, R_{-i})[s_{-i}]$.

Let $\mu_i \in \Delta(\theta_i \times T^*_{-i} \times S_{-i})$ be a distribution such that

$$\mu_i[\omega, t_{-i}, s_{-i}] = \kappa^*_i(t_i)[\omega, t_{-i}] f_i(\omega, \mathbf{W}^{n-1}_{-i}(t_{-i}))[s_{-i}].$$

It follows that $\text{marg}_{\theta_i\times T^*_{-i}} \mu_i = \kappa^*_i(t_i)$ and $\mu_i\big[\{(\omega, t_{-i}, s_{-i}) : (t_{-i}, s_{-i}) \in \text{graph}\left(\mathbf{W}^{n-1}_{-i}\right)\}\big] = 1$. Moreover, we have

$$\begin{aligned}
&\sum_{t_{-i}\in T^*_{-i}} \mu_i[\omega, t_{-i}, s_{-i}] \\
&= \sum_{R_{-i}\in 2^{S_{-i}}} \kappa^*_i(t_i)\left[(\omega, t_{-i}) : \mathbf{W}^{n-1}_{-i}(t_{-i}) = R_{-i}\right] f_i(\omega, R_{-i})[s_{-i}] \\
&= \sum_{R_{-i}\in 2^{S_{-i}}} \overline{\nu}_i(\omega, R_{-i}) f_i(\omega, R_{-i})[s_{-i}] \\
&= \pi_i(\phi)[\omega, s_{-i}].
\end{aligned}$$

Therefore, any $s_i \in r_i(\pi_i \,|\, \theta_i)$ would satisfy $s_i \in \mathbf{W}^n_i(t_i)$. Since $R_i \supseteq \mathbf{W}^n_i(t_i)$, we conclude that $R_i \supseteq r_i(\pi_i \,|\, \theta_i)$ for arbitrary $\pi_i(\phi) \in \text{C}_i(\overline{\nu}_i \,|\, \theta_i)$. Hence, $R_i \in \overline{\mathcal{R}}^{\uparrow,n}_i(\theta_i)$. □

We now turn to the proof of Proposition 1.

Proof of Proposition 1. (i) $\overline{\mathcal{R}}^{\uparrow}_i(\theta_i) = \bigcup_{n\geq 0} \overline{\mathcal{R}}^{\uparrow,n}_i(\theta_i)$: For "$\subseteq$", let $R_i \in \overline{\mathcal{R}}^{\uparrow}_i(\theta_i)$. Then there exist $t_i \in T^*_i$ and $m \in \mathbb{N}$ such that $\vartheta^*_i(t_i) = \theta_i$ and $R_i \supseteq \mathbf{W}_i(t_i) = \mathbf{W}^m_i(t_i)$. By Lemma 3, we have $R_i \in \overline{\mathcal{R}}^{\uparrow,m}_i(\theta_i)$. For "$\supseteq$", take $R_i \in \overline{\mathcal{R}}^{\uparrow,n}_i(\theta_i)$ for some n. By Lemma 3, there exists $t_i \in T^*_i$ such that $\vartheta^*_i(t_i) = \theta_i$ and $R_i \supseteq \mathbf{W}^n_i(t_i) \supseteq \mathbf{W}_i(t_i)$. Hence, $R_i \in \overline{\mathcal{R}}^{\uparrow}_i(\theta_i)$.

(ii) $\overline{\mathcal{R}}^{\downarrow}_i(\theta_i) = \bigcap_{n\geq 0} \overline{\mathcal{R}}^{\downarrow,n}_i(\theta_i)$: For "$\subseteq$", let $R_i \in \overline{\mathcal{R}}^{\downarrow}_i(\theta_i)$. Then there exists $t_i \in T^*_i$ such that $\vartheta^*_i(t_i) = \theta_i$ and $R_i \subseteq \mathbf{W}_i(t_i) \subseteq \mathbf{W}^n_i(t_i)$ for all n. By Lemma 3, we have $R_i \in \overline{\mathcal{R}}^{\downarrow,n}_i(\theta_i)$ for all n. For "$\supseteq$", take $R_i \in \bigcap_{n\geq 0} \overline{\mathcal{R}}^{\downarrow,n}_i(\theta_i)$. By Lemma 3, there exists a sequence $\{t_{i,n}\} \subseteq T^*_i$ such that $\vartheta^*_i(t_{i,n}) = \theta_i$ and $R_i \subseteq \mathbf{W}^n_i(t_{i,n})$ for each n. Since T^*_i is a compact metric space, there exists a convergent subsequence $\{t_{i,n_k}\}$, and let t_i denote its limit. Fix any n, and therefore for all $n_k \geq n$, we have $R_i \subseteq \mathbf{W}^{n_k}_i(t_{i,n_k}) \subseteq \mathbf{W}^n_i(t_{i,n_k})$. Since $\mathbf{W}^n_i$ is upper hemicontinuous for each n (Lemma 2) and $t_{i,n_k} \to t_i$, we have $R_i \subseteq \mathbf{W}^n_i(t_i)$. This is true for all n, and thus we conclude that $R_i \subseteq \mathbf{W}_i(t_i)$, which implies $R_i \in \overline{\mathcal{R}}^{\downarrow}_i(\theta_i)$. □

A.1.2 Proof of Proposition 2

By Definition 4 and the definition of $\mathcal{R}^{\text{loc}}_i(t_i)$.

A.1.3 Proof of Proposition 3

Since each $\mathbf{W}_i(\cdot)$ is an upper hemicontinuous correspondence on T_i^* (Lemma 2), the proof of Proposition 1 from Chen et al. (2022) can be directly replicated to show the following lemma.

Lemma 4. *Given a finite type space $\{(T_i, \vartheta_i, \kappa_i)_{i\in I}\}$. A prediction P is a robust refinement of $\mathbf{W}$ if and only if for every $i \in I$ and $t_i \in T_i$, there exists an open neighborhood $E_{t_i} \subseteq T_i^*$ of t_i such that $P_i(t_i) \cap \mathbf{W}_i(t_i^*) \neq \varnothing$ for every $t_i^* \in E_{t_i}$.*

We now use this lemma to prove Proposition 3. For "$\Rightarrow$", suppose P is robust. Fix any $i \in I$ and $t_i \in T_i$. By definition, for any $R_i \in \mathcal{R}_i^{\text{loc}}(t_i)$, there exists a sequence $\{t_{i,m}\}_{m\in\mathbb{N}} \subseteq T_i^*$ such that $t_{i,m} \to t_i$ and $R_i \supseteq \mathbf{W}_i(t_{i,m})$ for all m. Since P is a robust refinement, by Lemma 4, $P_i(t_i) \cap \mathbf{W}_i(t_{i,m}) \neq \varnothing$ for sufficiently large m, which means $P_i(t_i) \cap R_i \neq \varnothing$.

For the converse "$\Leftarrow$", suppose P is not a robust refinement. Then by Lemma 4, there exists a type $t_i \in T_i$ and a sequence $\{t_{i,m_k}\}_{k\in\mathbb{N}} \subseteq T_i^*$ such that $t_{i,m_k} \to t_i$ and $P_i(t_i) \cap \mathbf{W}_i(t_{i,m_k}) = \varnothing$ for all k. This means $\bigcup_{k\geq 0} (P_i(t_i) \cap \mathbf{W}_i(t_{i,m_k})) = P_i(t_i) \cap \left(\bigcup_{k\geq 0} \mathbf{W}_i(t_{i,m_k})\right) = \varnothing$. But $\bigcup_{k\geq 0} \mathbf{W}_i(t_{i,m_k}) \in \mathcal{R}_i^{\text{loc}}(t_i)$ by definition.

A.1.4 Proof of Proposition 4

Fixing a type $t_i \in T_i$ in the finite type space $\{(T_i, \vartheta_i, \kappa_i)_{i\in I}\}$, we split the statement of Proposition 4 into two lemmas and prove them in order.

Lemma 5. *$\mathcal{R}_i^{\text{loc}}(t_i) \supseteq \bigcap_{n\geq 0} \mathcal{R}_i^{\text{loc},n}(t_i)$.*

Proof. Let $\tilde{\mathcal{R}}_i^{\text{loc},0}(t_i) = \overline{\mathcal{R}}_i^{\uparrow}(\vartheta_i(t_i))$, and define for each n,

$$\tilde{\mathcal{R}}_i^{\text{loc},n}(t_i) = \left\{ R_i \in \overline{\mathcal{R}}_i^{\uparrow}(\vartheta_i(t_i)) : \begin{array}{l} \exists \{t_{i,m}\}_{m\in\mathbb{N}} \subset T_i^* \text{ s.t.} \\ \text{(i) } \vartheta_i^*(t_{i,m}) = \vartheta_i(t_i)\ \forall m, \text{ and } \tau_i^n(t_{i,m}) \to \tau_i^n(t_i) \text{ as } m \to \infty \\ \text{(ii) } R_i \supseteq \mathbf{W}_i(t_{i,m})\ \forall m \end{array} \right\}.$$

Recall that $\tau_i^n(t_i)$ denotes the n-th order belief induced by type t_i. It suffices to show that $\tilde{\mathcal{R}}_i^{\text{loc},n}(t_i) \supseteq \mathcal{R}_i^{\text{loc},n}(t_i)$ for all n, and then $\mathcal{R}_i^{\text{loc}}(t_i) \supseteq \bigcap_{n\geq 0} \mathcal{R}_i^{\text{loc},n}(t_i)$ is implied by taking a diagonal sequence.

We prove by induction. For $n = 0$, we have $\tilde{\mathcal{R}}_i^{\text{loc},0}(t_i) = \mathcal{R}_i^{\uparrow}(\vartheta_i(t_i)) \supseteq \mathcal{R}_i^{\text{loc},0}(t_i)$ by definition. Suppose the inclusion holds for $n-1$; i.e., $\tilde{\mathcal{R}}_i^{\text{loc},n-1}(t_i) \supseteq \mathcal{R}_i^{\text{loc},n-1}(t_i)$ for all $i \in I$ and all $t_i \in T_i$. Take $R_i \in \mathcal{R}_i^{\text{loc},n}(t_i)$, and we need to show that $R_i \in \tilde{\mathcal{R}}_i^{\text{loc},n}(t_i)$. By definition of $\mathcal{R}_i^{\text{loc},n}(t_i)$, for each $m \in \mathbb{N}$, there exists $(\nu_{i,m}, \overline{\nu}_{i,m}) \in \Delta\big(\vartheta_i(t_i) \times T_{-i} \times 2^{S_{-i}}\big) \times \Delta\big(\vartheta_i(t_i) \times 2^{S_{-i}}\big)$ such that

(i) $\text{marg}_{\vartheta_i(t_i)\times T_{-i}} \nu_{i,m} = \kappa_i(t_i)$;

(ii) $\nu_{i,m}\big[\big\{(\omega,t_{-i},R_{-i}):(t_{-i},R_{-i})\in\text{graph}\,\big(\mathcal{R}^{\text{loc},n-1}_{-i}(t_{-i})\big)\big\}\big]=1$;

(iii) $\overline{\nu}_{i,m}\big[\big\{(\omega,R_{-i}):(\theta_{-i}(\omega),R_{-i})\in\text{graph}\,\big(\overline{\mathcal{R}}^{\uparrow}_{-i}\big)\big\}\big]=1$;

(iv) $R_i\supseteq\bigcup_{\pi_i(\phi)\in \mathrm{C}_i(\nu_i,\overline{\nu}_i,\frac{1}{m+1}\,|\,t_i)} r_i(\pi_i\,|\,\vartheta_i(t_i))$.

For each $R_{-i}\in\overline{\mathcal{R}}^{\uparrow}_{-i}(\theta_{-i})$, there exists $y^{(\theta_{-i},R_{-i})}_{-i}\in T^*_{-i}$ such that $\vartheta^*_{-i}\big(y^{(\theta_{-i},R_{-i})}_{-i}\big)=\theta_{-i}$ and $R_{-i}\supseteq\mathbf{W}_{-i}\big(y^{(\theta_{-i},R_{-i})}_{-i}\big)$. For each $t_{-i}\in T_{-i}$ and $R_{-i}\in\mathcal{R}^{\text{loc},n-1}_{-i}(t_{-i})$, by the induction hypothesis, there exists a sequence $\big\{y^{(t_{-i},R_{-i})}_{-i,m}\big\}_{m\in\mathbb{N}}\subseteq T^*_{-i}$ such that $\vartheta^*_{-i}\big(y^{(t_{-i},R_{-i})}_{-i,m}\big)=\vartheta_{-i}(t_{-i})$ for all m, $\tau^{n-1}_{-i}\big(y^{(t_{-i},R_{-i})}_{-i,m}\big)\to\tau^{n-1}_{-i}(t_{-i})$ as $m\to\infty$, and $R_{-i}\supseteq\mathbf{W}_{-i}\big(y^{(t_{-i},R_{-i})}_{-i,m}\big)$ for all m. We now define, for each m, a type $t_{i,m}\in T^*_i$ such that $\vartheta^*_i(t_{i,m})=\vartheta_i(t_i)$, and

$$\begin{aligned}\kappa^*_i(t_{i,m})[\omega,y_{-i}]=&\frac{m}{m+1}\nu_{i,m}\left[\left\{(\omega,t_{-i},R_{-i}):y^{(t_{-i},R_{-i})}_{-i,m}=y_{-i}\right\}\right]\\&+\frac{1}{m+1}\overline{\nu}_{i,m}\left[\left\{(\omega,R_{-i}):y^{(\theta_{-i}(\omega),R_{-i})}_{-i}=y_{-i}\right\}\right].\end{aligned}\tag{4}$$

for every $(\omega,y_{-i})\in\Omega\times T^*_{-i}$. Notice that each $\kappa^*_i(t_{i,m})$ has finite support and satisfies $\kappa^*_i(t_{i,m})\big[\big\{(\omega,t_{-i}):\omega\in\bigcap_{j\neq i}\vartheta_j(t_j)\big\}\big]=1$. Since $\tau^{n-1}_{-i}\big(y^{(t_{-i},R_{-i})}_{-i,m}\big)\to\tau^{n-1}_{-i}(t_{-i})$ as $m\to\infty$ and $\text{marg}_{\vartheta_i(t_i)\times T_{-i}}\,\nu_{i,m}=\kappa_i(t_i)$ for each m, we have $\tau^n_i(t_{i,m})\to\tau^n_i(t_i)$ as $m\to\infty$.

We need to show that $R_i\supseteq\mathbf{W}_i(t_{i,m})$ for each m. By Definition 3 of weak rationalizability, for each $s_i\in\mathbf{W}_i(t_{i,m})$, there exist a distribution $\mu_i\in\Delta(\vartheta_i(t_i)\times T^*_{-i}\times S_{-i})$ and a conjecture $\pi_i\in\Delta^{\mathcal{H}_i}(\vartheta_i(t_i)\times S_{-i})$ such that

(i) $\text{marg}_{\vartheta_i(t_i)\times T^*_{-i}}\,\mu_i=\kappa^*_i(t_{i,m})$;

(ii) $\mu_i\left[\big\{(\omega,t_{-i},s_{-i}):(t_{-i},s_{-i})\in\text{graph}\,\big(\mathbf{W}^{n-1}_{-i}\big)\big\}\right]=1$;

(iii) $\pi_i(\phi)=\text{marg}_{\vartheta_i(t_i)\times S_{-i}}\,\mu_i$;

(iv) $s_i\in r_i(\pi_i\,|\,\vartheta_i(t_i))$.

For $\nu_{i,m}\in\Delta\big(\vartheta_i(t_i)\times T_{-i}\times 2^{S_{-i}}\big)$, define a function $f_i:\vartheta_i(t_i)\times T_{-i}\times 2^{S_{-i}}\to\Delta(S_{-i})$ as follows: If (t_{-i},R_{-i}) are such that $y^{(t_{-i},R_{-i})}_{-i,m}=y_{-i}$, then let $f_i(\omega,t_{-i},R_{-i})[s_{-i}]=\mu_i[s_{-i}\,|\,\omega,y_{-i}]$. Because $R_{-i}\supseteq\mathbf{W}_{-i}\big(y^{(t_{-i},R_{-i})}_{-i,m}\big)$, we have $f_i(\omega,t_{-i},R_{-i})[R_{-i}]=1$. Then we define

$$\begin{aligned}\lambda_i[\omega,s_{-i}]&=\sum_{t_{-i}:\omega\in\vartheta_{-i}(t_{-i})}\ \sum_{R_{-i}\in 2^{S_{-i}}}\nu_{i,m}[\omega,t_{-i},R_{-i}]f_i(\omega,t_{-i},R_{-i})[s_{-i}]\\&=\sum_{y_{-i}:\omega\in\vartheta^*_{-i}(y_{-i})}\nu_{i,m}\left[\left\{(\omega,t_{-i},R_{-i}):y^{(t_{-i},R_{-i})}_{-i,m}=y_{-i}\right\}\right]\mu_i[s_{-i}\,|\,\omega,y_{-i}].\end{aligned}\tag{5}$$

Similarly, for $\overline{\nu}_i\in\Delta\big(\vartheta_i(t_i)\times 2^{S_{-i}}\big)$, define a function $\bar{f}_i:\vartheta_i(t_i)\times 2^{S_{-i}}\to\Delta(S_{-i})$ as follows: If $(\theta_{-i}(\omega),R_{-i})$ is such that $y^{(\theta_{-i}(\omega),R_{-i})}_{-i}=y_{-i}$, then let $\bar{f}_i(\omega,R_{-i})[s_{-i}]=\mu_i[s_{-i}\,|\,\omega,y_{-i}]$.

Because $R_{-i} \supseteq \mathbf{W}_{-i}\big(y_{-i}^{(\theta_{-i}(\omega),R_{-i})}\big)$, we have $\bar{f}_i(\omega, R_{-i})[R_{-i}] = 1$. Then we define

$$\begin{aligned}\lambda_i'[\omega, s_{-i}] &= \sum_{R_{-i}\in 2^{S_{-i}}} \overline{\nu}_i[\omega, R_{-i}]\bar{f}_i(\omega, R_{-i})[s_{-i}] \\ &= \sum_{y_{-i}:\omega\in\vartheta_{-i}^*(y_{-i})} \overline{\nu}_{i,m}\left[\left\{(\omega, R_{-i}) : y_{-i}^{(\theta_{-i}(\omega),R_{-i})} = y_{-i}\right\}\right]\mu_i[s_{-i}\,|\,\omega, y_{-i}]. \qquad (6)\end{aligned}$$

It is left to show that $\pi_i(\phi) \in \mathrm{C}_i(\nu_i, \overline{\nu}_i, \frac{1}{m+1}\,|\,t_i)$. By construction, $\lambda_i, \lambda_i' \in \Delta(\vartheta_i(t_i) \times S_{-i})$ are consistent with ν_i and $\overline{\nu}_i$, respectively. Moreover, we have

$$\begin{aligned}\pi_i(\phi)[\omega, s_{-i}] &= \sum_{y_{-i}:\omega\in\vartheta_{-i}^*(y_{-i})} \mu_i[\omega, y_{-i}, s_{-i}] \\ &= \sum_{y_{-i}:\omega\in\vartheta_{-i}^*(y_{-i})} \kappa_i^*(t_{i,m})[\omega, y_{-i}]\mu_i[s_{-i}\,|\,\omega, y_{-i}] \\ &= \frac{m}{m+1}\sum_{y_{-i}:\omega\in\vartheta_{-i}^*(y_{-i})} \nu_{i,m}\left[\left\{(\omega, t_{-i}, R_{-i}) : y_{-i,m}^{(t_{-i},R_{-i})} = y_{-i}\right\}\right]\mu_i[s_{-i}\,|\,\omega, y_{-i}] \\ &\quad + \frac{1}{m+1}\sum_{y_{-i}:\omega\in\vartheta_{-i}^*(y_{-i})} \overline{\nu}_{i,m}\left[\left\{(\omega, R_{-i}) : y_{-i}^{(\theta_{-i},R_{-i})} = y_{-i}\right\}\right]\mu_i[s_{-i}\,|\,\omega, y_{-i}] \\ &= \frac{m}{m+1}\lambda_i[\omega, s_{-i}] + \frac{1}{m+1}\lambda_i'[\omega, s_{-i}].\end{aligned}$$

The second equality is by the fact that $\mathrm{marg}_{\vartheta_i(t_i)\times T_{-i}^*}\,\mu_i = \kappa_i^*(t_{i,m})$; the third equality is by equation (4); and the last one is due to equations (5) and (6).

Therefore, we have $\pi_i(\phi) \in \mathrm{C}_i(\nu_i, \overline{\nu}_i, \frac{1}{m+1}\,|\,t_i)$. But then $R_i \supseteq r_i(\pi_i\,|\,\vartheta_i(t_i))$, which means $s_i \in R_i$. Since $s_i \in \mathbf{W}_i(t_{i,m})$ is arbitrary, we have $R_i \supseteq \mathbf{W}_i(t_{i,m})$, completing the proof of this lemma. □

Lemma 6. $\mathcal{R}_i^{\mathrm{loc}}(t_i) \subseteq \bigcap_{n\geq 0} \mathcal{R}_i^{\mathrm{loc},n}(t_i)$.

Proof. We prove this by showing that the profile $\big\{\big(\mathcal{R}_i^{\mathrm{loc}}(t_i)\big)_{t_i\in T_i}\big\}_{i\in I}$ survives each round of definition (3). Since every $R_i \in \mathcal{R}_i^{\mathrm{loc}}(t_i)$ satisfies $R_i \in \overline{\mathcal{R}}_i^{\uparrow}(\vartheta_i(t_i))$ by definition, we have $\mathcal{R}_i^{\mathrm{loc}}(t_i) \subseteq \mathcal{R}_i^{\mathrm{loc},0}(t_i)$. Now suppose $\mathcal{R}_i^{\mathrm{loc}}(t_i) \subseteq \mathcal{R}_i^{\mathrm{loc},n-1}(t_i)$ for all $i \in I$ and $t_i \in T_i$. We want to show that, if $R_i \in \mathcal{R}_i^{\mathrm{loc}}(t_i)$, then for any $\varepsilon \in (0, 1]$, there exists a pair $(\nu_i, \overline{\nu}_i) \in \Delta\big(\vartheta_i(t_i) \times T_{-i} \times 2^{S_{-i}}\big) \times \Delta\big(\vartheta_i(t_i) \times 2^{S_{-i}}\big)$ such that the following are true:

(i) $\mathrm{marg}_{\vartheta_i(t_i)\times T_{-i}}\,\nu_i = \kappa_i(t_i)$;

(ii) $\nu_i\big[\big\{(\omega, t_{-i}, R_{-i}) : (t_{-i}, R_{-i}) \in \mathrm{graph}\,\big(\mathcal{R}_{-i}^{\mathrm{loc},n-1}(t_{-i})\big)\big\}\big] = 1$;

(iii) $\overline{\nu}_i\big[\big\{(\omega, R_{-i}) : (\theta_{-i}(\omega), R_{-i}) \in \mathrm{graph}\,\big(\overline{\mathcal{R}}_{-i}^{\uparrow}\big)\big\}\big] = 1$;

(iv) $R_i \supseteq \bigcup_{\pi_i(\phi)\in\mathrm{C}_i(\nu_i,\overline{\nu}_i,\varepsilon\,|\,t_i)} r_i(\pi_i\,|\,\vartheta_i(t_i))$.

Since $R_i \in \mathcal{R}_i^{\mathrm{loc}}(t_i)$, by definition, there exists a sequence $\{t_{i,m}\}_{m\in\mathbb{N}} \subset T_i^*$ such that $t_{i,m} \to t_i$, and $R_i \supseteq \mathbf{W}_i(t_{i,m})$ for all m. Because Θ_i is finite, we can assume $\vartheta_i^*(t_{i,m}) = \vartheta_i(t_i)$ for all m without loss of generality. For each m, define $\nu_{i,m} \in \Delta\big(\vartheta_i(t_i) \times T_{-i}^* \times 2^{S_{-i}}\big)$ by

$$\nu_{i,m}[\{(\omega, y_{-i}, R_{-i}) : y_{-i} \in E_{-i}\}] = \kappa_i^*(t_{i,m})\,[\{(\omega, y_{-i}) : y_{-i} \in E_{-i} \text{ and } \mathbf{W}_{-i}(y_{-i}) = R_{-i}\}]\,, \tag{7}$$

for every measurable $E_{-i} \subseteq T_{-i}^*$ and every $(\omega, R_{-i}) \in \vartheta_i^*(t_{i,m}) \times 2^{S_{-i}}$. Because the space of probability measures $\Delta\big(\vartheta_i(t_i) \times T_{-i}^* \times 2^{S_{-i}}\big)$ is weak* compact metrizable, $\{\nu_{i,m}\}$ has a convergent subsequence $\{\nu_{i,m_k}\}$, and let ν_i denote its limit. We first verify that ν_i satisfies conditions (i) and (ii) above. For (i), since $\mathrm{marg}_{\vartheta_i(t_i)\times T_{-i}^*}\,\nu_{i,m_k} = \kappa_i^*(t_{i,m_k})$ by definition, $t_{i,m_k} \to t_i$ and $\nu_{i,m_k} \to \nu_i$ as $k \to \infty$, and κ_i^* is continuous, we have $\mathrm{marg}_{\vartheta_i(t_i)\times T_{-i}}\,\nu_i = \kappa_i(t_i)$. For (ii), let $\ell \in \mathbb{N}$ and define

$$\begin{aligned} F_\ell &= \mathrm{cl}\left\{(\omega, y_{-i}, R_{-i}) : \exists y'_{-i} \in T_{-i}^* \text{ s.t. } d_{-i}(y_{-i}, y'_{-i}) \le \frac{1}{\ell} \text{ and } \mathbf{W}_{-i}(y'_{-i}) = R_{-i}\right\},\\ F_\infty &= \big(\vartheta_i(t_i) \times T_{-i} \times 2^{S_{-i}}\big) \cap \bigcap_{\ell\ge 1} F_\ell, \end{aligned}$$

where d_{-i} is the metric on T_{-i}^*. Observe that

$$\begin{aligned} F_\ell &\supseteq \big\{(\omega, y_{-i}, R_{-i}) : y_{-i} \in T_{-i}^* \text{ and } \mathbf{W}_{-i}(y_{-i}) = R_{-i}\big\}\,,\ \forall \ell \ge 0\\ F_\infty &\subseteq \big\{(\omega, y_{-i}, R_{-i}) : y_{-i} \in T_{-i} \text{ and } R_{-i} \in \mathcal{R}_{-i}^{\mathrm{loc}}(t_{-i})\big\}\,. \end{aligned}$$

By definition, $\nu_{i,m}[\{(\omega, y_{-i}, R_{-i}) : \mathbf{W}_{-i}(y_{-i}) = R_{-i}\}] = 1$, so $\nu_{i,m}[F_\ell] = 1$ for all ℓ. Since F_ℓ is closed and $\nu_{i,m_k} \to \nu_i$ as $k \to \infty$, we have $\nu_i[F_\ell] \ge \limsup \nu_{i,m_k}[F_\ell] = 1$. Because $\mathrm{marg}_{\vartheta_i(t_i)\times T_{-i}}\,\nu_i = \kappa_i(t_i)$, we also have $\nu_i\big[\vartheta_i(t_i) \times T_{-i} \times 2^{S_{-i}}\big] = 1$. Therefore, we conclude that $\nu_i[F_\infty] = 1$, which implies $\nu_i\big[\big\{(\omega, t_{-i}, R_{-i}) : R_{-i} \in \mathcal{R}_{-i}^{\mathrm{loc}}(t_{-i})\big\}\big] = 1$. Condition (ii) above is then implied by the induction hypothesis $\mathcal{R}_i^{\mathrm{loc}}(t_i) \subseteq \mathcal{R}_i^{\mathrm{loc},n-1}(t_i)$ for all $i \in I$ and $t_i \in T_i$.

We next construct $\overline{\nu}_i \in \Delta\big(\vartheta_i(t_i) \times 2^{S_{-i}}\big)$ as follows:

$$\overline{\nu}_i[\omega, R_{-i}] = \frac{1}{\varepsilon}\left(\nu_{i,m}[\{(\omega, y_{-i}, R_{-i}) : y_{-i} \in T_{-i}^*\}] - (1-\varepsilon)\sum_{t_{-i}\in T_{-i}} \nu_i[\omega, t_{-i}, R_{-i}]\right). \tag{8}$$

Since $\nu_{i,m_k} \to \nu_i$, we can choose a sufficiently large m so that $\overline{\nu}_i[\omega, R_i] \ge 0$ for every (ω, R_i). When $\overline{\nu}_i[\omega, R_i] > 0$, we must have $\nu_{i,m}[\{(\omega, y_{-i}, R_{-i}) : y_{-i} \in T_{-i}^*\}] > 0$. By definition of $\nu_{i,m}$, there exists some type $y_{-i} \in T_{-i}^*$ such that $\vartheta_{-i}^*(y_{-i}) = \theta_{-i}(\omega)$ and $R_{-i} = \mathbf{W}_{-i}(y_{-i})$, so $R_{-i} \in \overline{\mathcal{R}}_{-i}^{\uparrow}(\theta_{-i}(\omega))$. Hence, condition (iii) holds.

Finally, we need to show that for any conjecture $\pi_i(\phi) \in \mathrm{C}_i(\nu_i, \overline{\nu}_i, \varepsilon \,|\, t_i)$, we have $R_i \supseteq r_i(\pi_i \,|\, \vartheta_i(t_i))$. Let $\lambda_i, \lambda'_i \in \Delta(\vartheta_i(t_i) \times S_{-i})$ be the distributions consistent with ν_i and $\overline{\nu}_i$, respectively, such that

$$\pi_i(\phi) = (1-\varepsilon)\lambda_i + \varepsilon\lambda'_i.$$

We now define a distribution $\mu_i \in \Delta\big(\vartheta_i(t_i) \times T^*_{-i} \times S_{-i}\big)$ as follows: For every $\omega \in \vartheta_i(t_i)$, $s_{-i} \in S_{-i}$, and measurable $E_{-i} \subseteq T^*_{-i}$, let

$$\mu_i[\{(\omega, y_{-i}, s_{-i}) : y_{-i} \in E_{-i}\}] = K_i(\omega, E_{-i}) \cdot ((1-\varepsilon)\lambda_i + \varepsilon\lambda'_i)\,[\omega, s_{-i}]. \tag{9}$$

where $K_i(\omega, E_{-i})$ is a multiplier defined by

$$K_i(\omega, E_{-i}) = \frac{\left(\operatorname{marg}_{\vartheta_i(t_i)\times T^*_{-i}} \nu_{i,m}\right)[\{(\omega, y_{-i}) : y_{-i} \in E_{-i}\}]}{\left(\operatorname{marg}_{\vartheta_i(t_i)\times T^*_{-i}} \nu_{i,m}\right)[\{(\omega, y_{-i}) : y_{-i} \in T^*_{-i}\}]}. \tag{10}$$

It can be checked that the probability distribution μ_i is well-defined. Moreover, for every $\omega \in \vartheta_i(t_i)$ and measurable $E_{-i} \subseteq T^*_{-i}$, we have

$$\begin{aligned}
&\left(\operatorname*{marg}_{\vartheta_i(t_i)\times T^*_{-i}} \mu_i\right)[\{(\omega, y_{-i}) : y_{-i} \in E_{-i}\}] \\
&= K_i(\omega, E_{-i}) \cdot \left(\operatorname*{marg}_{\vartheta_i(t_i)\times T^*_{-i}} \nu_{i,m}\right)[\{(\omega, y_{-i}) : y_{-i} \in T^*_{-i}\}] \\
&= \left(\operatorname*{marg}_{\vartheta_i(t_i)\times T^*_{-i}} \nu_{i,m}\right)[\{(\omega, y_{-i}) : y_{-i} \in E_{-i}\}] \\
&= \kappa^*_i(t_{i,m})[\{(\omega, y_{-i}) : y_{-i} \in E_{-i}\}],
\end{aligned}$$

where the first equality is by equations (9) and (8); the second equality is by equation (10); and the third equality is by equation (7). The above implies $\operatorname{marg}_{\vartheta_i(t_i)\times T^*_{-i}} \mu_i = \kappa^*_i(t_{i,m})$. Also, by the definitions of ν_i and $\overline{\nu}_i$, the fact that λ_i and λ'_i are consistent with ν_i and $\overline{\nu}_i$ respectively, and $\nu_{i,m}[\{(\omega, y_{-i}, R_{-i}) : \mathbf{W}_{-i}(y_{-i}) = R_{-i}\}] = 1$, we have $\mu_i[\{(\omega, y_{-i}, s_{-i}) : s_{-i} \in \mathbf{W}_{-i}(y_{-i})\}] = 1$.

Our last step is to show that $r_i(\pi_i \,|\, \vartheta_i(t_i)) \subseteq \mathbf{W}_i(t_{i,m})$, but this can be obtained by invoking Definition 3 of $\mathbf{W}$ and by simply noting that for each pair (ω, s_{-i}),

$$\begin{aligned}
\mu_i\left[\{(\omega, y_{-i}, s_{-i}) : y_{-i} \in T^*_{-i}\}\right] &= (1-\varepsilon)\lambda_i[\omega, s_{-i}] + \varepsilon\lambda'_i[\omega, s_{-i}] \\
&= \pi_i(\phi)[\omega, s_{-i}],
\end{aligned}$$

where the first equality is by combining equations (9) and (10). Since $R_i \supseteq \mathbf{W}_i(t_{i,m})$, we can conclude that $R_i \supseteq r_i(\pi_i \,|\, \vartheta_i(t_i))$, which wraps up the proof. □

A.1.5 Proof of Proposition 5

(1) ⇒ (2) We split this direction into two steps.

First, we construct a finite type space such that the collection of $\mathbf{W}$ sets contains all maximal sets in $\overline{\mathcal{R}}^{\downarrow}_i(\theta_i)$. To achieve this, for each $i \in I$, partition the space T^*_i into

$\mathbf{T}_i^{\mathbf{W}} = \{T_i^{(\theta_i,R_i)}\}$ by payoff types and the $\mathbf{W}$ correspondence; i.e., $t_i, t_i' \in T_i^{(\theta_i,R_i)}$ if and only if $\vartheta_i^*(t_i) = \vartheta_i^*(t_i') = \theta_i$ and $\mathbf{W}_i(t_i) = \mathbf{W}_i(t_i') = R_i$. Note that the partition $\mathbf{T}_i^{\mathbf{W}}$ is finite and measurable by continuity of ϑ_i^*, upper hemicontinuity of $\mathbf{W}$, and finiteness of Θ_i and S_i. Now for each $T_i^{(\theta_i,R_i)} \in \mathbf{T}_i^{\mathbf{W}}$, fix a type $t_i^{(\theta_i,R_i)} \in T_i^{(\theta_i,R_i)}$. Define a finite type space $\left\{\left(\mathbf{T}_i^{\mathbf{W}}, \vartheta_i, \kappa_i\right)_{i\in I}\right\}$ such that $\vartheta_i\big(T_i^{(\theta_i,R_i)}\big) = \theta_i$, and

$$\kappa_i\left(T_i^{(\theta_i,R_i)}\right)\left[\omega, T_{-i}^{(\theta_{-i},R_{-i})}\right] = \kappa_i^*\left(t_i^{(\theta_i,R_i)}\right)\left[\left\{(\omega, t_{-i}) : t_{-i} \in T_{-i}^{(\theta_{-i},R_{-i})}\right\}\right]. \tag{11}$$

Note that $T_i^{(\theta_i,R_i)}$ denotes both a type in $\mathbf{T}_i^{\mathbf{W}}$ (on the left-hand-side) and a subset of T_i^* (on the right-hand-side). We claim that $\mathbf{W}_i\big(T_i^{(\theta_i,R_i)}\big) \supseteq R_i$ for every $T_i^{(\theta_i,R_i)} \in \mathbf{T}_i^{\mathbf{W}}$. To prove this, we will show that $\mathbf{W}_i^n\big(T_i^{(\theta_i,R_i)}\big) \supseteq R_i$ for all $n \geq 0$. This is clearly true for $n = 0$. Suppose it holds for $n-1$. Take $t_i^{(\theta_i,R_i)} \in T_i^{(\theta_i,R_i)}$ and $s_i \in R_i = \mathbf{W}_i(t_i)$. Then by Definition 3, there exists a distribution $\mu_i \in \Delta\big(\theta_i \times T_{-i}^* \times S_{-i}\big)$ and $\pi_i \in \Delta^{\mathcal{H}_i}(\theta_i \times S_{-i})$ such that

(i) $\operatorname{marg}_{\theta_i \times T_{-i}^*} \mu_i = \kappa_i^*\left(t_i^{(\theta_i,R_i)}\right)$;

(ii) $\mu_i[\{(\omega, t_{-i}, s_{-i}) : (t_{-i}, s_{-i}) \in \operatorname{graph}(\mathbf{W}_{-i})\}] = 1$;

(iii) $\pi_i(\phi) = \operatorname{marg}_{\theta_i \times S_{-i}} \mu_i$;

(iv) $s_i \in r_i(\pi_i \,|\, \theta_i)$.

Now define a new distribution $\mu_i' \in \Delta(\theta_i \times \mathbf{T}_{-i}^{\mathbf{W}} \times S_{-i})$ for type $T_i^{(\theta_i,R_i)} \in \mathbf{T}_i^{\mathbf{W}}$ by

$$\mu_i'\left[\omega, T_{-i}^{(\theta_{-i},R_{-i})}, s_{-i}\right] = \mu_i\left[\left\{(\omega, t_{-i}, s_{-i}) : t_{-i} \in T_{-i}^{(\theta_{-i},R_{-i})}\right\}\right].$$

By equation (11) and condition (i) above, we have $\operatorname{marg}_{\theta_i \times \mathbf{T}_{-i}^{\mathbf{W}}} \mu_i' = \kappa_i\big(T_i^{(\theta_i,R_i)}\big)$. Since for every $t_{-i} \in T_{-i}^{(\theta_{-i},R_{-i})}$, we have $\mathbf{W}_{-i}(t_{-i}) = R_{-i}$, condition (ii) above and the induction hypothesis $\mathbf{W}_i^{n-1}\big(T_i^{(\theta_i,R_i)}\big) \supseteq R_i$ together imply that

$$\mu_i'\left[\left(\omega, T_{-i}^{(\theta_{-i},R_{-i})}, s_{-i}\right) : (T_{-i}^{(\theta_{-i},R_{-i})}, s_{-i}) \in \operatorname{graph}\left(\mathbf{W}_{-i}^{n-1}\right)\right] = 1.$$

Moreover, by construction,

$$\operatorname*{marg}_{\theta_i \times S_{-i}} \mu_i' = \operatorname*{marg}_{\theta_i \times S_{-i}} \mu_i = \pi_i(\phi).$$

Therefore, the fact that $s_i \in r_i(\pi_i \,|\, \theta_i)$ implies $s_i \in \mathbf{W}_i^n\big(T_i^{(\theta_i,R_i)}\big)$, and thus we have proved our claim $\mathbf{W}_i^n\big(T_i^{(\theta_i,R_i)}\big) \supseteq R_i$.

To summarize the first step, for any $R_i \in \overline{\mathcal{R}}_i^{\downarrow}(\theta_i)$, there is a *finite* type t_i such that $\vartheta_i^*(t_i) = \theta_i$ and $R_i \subseteq \mathbf{W}(t_i)$. The second step is simple. Since any strategy $s_i \in \mathbf{W}(t_i)$ can be uniquely selected for t_i, we have $\{s_i\} \in \overline{\mathcal{R}}_i^{\uparrow}(\theta_i)$ for all $s_i \in R_i$.

(2) ⇒ (1) Fixing a finite type space $\{(T_i, \vartheta_i, \kappa_i)_{i\in I}\}$, we need to show that for each $t_i \in T_i$ and any strategy $s_i \in \mathbf{W}_i(t_i)$, the singleton set $\{s_i\} \in \mathcal{R}_i^{\mathrm{loc}}(t_i)$. By Proposition 4, we can prove this by showing $\{s_i\} \in \mathcal{R}_i^{\mathrm{loc},n}(t_i)$ for all $n \geq 0$. For $n = 0$, note that $\mathbf{W}_i(t_i) \in \overline{\mathcal{R}}_i^{\downarrow}(\vartheta_i(t_i))$, and hence $\mathbf{W}_i(t_i) \subseteq \overline{\mathbf{W}}_i^{\mathrm{u}}(\vartheta_i(t_i))$ by assumption. This means each $s_i \in \mathbf{W}_i(t_i)$ forms a singleton set in $\overline{\mathcal{R}}_i^{\uparrow}(\vartheta_i(t_i))$. But $\mathcal{R}_i^{\mathrm{loc},0}(t_i) = \overline{\mathcal{R}}_i^{\uparrow}(\vartheta_i(t_i))$ by definition, so $\{s_i\} \in \mathcal{R}_i^{\mathrm{loc},0}(t_i)$ is established.

Suppose next for each $t_i \in T_i$, $s_i \in \mathbf{W}_i(t_i)$ implies $\{s_i\} \in \mathcal{R}_i^{\mathrm{loc},n-1}(t_i)$. We now show that the same statement holds for n. Fix a type $t_i \in T_i$. If $s_i \in \mathbf{W}_i(t_i)$, by Definition 3, there exists a distribution $\mu_i \in \Delta\big(\vartheta_i(t_i) \times T_{-i} \times S_{-i}\big)$ and a conjecture $\pi_i \in \Delta^{\mathcal{H}_i}(\vartheta_i(t_i) \times S_{-i})$ such that

(i) $\operatorname{marg}_{\vartheta_i(t_i)\times T_{-i}} \mu_i = \kappa_i(t_i)$;

(ii) $\mu_i[\{(\omega, t_{-i}, s_{-i}) : (t_{-i}, s_{-i}) \in \operatorname{graph}(\mathbf{W}_{-i})\}] = 1$;

(iii) $\pi_i(\phi) = \operatorname{marg}_{\vartheta_i(t_i)\times S_{-i}} \mu_i$;

(iv) $s_i \in r_i(\pi_i \,|\, \vartheta_i(t_i))$.

Now define $\nu_i \in \Delta\big(\vartheta_i(t_i) \times T_{-i} \times 2^{S_{-i}}\big)$ by

$$\nu_i[\omega, t_{-i}, \{s_{-i}\}] = \mu_i[\omega, t_{-i}, s_{-i}].$$

By construction, we have $\operatorname{marg}_{\vartheta_i(t_i)\times T_{-i}} \nu_i = \kappa_i(t_i)$, and the *only* λ_i consistent with ν_i is the one such that $\lambda_i = \pi_i(\phi)$. Moreover, by the induction hypothesis, we also have

$$\nu_i\left[\left\{(\omega, t_{-i}, R_{-i}) : R_{-i} \in \mathcal{R}_{-i}^{\mathrm{loc},n-1}(t_{-i})\right\}\right] = 1.$$

Since $s_i \in \overline{\mathbf{W}}_i^{\mathrm{u}}(\vartheta_i(t_i))$ by assumption, there exists a type $t_i^{s_i} \in T_i^*$ such that $\vartheta_i^*\big(t_i^{s_i}\big) = \vartheta_i(t_i)$ and $\mathbf{W}_i\big(t_i^{s_i}\big) = \{s_i\}$. We now define $\overline{\nu}_i \in \Delta\big(\vartheta_i(t_i) \times 2^{S_{-i}}\big)$ by

$$\overline{\nu}_i[\omega, R_{-i}] = \kappa_i^*\big(t_i^{s_i}\big)[\{(\omega, t_{-i}) : \mathbf{W}_{-i}(t_{-i}) = R_{-i}\}].$$

Note that this construction ensures

$$\overline{\nu}_i\left[\left\{(\omega, R_{-i}) : (\theta_{-i}(\omega), R_{-i}) \in \operatorname{graph}\left(\overline{\mathcal{R}}_{-i}^{\uparrow}\right)\right\}\right] = 1.$$

Moreover, for any conjecture $\pi_i' \in \Delta^{\mathcal{H}_i}(\vartheta_i(t_i) \times S_{-i})$ that satisfies $\pi_i'(\phi) \in \mathrm{C}_i(\overline{\nu}_i \,|\, \vartheta_i(t_i))$, there exists a distribution $\mu_i' \in \Delta(\vartheta_i(t_i) \times T_{-i}^* \times S_{-i})$ such that

(i) $\operatorname{marg}_{\vartheta_i(t_i)\times T_{-i}^*} \mu_i' = \kappa_i^*\big(t_i^{s_i}\big)$;

(ii) $\mu_i'[\{(\omega, t_{-i}, s_{-i}) : (t_{-i}, s_{-i}) \in \operatorname{graph}(\mathbf{W}_{-i})] = 1$;

(iii) $\pi_i'(\phi) = \operatorname{marg}_{\vartheta_i(t_i)\times S_{-i}} \mu_i'$.

Because $\mathbf{W}_i(t_i^{s_i}) = \{s_i\}$, we must have $\{s_i\} = r_i(\pi_i' \,|\, \vartheta_i(t_i))$.

Finally, for any $\varepsilon \in (0, 1]$, take the distributions ν_i and $\overline{\nu}_i$ above. We claim that for any $\hat{\pi}_i \in \Delta^{\mathcal{H}_i}(\vartheta_i(t_i) \times S_{-i})$ such that $\hat{\pi}_i(\phi) \in \mathrm{C}_i(\nu_i, \overline{\nu}_i, \varepsilon \,|\, t_i)$, we must have $\{s_i\} = r_i(\hat{\pi}_i)$. To see this, write $\hat{\pi}_i(\phi) = (1-\varepsilon)\lambda_i + \varepsilon\lambda_i'$, where $(\lambda_i, \lambda_i') \in \mathrm{C}_i(\nu_i \,|\, t_i) \times \mathrm{C}_i(\overline{\nu}_i \,|\, \vartheta_i(t_i))$. For every $h \in \mathcal{H}_i(s_i)$, consider three cases:

- First, suppose $[h] \cap \operatorname{supp} \lambda_i' = \varnothing$; that is, information set h is off-path for the initial belief $\pi_i'(\phi) = \lambda_i'$. It is easy to construct a conjecture $\tilde{\pi}_i \in \Delta^{\mathcal{H}_i}(\vartheta_i(t_i) \times S_{-i})$ such that $\tilde{\pi}_i(\phi) = \pi_i'(\phi) \in \mathrm{C}_i(\overline{\nu}_i \,|\, \vartheta_i(t_i))$ and $\tilde{\pi}_i(h) = \hat{\pi}_i(h)$. Because $\{s_i\} = r_i(\tilde{\pi}_i \,|\, \vartheta_i(t_i))$, s_i must be the unique best response against $\hat{\pi}_i(h)$ at information set h;

- Next, suppose $[h] \cap \operatorname{supp} \lambda_i' \neq \varnothing$ but $[h] \cap \operatorname{supp} \lambda_i = \varnothing$. This means $\hat{\pi}_i(h) = \pi_i'(h)$, and therefore s_i is the unique best response against $\hat{\pi}_i(h)$;

- Suppose $[h] \cap \operatorname{supp} \lambda_i \cap \operatorname{supp} \lambda_i' \neq \varnothing$. Then, $\hat{\pi}(h)$ is a convex combination of distributions such that (with probability $1-\varepsilon$) s_i is a best response and (with probability ε) s_i is the only best response. This guarantees that s_i is the unique best response against $\hat{\pi}_i(h)$.

Hence, for all information sets $h \in \mathcal{H}_i(s_i)$,

$$\{s_i\} = \operatorname*{arg\,max}_{s_i' \in S_i(h)} \sum_{\omega, s_{-i}} u_i(\omega, z(s_i', s_{-i}))\hat{\pi}_i(h)[\omega, s_{-i}].$$

Therefore, $\{s_i\} = \bigcup_{\hat{\pi}_i(\phi) \in \mathrm{C}_i(\nu_i, \overline{\nu}_i, \varepsilon \,|\, t_i)} r_i(\hat{\pi}_i \,|\, \vartheta_i(t_i))$. This means $\{s_i\} \in \mathcal{R}_i^{\mathrm{loc},n}(t_i)$, completing the proof.

A.1.6 Proof of Proposition 6

(1) ⇒ (2) Fix a player $i \in I$ and $\theta_i \in \Theta_i$. If $R_i \in \overline{\mathcal{R}}_i^{\uparrow}(\theta_i)$, there exists a type $t_i \in T_i^*$ such that $\vartheta_i^*(t_i) = \theta_i$ and $R_i \supseteq \mathbf{W}_i(t_i)$. By denseness of $\mathcal{U}_i$, there is a sequence $\{t_{i,m}\}_{m \in \mathbb{N}} \subseteq T_i^*$ such that $t_{i,m} \to t_i$ as $m \to \infty$, and $|\mathbf{W}_i(t_{i,m})| = 1$. Since Θ_i and S_i are finite, there exists a subsequence $\{t_{i,m_k}\}$ such that $\vartheta_i^*(t_{i,m_k}) = \theta_i$ and $\{s_i\} = \mathbf{W}_i(t_{i,m_k})$ for some $s_i \in S_i$. This means $\{s_i\} \in \overline{\mathcal{R}}_i^{\uparrow}(\theta_i)$. By upper hemicontinuity of $\mathbf{W}$, we must have $s_i \in \mathbf{W}_i(t_i)$. Since $R_i \supseteq \mathbf{W}_i(t_i)$, we also have $s_i \in R_i$.

(2) ⇒ (1) Openness of $\mathcal{U}_i$ is implied by the fact that $\mathbf{W}_i(\cdot)$ is upper hemicontinuous and nonempty on T_i^*. Denseness of $\mathcal{U}_i$ is proved in two steps.

First, we show that for every $t_i \in T_i$ in any finite type space $\{(T_i, \vartheta_i, \kappa_i)_{i \in I}\}$, the set $\mathbf{W}_i(t_i)$ contains some strategies that can be uniquely selected. To this end, we define a refinement of $\mathbf{W}$ which collects all strategies that can be uniquely selected for a type. For each $i \in I$

and $\theta_i \in \Theta_i$, let $\overline{\mathbf{W}}_i^{\mathrm{u}}(\theta_i) = \{s_i \in S_i : \{s_i\} \in \overline{\mathcal{R}}_i^{\uparrow}(\theta_i)\}$ denote the set of uniquely rationalizable strategies for types with information θ_i. For each $t_i \in T_i$, let $\mathbf{W}_i^{\mathrm{u},0}(t_i) = \overline{\mathbf{W}}_i^{\mathrm{u}}(\vartheta_i(t_i))$, which is nonempty by assumption. For $n \geq 1$, let $\mathbf{W}_{-i}^{\mathrm{u},n-1}(t_{-i}) = \bigtimes_{j \neq i} \mathbf{W}_j^{\mathrm{u},n-1}(t_j)$, and define

$$\mathbf{W}_i^{\mathrm{u},n}(t_i) = \left\{ s_i \in \overline{\mathbf{W}}_i^{\mathrm{u}}(\vartheta_i(t_i)) : \begin{array}{l} \exists (\pi_i, \mu_i) \in \Delta^{\mathcal{H}_i}(\vartheta_i(t_i) \times S_{-i}) \times \Delta(\vartheta_i(t_i) \times T_{-i} \times S_{-i}) \text{ s.t.} \\ \text{(i) } \operatorname{marg}_{\vartheta_i(t_i) \times T_{-i}} \mu_i = \kappa_i(t_i); \\ \text{(ii) } \mu_i \left[\left\{ (\omega, t_{-i}, s_{-i}) : (t_{-i}, s_{-i}) \in \operatorname{graph}\left(\mathbf{W}_{-i}^{\mathrm{u},n-1}\right) \right\} \right] = 1; \\ \text{(iii) } \pi_i(\phi) = \operatorname{marg}_{\vartheta_i(t_i) \times S_{-i}} \mu_i; \\ \text{(iv) } s_i \in r_i(\pi_i \,|\, \vartheta_i(t_i)) \end{array} \right\}.$$

Finally, let $\mathbf{W}_i^{\mathrm{u}}(t_i) = \bigcap_{n \geq 0} \mathbf{W}_i^{\mathrm{u},n}(t_i)$. Observe that $\{\mathbf{W}_i^{\mathrm{u},n}(t_i)\}_{n \in \mathbb{N}} \subseteq \overline{\mathbf{W}}_i^{\mathrm{u}}(\vartheta_i(t_i))$ is a decreasing sequence for each t_i and converges in finitely many steps.

Claim 1. *For all $n \geq 0$, $i \in I$, and $t_i \in T_i$, $\mathbf{W}_i^{\mathrm{u},n}(t_i) \subseteq \mathbf{W}_i^n(t_i)$. Hence, $\mathbf{W}_i^{\mathrm{u}}(t_i) \subseteq \mathbf{W}_i(t_i)$.*

Proof. For $n = 0$, we have $\overline{\mathbf{W}}_i^{\mathrm{u}}(\vartheta_i(t_i)) = \mathbf{W}_i^{\mathrm{u},0}(t_i) \subseteq \mathbf{W}_i^0(t_i) = S_i$ by definition. Suppose $\mathbf{W}_i^{\mathrm{u},n-1}(t_i) \subseteq \mathbf{W}_i^{n-1}(t_i)$ for all $t_i \in T_i$ and $i \in I$. Then $\operatorname{graph}\left(\mathbf{W}_{-i}^{\mathrm{u},n-1}\right) \subseteq \operatorname{graph}\left(\mathbf{W}_{-i}^{n-1}\right)$, which implies $\mathbf{W}_i^{\mathrm{u},n}(t_i) \subseteq \mathbf{W}_i^n(t_i)$. □

Claim 2. *For all $n \geq 0$, $\mathbf{W}_i^{\mathrm{u},n}(t_i) \neq \varnothing$.*

Proof. We prove this by induction. Each $\mathbf{W}_i^{\mathrm{u},0}(t_i) = \overline{\mathbf{W}}_i^{\mathrm{u}}(\vartheta_i(t_i))$ is nonempty by assumption. Suppose $\mathbf{W}_i^{\mathrm{u},n-1}(t_i)$ is nonempty for every $i \in I$ and $t_i \in T_i$. Take any distribution $\mu_i \in \Delta(\vartheta_i(t_i) \times T_{-i} \times S_{-i})$ such that $\operatorname{marg}_{\vartheta_i(t_i) \times T_{-i}} \mu_i = \kappa_i(t_i)$ and $\mu_i\big[\big\{(\omega, t_{-i}, s_{-i}) : (t_{-i}, s_{-i}) \in \operatorname{graph}\left(\mathbf{W}_{-i}^{\mathrm{u},n-1}\right)\big\}\big] = 1$. We need to show that there exists a conjecture $\pi_i \in \Delta^{\mathcal{H}_i}(\vartheta_i(t_i) \times S_{-i})$ such that $\pi_i(\phi) = \operatorname{marg}_{\vartheta_i(t_i) \times S_{-i}} \mu_i$ and $r_i(\pi_i \,|\, \vartheta_i(t_i)) \cap \overline{\mathbf{W}}_i^{\mathrm{u}}(\vartheta_i(t_i)) \neq \varnothing$. Define $\bar{\nu}_i \in \Delta\big(\vartheta_i(t_i) \times 2^{S_{-i}}\big)$ by

$$\bar{\nu}_i[\omega, \{s_{-i}\}] = \sum_{t_{-i} \in T_{-i}} \mu_i[\omega, t_{-i}, s_{-i}].$$

Because $\mathbf{W}_i^{\mathrm{u},n-1}(t_i)$ is a nonempty subset of $\overline{\mathbf{W}}_i^{\mathrm{u}}(\vartheta_i(t_i))$, this $\bar{\nu}_j$ is well-defined and satisfies

$$\bar{\nu}_i \left[\left\{ (\omega, R_{-i}) : (\theta_{-i}(\omega), R_{-i}) \in \operatorname{graph}\left(\overline{\mathcal{R}}_{-i}^{\uparrow}\right) \right\} \right] = 1.$$

By construction, any conjecture π_i with $\pi_i(\phi) \in \mathrm{C}_i(\bar{\nu}_i \,|\, \vartheta_i(t_i))$ satisfies $\pi_i(\phi) = \operatorname{marg}_{\vartheta_i(t_i) \times S_{-i}} \mu_i$. By definition of $\overline{\mathcal{R}}_i^{\uparrow}$ and its convergence in finitely many steps, we have $\bigcup_{\pi_i(\phi) \in \mathrm{C}_i(\bar{\nu}_i \,|\, \vartheta_i(t_i))} r_i(\pi_i \,|\, \vartheta_i(t_i)) \in \overline{\mathcal{R}}_i^{\uparrow}(\vartheta_i(t_i))$. By assumption, we then have

$$\left(\bigcup_{\pi_i(\phi) \in \mathrm{C}_i(\bar{\nu}_i \,|\, \vartheta_i(t_i))} r_i(\pi_i \,|\, \vartheta_i(t_i)) \right) \cap \overline{\mathbf{W}}_i^{\mathrm{u}}(\vartheta_i(t_i)) \neq \varnothing.$$

Therefore, $\mathbf{W}_i^{\mathrm{u},n}(t_i) \neq \varnothing$, which means the iterative procedure converges to a nonempty set of strategies for every $t_i \in T_i$. □

We claim that for every $s_i \in \mathbf{W}_i^{\mathrm{u}}(t_i)$, the singleton $\{s_i\} \in \mathcal{R}_i^{\mathrm{loc}}(t_i)$. We prove this by induction. Since any $s_i \in \mathbf{W}_i^{\mathrm{u},0}(t_i)$ is in the set $\mathbf{W}_i^{\mathrm{u}}(\vartheta_i(t_i))$, we know $\{s_i\} \in \mathcal{R}_i^{\mathrm{loc},0}(t_i)$ by definition. Suppose $s_i \in \mathbf{W}_i^{\mathrm{u},n-1}(t_i)$ implies $\{s_i\} \in \mathcal{R}_i^{\mathrm{loc},n-1}(t_i)$ and take any $s_i \in \mathbf{W}_i^{\mathrm{u},n}(t_i)$. Constructions of ν_i and $\bar{\nu}_i$ in the proof of Proposition 5, with obvious modifications, can be used to show that $\{s_i\} \in \mathcal{R}_i^{\mathrm{loc},n}(t_i)$ (see Appendix A.1.5).

Therefore, our first step shows that for any finite type $t_i \in T_i^*$, there is some strategy $s_i \in \mathbf{W}_i(t_i)$ that can be uniquely selected. The second step is standard: Let $T_i^{\mathrm{f}} \subseteq T_i^*$ be the collection of all finite types of player i. Since $t_i \in \mathrm{cl}(\mathcal{U}_i)$ for all $t_i \in T_i^{\mathrm{f}}$, we have $T_i^{\mathrm{f}} \subseteq \mathrm{cl}(\mathcal{U}_i)$. Because T_i^{f} is dense in T_i^* (Mertens and Zamir, 1985), we know that $T_i^* = \mathrm{cl}(T_i^{\mathrm{f}}) \subseteq \mathrm{cl}(\mathcal{U}_i)$, which means $\mathcal{U}_i$ is also dense in T_i^*.

A.2 Proofs for Section 4: Applications

A.2.1 Proof of Proposition 7

We first establish the following lemma.

Lemma 7. *For every $i \in I$, $\theta_i \in \Theta_i$, and every $n \geq 0$, we have $\overline{\mathbf{S}}_i^n(\theta_i) \cap R_i \neq \emptyset$ for all $R_i \in \overline{\mathcal{R}}_i^{\uparrow,n}(\theta_i)$. Therefore, $\overline{\mathbf{S}}_i(\theta_i) \cap R_i \neq \emptyset$ for all $R_i \in \overline{\mathcal{R}}_i^{\uparrow}(\theta_i)$.*

Proof. By definition, we have $\overline{\mathbf{S}}_i^0(\theta_i) = S_i$ and $\overline{\mathcal{R}}_i^{\uparrow,0}(\theta_i) = \{S_i\}$, so the statement holds for $n = 0$. By induction, suppose the statement holds for $n-1$, and let $R_i \in \overline{\mathcal{R}}_i^{\uparrow,n}(\theta_i)$. Then there exists a distribution $\bar{\nu}_i \in \Delta(\theta_i \times 2^{S_{-i}})$ such that $\bar{\nu}_i\big[\big\{(\omega, R_{-i}) : (\theta_{-i}(\omega), R_{-i}) \in \mathrm{graph}\,\big(\overline{\mathcal{R}}_{-i}^{\uparrow,n-1}\big)\big\}\big] = 1$ and $R_i \supseteq \bigcup_{\pi_i(\phi)\in \mathrm{C}_i(\bar{\nu}_i \,|\, \theta_i)} r_i(\pi_i \,|\, \theta_i)$. Let $\bar{f} : \theta_i \times 2^{S_{-i}} \to \Delta(S_{-i})$ be such that $\bar{f}_i(\omega, R_{-i})[\overline{\mathbf{S}}_{-i}^{n-1}(\theta_{-i}) \cap R_{-i}] = 1$ whenever $\bar{\nu}_i[\omega, R_{-i}] > 0$. By the induction hypothesis, this function is well-defined. Now define a conjecture $\pi_i \in \Delta^{\mathcal{H}_i}(\theta_i \times S_{-i})$ that satisfies the following two conditions:

(i) $\pi_i(\phi)[\omega, s_{-i}] = \sum_{R_{-i}\in 2^{S_{-i}}} \bar{\nu}_i[\omega, R_{-i}]\bar{f}_i(\omega, R_{-i})[s_{-i}]$;

(ii) $\forall h \in \mathcal{H}_i$, $\big(\bigcup_{\omega\in\theta_i} \theta_{-i}(\omega) \times S_{-i}(h)\big) \cap \mathrm{graph}\,\big(\overline{\mathbf{S}}_{-i}^{n-1}\big) \neq \varnothing$ implies

$$\pi_i(h)\left[\left\{(\omega, s_{-i}) : (\theta_{-i}(\omega), s_{-i}) \in \mathrm{graph}\left(\overline{\mathbf{S}}_{-i}^{n-1}\right)\right\}\right] = 1.$$

Clearly, $\pi_i(\phi) \in \mathrm{C}_i(\bar{\nu}_i \,|\, \theta_i)$, which implies that $r_i(\pi_i \,|\, \theta_i) \subseteq \overline{\mathbf{S}}_i^n(\theta_i) \cap R_i \neq \emptyset$. □

This lemma can be used to show the next result.

Lemma 8. *Fix a finite type space $\{(T_i, \vartheta_i, \kappa_i)_{i\in I}\}$. For every $i \in I$, $t_i \in T_i$, and every $n \geq 0$, we have $\mathbf{S}_i^n(t_i) \cap R_i \neq \emptyset$ for all $R_i \in \mathcal{R}_i^{\mathrm{loc},n}(t_i)$. Therefore, $\mathbf{S}_i(t_i) \cap R_i \neq \emptyset$ for all $R_i \in \mathcal{R}_i^{\mathrm{loc}}(t_i)$.*

Proof. By definition, we have $\mathbf{S}_i^0(t_i) = \overline{\mathbf{S}}(\vartheta_i(t_i))$ and $\mathcal{R}_i^{\text{loc},0}(t_i) = \overline{\mathcal{R}}_i^{\uparrow}(\vartheta_i(t_i))$. By Lemma 7, the statement holds for $n = 0$. Suppose now the statement holds for $n-1$, and take an arbitrary $R_i \in \mathcal{R}_i^{\text{loc},n}(t_i)$. By definition of $\mathcal{R}_i^{\text{loc},n}(t_i)$, for each $m \in \mathbb{N}$, there exists $(\nu_{i,m}, \overline{\nu}_{i,m}) \in \Delta\big(\vartheta_i(t_i) \times T_{-i} \times 2^{S_{-i}}\big) \times \Delta\big(\vartheta_i(t_i) \times 2^{S_{-i}}\big)$ such that

(i) $\text{marg}_{\vartheta_i(t_i) \times T_{-i}} \nu_{i,m} = \kappa_i(t_i)$;

(ii) $\nu_{i,m}\big[\big\{(\omega, t_{-i}, R_{-i}) : (t_{-i}, R_{-i}) \in \text{graph}\left(\mathcal{R}_{-i}^{\text{loc},n-1}(t_{-i})\right)\big\}\big] = 1$;

(iii) $\overline{\nu}_{i,m}\big[\big\{(\omega, R_{-i}) : (\theta_{-i}(\omega), R_{-i}) \in \text{graph}\left(\overline{\mathcal{R}}_{-i}^{\uparrow}\right)\big\}\big] = 1$;

(iv) $R_i \supseteq \bigcup_{\pi_i(\phi) \in \mathrm{C}_i(\nu_i, \overline{\nu}_i, \frac{1}{m+1} \,|\, t_i)} r_i(\pi_i \,|\, \vartheta_i(t_i))$.

By the induction hypothesis and $\mathbf{S}_i^{n-1}(t_i) \subseteq \overline{\mathbf{S}}_i(\vartheta_i(t_i))$, we can define a function $f_{i,m} : \vartheta_i(t_i) \times T_{-i} \times 2^{S_{-i}} \to \Delta(S_{-i})$ such that $f_{i,m}(\omega, t_{-i}, R_{-i})[\overline{\mathbf{S}}_{-i}(\vartheta_{-i}(t_{-i})) \cap R_{-i}] = 1$ whenever $\nu_{i,m}[\omega, t_{-i}, R_{-i}] > 0$. Moreover, by Lemma 7, we can also define a function $\bar{f}_{i,m} : \vartheta_i(t_i) \times 2^{S_{-i}} \to \Delta(S_{-i})$ such that $\bar{f}_{i,m}(\omega, R_{-i})[\overline{\mathbf{S}}_{-i}(\theta_{-i}(\omega)) \cap R_{-i}] = 1$ whenever $\overline{\nu}_{i,m}[\omega, R_{-i}] > 0$. Now define a conjecture that satisfies the following conditions: First, at the initial information set,

$$\pi_{i,m}(\phi)[\omega, s_{-i}] = \frac{m}{m+1} \sum_{t_{-i} \in T_{-i}} \sum_{R_{-i} \in 2^{S_{-i}}} \nu_{i,m}[\omega, t_{-i}, R_{-i}] f_{i,m}(\omega, f_{-i}, R_{-i})[s_{-i}] + \frac{1}{m+1} \sum_{R_{-i} \in 2^{S_{-i}}} \overline{\nu}_{i,m}[\omega, R_{-i}] \bar{f}_{i,m}(\omega, R_{-i})[s_{-i}].$$

Note that this implies $\pi_{i,m}(\phi)\big[\big\{(\omega, s_{-i}) : (\theta_{-i}(\omega), s_{-i}) \in \text{graph}\left(\overline{\mathbf{S}}_{-i}\right)\big] = 1$. For all $h \in \mathcal{H}_i$ and all $k \geq 0$, whenever $\left(\bigcup_{\omega \in \theta_i} \theta_{-i}(\omega) \times S_{-i}(h)\right) \cap \text{graph}\left(\overline{\mathbf{S}}_{-i}^k\right) \neq \varnothing$, let

$$\pi_{i,m}(h)\left[\big\{(\omega, s_{-i}) : (\theta_{-i}(\omega), s_{-i}) \in \text{graph}\left(\overline{\mathbf{S}}_{-i}^k\right)\big\}\right] = 1.$$

Clearly, every $s_{i,m} \in r_i(\pi_{i,m} \,|\, \vartheta_i(t_i))$ satisfies $s_{i,m} \in \bigcap_{k \geq 0} \overline{\mathbf{S}}_i^k(\vartheta_i(t_i)) = \overline{\mathbf{S}}_i(\vartheta_i(t_i))$. Now take a converging subsequence $\{\pi_{i,m_k}\}$ and let π_i denote its limit. Moreover, by finiteness of S_i, let $s_i \in r_i(\pi_{i,m_k} \,|\, \vartheta_i(t_i))$ for all k. Since $r_i(\cdot \,|\, \vartheta_i(t_i))$ is upper hemicontinuous, we have $s_i \in r_i(\pi_i \,|\, \vartheta_i(t_i))$. Finally, let $\mu_{i,m_k} \in \Delta(\vartheta_i(t_i) \times T_{-i} \times S_{-i})$ be defined as

$$\mu_{i,m_k}[\omega, t_{-i}, s_{-i}] = \sum_{R_{-i} \in 2^{S_{-i}}} \nu_{i,m}[\omega, t_{-i}, R_{-i}] f_{i,m}(\omega, f_{-i}, R_{-i})[s_{-i}].$$

We have, by definition and our induction hypothesis, $\mu_{i,m_k}\big[\big\{(\omega, t_{-i}, s_{-i}) : (t_{-i}, s_{-i}) \in \text{graph}\left(\mathbf{S}_{-i}^{n-1}\right)\big\}\big] = 1$. Taking a further subsequence if necessary, let μ_i denote the limit of $\{\mu_{i,m_k}\}$. We then have

(i) $\text{marg}_{\vartheta_i(t_i) \times T_{-i}} \mu_i = \kappa_i(t_i)$;

(ii) $\mu_i\big[\big\{(\omega, t_{-i}, s_{-i}) : (t_{-i}, s_{-i}) \in \text{graph}\,\big(\mathbf{S}^{n-1}_{-i}\big)\big\}\big] \geq \limsup \mu_{i,m_k}\big[\big\{(\omega, t_{-i}, s_{-i}) : (t_{-i}, s_{-i}) \in \text{graph}\,\big(\mathbf{S}^{n-1}_{-i}\big)\big\}\big] = 1$;

(iii) $\pi_i(\phi)[\omega, s_{-i}] = \sum_{t_{-i}\in T_{-i}} \mu_i[\omega, t_{-i}, s_{-i}]$.

These together imply that $s_i \in \mathbf{S}^n_i(t_i) \cap R_i \neq \emptyset$, which completes the proof. $\square$

A.2.2 Proof of Proposition 8

We prove this proposition by showing the following: For each $i = 1, 2$ and $\theta_i \in \Theta_i$, all minimal elements in $\mathcal{R}^{\uparrow}_i(\theta_i)$ are singletons. Once this is claim is established, Proposition 8 is then implied by Proposition 6. Throughout the proof, denote $S_1 = M$ and $S_2 = A^M$.

By Assumption 1, for each m and ω^k_o, there is a unique response $a(\omega^k_o, m) \in A$ such that $\{a(\omega^k_o, m)\} = \arg\max_{a'\in A} u_2(\omega^k_o, m, a')$. For each $\theta^k_2 \in \Theta_2$, it is immediate that $\overline{\mathcal{R}}^{\uparrow,n}_2(\theta^k_2) = \{s^k_2\}$ for all $n \geq 1$, where $s^k_2 : m \mapsto a(\omega^k_o, m)$.

Claim 3. *For every $\theta^k_1 \in \Theta_1$ and every $R_1 \in \overline{\mathcal{R}}^{\uparrow,0}_1(\theta^k_1)$, there exists an $m^{(k,R_1)} \in R_1$ such that $\{m^{(k,R_1)}\} \in \overline{\mathcal{R}}^{\uparrow,2}_1(\theta^k_1)$. Moreover, for every $m \in M$, there exists a $\theta^{k(m)}_1 \in \Theta_1$ such that $\{m\} \in \overline{\mathcal{R}}^{\uparrow,2}_1(\theta^{k(m)}_1)$.*

Proof. For each $\theta^k_1 \in \Theta_1$, let $\bar{\nu}_1 \in \Delta(\theta^k_1 \times \mathcal{S}_2)$ such that $\bar{\nu}_1[(\omega^k_o, \{a(\omega^k_o, m)\})] = 1$. We then have

(i) $\bar{\nu}_1\big[\big\{(\omega, R_2) : (\theta_2(\omega), R_2) \in \text{graph}\,\big(\overline{\mathcal{R}}^{\uparrow,1}_2\big)\big\}\big] = 1$;

(ii) $\bigcup_{\pi_1(\phi)\in \mathrm{C}_1(\bar{\nu}_1\,|\,\theta^k_1)} r_1(\pi_1\,|\,\theta^k_1) = \arg\max_{m\in M} u_1(\omega^k_o, m, a(\omega^k_o, m)) \equiv \{m^k\}$.

Condition (ii) comes from part (i) of Assumption 1. Hence, we have $\{m^k\} \in \overline{\mathcal{R}}^{\uparrow,2}_1(\theta^k_1)$ for every $\theta^k_1 \in \Theta_1$. The first part of the claim is then proved by noting that, by definition, $\overline{\mathcal{R}}^{\uparrow,0}_1(\theta^k_1) = \{S_1\}$ and $m^k \in S_1$ for every $\theta^k_1 \in \Theta_1$. For the second part, note that by part (ii) of Assumption 1, for each message $m \in M$, there exists a state $\omega^{k(m)}_o$ such that $\arg\max_{m\in M} u_1(\omega^{k(m)}_o, m, a(\omega^{k(m)}_o, m)) = \{m\}$. Then letting $\theta^k_1 = \theta^{k(m)}_1$ yields $m^k = m$ in the construction above and thus $\{m\} \in \overline{\mathcal{R}}^{\uparrow,2}_1(\theta^{k(m)}_1)$. $\square$

Claim 4. *For each even $n \geq 0$, if for every $\theta^k_1 \in \Theta_1$ and every $R_1 \in \overline{\mathcal{R}}^{\uparrow,n}_1(\theta^k_1)$, there exists an $m^{(k,R_1)} \in R_1$ such that $\{m^{(k,R_1)}\} \in \overline{\mathcal{R}}^{\uparrow,n+2}_1(\theta^k_1)$, then for every $R_2 \in \overline{\mathcal{R}}^{\uparrow,n+1}_2(\theta^p_2)$, there exists an $s^{R_2}_2 \in R_2$ such that $\{s^{R_2}_2\} \in \overline{\mathcal{R}}^{\uparrow,n+3}_2(\theta^p_2)$*

Proof. Take an arbitrary $R_2 \in \overline{\mathcal{R}}^{\uparrow,n+1}_2(\theta^p_2)$. By definition, there is a $\bar{\nu}_2 \in \Delta(\theta^p_2 \times \mathcal{S}_1)$ such that

(i) $\bar{\nu}_2\big[\big\{(\omega, R_1) : (\theta_1(\omega), R_1) \in \text{graph}\,\big(\overline{\mathcal{R}}^{\uparrow,n}_1\big)\big\}\big] = 1$;

(ii) $R_2 \supseteq \bigcup_{\pi_2(\phi)\in \mathrm{C}_2(\bar{\nu}_2\,|\,\theta^p_2)} r_2(\pi_2\,|\,\theta^p_2)$.

By assumption, for every $\omega_p^k \in \theta_2^p$ and $R_1 \in \overline{\mathcal{R}}_1^{\uparrow,n}(\theta_1^k)$, there exists an $m^{(k,R_1)} \in R_1$ such that $\{m^{(k,R_1)}\} \in \overline{\mathcal{R}}_1^{\uparrow,n+2}(\theta_1^k)$. For each $m \neq M$, fix $\omega_p^{k(m)}$ to be a state such that $\{m\} \in \overline{\mathcal{R}}_1^{\uparrow,2}(\theta_1(\omega_p^{k(m)}))$. The existence of such a state is ensured by Claim 3. We now construct a sequence $(\overline{\nu}_2^\ell)_{\ell\geq 0}$ by letting

$$\overline{\nu}_2^*[\omega, \{m^{(k,R_1)}\}] = \overline{\nu}_2[\omega, R_1]$$

for all $(\omega, R_1) \in \theta_2^p \times \mathcal{S}_1$, and for every $\ell \geq 0$

$$\overline{\nu}_2^\ell = \frac{\ell}{\ell+1}\overline{\nu}_2^* + \frac{1}{|M|(\ell+1)}\sum_{m\in M}\delta_{(\omega_p^{k(m)},\{m\})},$$

where $\delta_{(\omega_p^{k(m)},\{m\})}$ denotes the Dirac measure on $(\omega_p^{k(m)}, \{m\})$. With vanishing weights, the summation of the Dirac measures breaks ties both on path and off path. Letting π_2^ℓ denote the unique conjecture such that $\pi_2^\ell(\phi) \in \mathrm{C}_2(\overline{\nu}_2^\ell \,|\, \theta_2^p)$, then by Assumption 1, for sufficiently large ℓ, there exists a strategy of the receiver $s_2^{R_2} \in S_2$ such that $\{s_2^{R_2}\} = r_2(\pi_2^\ell \,|\, \theta_2^p)$. Note that $\overline{\nu}_2^\ell\big[\big\{(\omega, R_1) : (\theta_1(\omega), R_1) \in \operatorname{graph}\big(\overline{\mathcal{R}}_1^{\uparrow,n+2}\big)\big\}\big] = 1$, which means $\{s_2^{R_2}\} \in \overline{\mathcal{R}}_2^{\uparrow,n+3}(\theta_2^p)$. Moreover, as $\ell \to \infty$, π_2^ℓ converges to some conjecture π_2^* such that $\pi_2^*(\phi) \in \mathrm{C}_2(\overline{\nu}_2 \,|\, \theta_2^p)$. By upper hemicontinuity of the best response correspondence $r_2(\cdot \,|\, \theta_2^p)$, we conclude that $s_2^{R_2} \in R_2$. □

Claim 5. *For each odd $n \geq 0$, if for every $R_2 \in \overline{\mathcal{R}}_2^{\uparrow,n}(\theta_2^p)$, there exists an $s_2^{R_2} \in R_2$ such that $\{s_2^{R_2}\} \in \overline{\mathcal{R}}_2^{\uparrow,n+2}(\theta_2^p)$, then for every $\theta_1^k \in \Theta_1$ and every $R_1 \in \overline{\mathcal{R}}_1^{\uparrow,n+1}(\theta_1^k)$, there exists an $m^{(k,R_1)} \in R_1$ such that $\{m^{(k,R_1)}\} \in \overline{\mathcal{R}}_1^{\uparrow,n+3}(\theta_1^k)$.*

Proof. For every $\theta_1^k \in \Theta_1$, take an arbitrary $R_1 \in \overline{\mathcal{R}}_1^{\uparrow,n+1}(\theta_1^k)$. By definition, there is a $\overline{\nu}_1 \in \Delta(\theta_1^k \times \mathcal{S}_2)$ such that

(i) $\overline{\nu}_1\big[\big\{(\omega, R_2) : (\theta_2(\omega), R_2) \in \operatorname{graph}\big(\overline{\mathcal{R}}_2^{\uparrow,n}\big)\big\}\big] = 1$;

(ii) $R_1 \supseteq \bigcup_{\pi_1(\phi)\in \mathrm{C}_1(\overline{\nu}_1 \,|\, \theta_1^k)} r_1(\pi_1 \,|\, \theta_1^k)$.

By assumption, for every $R_2 \in \overline{\mathcal{R}}_2^{\uparrow,n}(\theta_2^p)$, there exists an $s_2^{R_2} \in R_2$ such that $\{s_2^{R_2}\} \in \overline{\mathcal{R}}_2^{\uparrow,n+2}(\theta_2^p)$. Define a sequence $(\overline{\nu}_1^\ell)_{\ell\geq 0}$ by letting $\overline{\nu}_1^*[\omega_o^k, \{s_2^k\}] = \overline{\nu}_1[\omega_o^k, \{s_2^k\}]$ and $\overline{\nu}_1^*[\omega_p^k, \{s_2^{R_2}\}] = \overline{\nu}_1[\omega_p^k, R_2]$ for every $R_2 \in \mathcal{S}_2$., and for every $\ell \geq 0$

$$\overline{\nu}_1^\ell = \frac{\ell}{\ell+1}\overline{\nu}_1^* + \frac{1}{(\ell+1)}\delta_{(\omega_o^k,\{s_2^k\})}.$$

Again, the second vanishing term is used to break potential ties induced by $\overline{\nu}_1^*$. Denoting by π_1^ℓ the unique conjecture such that $\pi_1^\ell(\phi) \in \mathrm{C}_1(\overline{\nu}_1^\ell \,|\, \theta_1^k)$, then by Assumption 1, for sufficiently large ℓ, there exists a message of the sender $m^{(k,R_1)} \in S_1$ such that $\{m^{(k,R_1)}\} = r_1(\pi_1^\ell \,|\, \theta_1^k)$. Because $\overline{\nu}_1^\ell\big[\big\{(\omega, R_2) : (\theta_2(\omega), R_2) \in \operatorname{graph}\big(\overline{\mathcal{R}}_2^{\uparrow,n+2}\big)\big\}\big] = 1$, we have $\{m^{(k,R_1)}\} \in \overline{\mathcal{R}}_1^{\uparrow,n+3}(\theta_1^k)$.

Moreover, as $\ell \to \infty$, π_1^ℓ converges to some conjecture π_1^* such that $\pi_1^*(\phi) \in \mathrm{C}_2(\overline{\nu}_1 \,|\, \theta_1^k)$. By upper hemicontinuity of the best response correspondence $r_1(\cdot \,|\, \theta_1^k)$, we conclude that $m^{(k,R_1)} \in R_1$. □

Because each sequence $(\overline{\mathcal{R}}_i^{\uparrow,n}(\theta_i))_{n\geq 0}$ is increasing in n and converges in finitely many rounds, the preceding claims together imply that for every $i = 1, 2$, $\theta_i \in \Theta_i$, and $R_i \in \overline{\mathcal{R}}_i^{\uparrow}(\theta_i)$, there exists some $s_i \in R_i$ such that $\{s_i\} \in \overline{\mathcal{R}}_i^{\uparrow}(\theta_i)$. Invoking Proposition 6 completes the proof.

A.2.3 Proof of Proposition 9

To prove part (i), it is sufficient to establish the following and then invoke Proposition 3: for every $i = 1, 2$ and $R_i \in \mathcal{R}_i^{\mathrm{loc}}(t_i^0)$, either $a_i^i \in R_i$ or $a_i^j \in R_i$. We first prove the following claim.

Claim 6. *For each $i = 1, 2$, $\mathcal{P}_i \in \Pi_i$, and $R_j \in \overline{\mathcal{R}}_j^{\uparrow}(\theta_j^{\mathcal{P}_i})$, there exists a strategy $s_j \in S_j(\theta_j^{\mathcal{P}_i})$ such that $s_j(\mathcal{P}_i(a_i^i)) = a_j^i$ and $s_j \in R_j$. Moreover, for each $i = 1, 2$ and $R_i \in \overline{\mathcal{R}}_i^{\uparrow}(\theta_i^0)$, we have $a_i^i \in R_i$ or $a_i^j \in R_i$.*

Proof. We proceed by induction. For each $i = 1, 2$ and $j \neq i$, the statement clearly holds for $\overline{\mathcal{R}}_i^{\uparrow,0}(\theta_i^0)$ and $\overline{\mathcal{R}}_j^{\uparrow,0}(\theta_j^{\mathcal{P}_i})$ by definition. Suppose it holds at step $n-1$.

We first consider $\overline{\mathcal{R}}_i^{\uparrow,n}(\theta_i^0)$. Let $\overline{\nu}_i \in \Delta\big(\bigcup_{\omega \in \theta_i^0}\{\omega\} \times \mathcal{S}_j(\theta_j(\omega))\big)$ satisfy $\overline{\nu}_i\big[\big\{(\omega, R_j) : (\theta_j(\omega), R_j) \in \operatorname{graph}\big(\overline{\mathcal{R}}_j^{\uparrow,n-1}\big)\big\}\big] = 1$. Define a function $\bar{f}_i$ as follows: (i) $\bar{f}_i(\omega_i^{\mathcal{P}_i}, R_j)[s_j] = 1$ for some $s_j \in S_j(\theta_j^{\mathcal{P}_i})$ such that $s_j(\mathcal{P}_i(a_i^i)) = a_j^i$ and (ii) $\bar{f}_i(\omega_0, R_j)[a_j^i] = 1$ if $a_j^i \in R_j$ and $\bar{f}_i(\omega_0, R_j)[a_j^j] = 1$ if $a_j^i \notin R_j$. This function is well defined by the induction hypothesis.

Consider the belief $\lambda_i \in \mathrm{C}_i(\overline{\nu}_i \,|\, \theta_i^0)$ induced by $\bar{f}_i$. If λ_i attaches high enough probability to state ω_0 and the opponent playing a_j^j, then the best response is a_i^j; otherwise, the opponent plays a_j^i with high probability and the best response is a_i^i. No other action can be optimal, because players strictly prefer any Nash equilibrium of $\hat{\Gamma}$ to any non-equilibrium profile and are indifferent over non-equilibrium profiles. Hence, for every $i = 1, 2$ and $R_i \in \overline{\mathcal{R}}_i^{\uparrow,n}(\theta_i^0)$, either $a_i^i \in R_i$ or $a_i^j \in R_i$.

Next consider $\overline{\mathcal{R}}_j^{\uparrow,n}(\theta_j^{\mathcal{P}_i})$. Let $\overline{\nu}_j \in \Delta\big(\{\omega_i^{\mathcal{P}_i}\} \times \mathcal{S}_j(\theta_j^{\mathcal{P}_i})\big)$ satisfy $\overline{\nu}_j\big[\big\{(\omega_i^{\mathcal{P}_i}, R_i) : (\theta_i^0, R_i) \in \operatorname{graph}\big(\overline{\mathcal{R}}_i^{\uparrow,n-1}\big)\big\}\big] = 1$. We can define a function $\bar{f}_j$ by $\bar{f}_j(\omega_i^{\mathcal{P}_i}, R_i)[a_i^j] = 1$ if $a_i^j \in R_i$ and $\bar{f}_i(\omega_i^{\mathcal{P}_i}, R_i)[a_i^i] = 1$ if $a_i^j \notin R_i$.

For the belief $\lambda_j \in \mathrm{C}_j(\overline{\nu}_j \,|\, \theta_j^{\mathcal{P}_i})$ induced by $\bar{f}_i$, it attaches probability one either to a_i^j or to a_i^i. In the latter case, the sequentially best response at information set $\mathcal{P}_i(a_i^i)$ satisfies $s_j(\mathcal{P}_i(a_i^i)) = a_j^i$. In the former case, this information set is initially assigned zero probability; upon reaching it, beliefs may be revised in a way that a_i^i is played and therefore the best response still satisfies $s_j(\mathcal{P}_i(a_i^i)) = a_j^i$. Thus, for any $R_j \in \overline{\mathcal{R}}_j^{\uparrow,n}(\theta_j^{\mathcal{P}_i})$, there exists a strategy $s_j \in S_j(\theta_j^{\mathcal{P}_i})$ such that $s_j(\mathcal{P}_i(a_i^i)) = a_j^i$ and $s_j \in R_j$. □

Claim 7. *For each $i = 1, 2$ and $R_i \in \mathcal{R}_i^{\text{loc}}(t_i^0)$, we have $a_i^i \in R_i$ or $a_i^j \in R_i$.*

Proof. We prove by induction again. By the previous lemma, we have $a_i^i \in R_i$ or $a_i^j \in R_i$ for every $R_i \in \mathcal{R}_i^{\text{loc},0}(t_i^0)$. Suppose the statement holds at step $n-1$, and consider an arbitrary $R_i \in \mathcal{R}_i^{\text{loc},n}(t_i^0)$. Let $\nu_i \in \Delta(\bigcup_{\omega \in \theta_i^0} \{\omega\} \times \{t_j^0\} \times \mathcal{S}_j(\theta_j(\omega))$ be any distribution supporting R_i as $\varepsilon \to 0$ in Definition (3). We can define a function f_i by $f_i(\omega_0, t_j^0, R_j)[a_j^i] = 1$ if $a_j^i \in R_j$ and $f_i(\omega_0, t_j^0, R_j)[a_j^j] = 1$ if $a_j^i \notin R_i$. It follows that $a_i^i \in R_i$ in the former case and a_i^j in the latter. This completes the proof. □

We now turn to part (ii) of Proposition 9. It suffices to prove the following claim.

Claim 8. *If there exists a partition $\mathcal{P}_i \in \Pi_i$ such that $\{a_i^i\} \in \mathcal{P}_i$ for both $i = 1, 2$, then for each $i = 1, 2$, $\{\{a_i^i\}, \{a_i^j\}\} \subseteq \mathcal{R}_i^{\text{loc}}(t_i^0)$.*

Proof. For each $i = 1, 2$ and any $s_i \in \overline{\mathcal{R}}_i^{\uparrow,1}(\theta_i^{\mathcal{P}_j})$ with $\{a_j^j\} \in \mathcal{P}_j$, we must have $s_i(\{a_j^j\}) = a_i^j$ by sequential rationality. Therefore, in Definition (1), we can pick a distribution $\overline{\nu}_i$ that assigns probability one to $\omega_j^{\mathcal{P}_j}$, under which a_i^i is the unique best response. That is, $\{a_i^i\} \in \overline{\mathcal{R}}_i^{\uparrow,2}(\theta_i^0)$ for each $i = 1, 2$. In the next step, choosing a distribution $\overline{\nu}_i$ assigning probability one to (ω_0, a_j^j) makes a_i^j as the unique best response. Hence, $\{a_i^j\} \in \overline{\mathcal{R}}_i^{\uparrow,3}(\theta_i^0)$ for each $i = 1, 2$. Consequently, the singleton sets $\{a_i^i\}$ and $\{a_i^j\}$ both belong to $\overline{\mathcal{R}}_i^{\uparrow}(\theta_i^0) = \mathcal{R}_i^{\text{loc},0}(t_i^0)$.

Now suppose that $\{\{a_i^i\}, \{a_i^j\}\} \subseteq \mathcal{R}_i^{\text{loc},n-1}(t_i^0)$ for each $i = 1, 2$. In Definition (3), for any $\varepsilon \in (0, 1]$, letting $\nu_i[\omega_0, t_j^0, \{a_j^j\}] = \overline{\nu}_i[\omega_0, \{a_j^j\}] = 1$ clearly yields a_i^j as the unique best response; similarly, letting $\nu_i[\omega_0, t_j^0, \{a_j^i\}] = \overline{\nu}_i[\omega_0, \{a_j^i\}] = 1$ yields a_i^i as the unique best response. This means $\{\{a_i^i\}, \{a_i^j\}\} \subseteq \mathcal{R}_i^{\text{loc},n}(t_i^0)$, which completes the proof. □

The claim above implies that any proper refinement of the prediction $P(t_i^0) = \{a_i^i, a_i^j\}$ for each i violates the condition in Proposition 3 and therefore fails to be robust. Part (ii) follows.

Finally, part (iii) follows from the claim below. Apart from the asymmetry, the proof proceeds along the same lines as those of the three preceding claims, and we therefore omit the details.

Claim 9. *If there exists a partition $\mathcal{P}_i \in \Pi_i$ such that $\{a_i^i\} \in \mathcal{P}_i$ and $\Pi_j = \emptyset$, then $a_i^i \in R_i$ for all $R_i \in \mathcal{R}_i^{\text{loc}}(t_i^0)$ and $a_j^i \in R_j$ for all $R_j \in \mathcal{R}_j^{\text{loc}}(t_j^0)$. Moreover, $\{a_i^i\} \in \mathcal{R}_i^{\text{loc}}(t_i^0)$ and $\{a_j^i\} \in \mathcal{R}_j^{\text{loc}}(t_j^0)$.*

References

Basu, K. and J. W. Weibull (1991): "Strategy Subsets Closed under Rational Behavior," *Economic Letters*, 36, 141–146.

Battigalli, P. (1993): "Strategic Independence and Perfect Bayesian Equilibria," Tech. rep., Dipartimento di Economia e Produzione Rapporto Interno 93-011, Politecnico di Milano.

——— (1996): "Strategic Rationality Orderings and the Best Rationalization Principle," *Games and Economic Behavior*, 13, 178–200.

——— (1997): "On Rationalizability in Extensive Games," *Journal of Economic Theory*, 74, 40–61.

Battigalli, P. and M. Siniscalchi (1999): "Interactive Beliefs, Epistemic Independence and Strong Rationalizability," *Research in Economics*, 53, 247–273.

——— (2002): "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory*, 106, 356–391.

——— (2003): "Rationalization and Incomplete Information," *Advances in Theoretical Economics*, 3, Article 3.

——— (2007): "Interactive Epistemology in Games with Payoff Uncertainty," *Research in Economics*, 61, 165–184.

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

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

Carlsson, H. and E. van Damme (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61, 989–1018.

Chen, Y.-C. (2012): "A Structure Theorem for Rationalizability in the Normal Form of Dynamic Games," *Games and Economic Behavior*, 75, 587–597.

Chen, Y.-C., A. Di Tillio, E. Faingold, and S. Xiong (2010): "Uniform Topologies on Types," *Theoretical Economics*, 5, 445–478.

Chen, Y.-C., S. Takahashi, and S. Xiong (2014): "The Robust Selection of Rationalizability," *Journal of Economic Theory*, 151, 448–475.

——— (2022): "Robust Refinement of Rationalizability with Arbitrary Payoff Uncertainty," *Games and Economic Behavior*, 136, 485–504.

Cho, I.-K. and D. M. Kreps (1987): "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102, 179–222.

Cooper, R., D. V. DeJong, R. Forsythe, and T. W. Ross (1993): "Forward Induction in the Battle-of-the-Sexes Games," *American Economic Review*, 83, 1303–1316.

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

Dekel, E., D. Fudenberg, and S. Morris (2006): "Topologies on Types," *Theoretical Economics*, 1, 275–309.

Doval, L. and J. C. Ely (2020): "Sequential Information Design," *Econometrica*, 88, 2575–2608.

Ely, J. C. and M. Peski (2011): "Critical Types," *Review of Economic Studies*, 78, 907–937.

Fudenberg, D., D. M. Kreps, and D. K. Levine (1988): "On the Robustness of Equilibrium Refinements," *Journal of Economic Theory*, 44, 354–380.

Fudenberg, D. and J. Tirole (1991): *Game Theory*, Cambridge, Mass. and London: The MIT Press.

Germano, F., J. Weinstein, and P. Zuazo-Garin (2020): "Uncertain Rationality, Depth of Reasoning and Robustness in Games with Incomplete Information," *Theoretical Economics*, 15, 89–122.

Harsanyi, J. C. (1981): "Solutions for Some Bargaining Games under the Harsanyi–Selten Solution Theory, Part II: Analysis of Specific Bargaining Games," *Mathematical Social Sciences*, 3, 259–279.

Heifetz, A. and W. Kets (2018): "Robust Multiplicity with a Grain of Naiveté," *Theoretical Economics*, 13, 415–465.

Kajii, A. and S. Morris (1997): "The Robustness of Equilibria to Incomplete Information," *Econometrica*, 65, 1283–1309.

Kalai, E. (2004): "Large Robust Games," *Econometrica*, 72, 1631–1665.

Kalai, E. and D. Samet (1984): "Persistent Equilibria in Strategic Games," *International Journal of Game Theory*, 13, 129–144.

Kuhn, H. W. (1953): "Extensive Games and the Problem of Information," in *Contributions to the Theory of Games*, Princeton, NJ: Princeton University Press, vol. 2, 193–216.

Makris, M. and L. Renou (2023): "Information Design in Multistage Games," *Theoretical Economics*, 18, 1475–1509.

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

Morris, S. and H. S. Shin (1998): "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks," *American Economic Review*, 88, 587–597.

——— (2000): "Rethinking Multiple Equilibria in Macroeconomic Modeling," *NBER Macroeconomics Annual*, 15, 139–161.

Morris, S. and T. Ui (2005): "Generalized Potentials and Robust Sets of Equilibria," *Journal of Economic Theory*, 124, 45–78.

Myerson, R. B. (1986): "Multistage Games with Communication," *Econometrica*, 54, 323–358.

Osborne, M. J. and A. Rubinstein (1994): *A Course in Game Theory*, Cambridge, Mass. and London: The MIT Press.

Oury, M. and O. Tercieux (2012): "Continuous Implementation," *Econometrica*, 80, 1605–1637.

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

Penta, A. (2010): "Higher Order Beliefs in Dynamic Environments," Tech. rep., University of Wisconsin–Madison.

——— (2012): "Higher Order Uncertainty and Information: Static and Dynamic Games," *Econometrica*, 80, 631–660.

——— (2013): "On the Structure of Rationalizability for Arbitrary Spaces of Uncertainty," *Theoretical Economics*, 8, 405–430.

Penta, A. and P. Zuazo-Garin (2022): "Rationalizability, Observability and Common Knowledge," *Review of Economic Studies*, 89, 948–975.

Piermont, E. and P. Zuazo-Garin (2026): "Misspecified Information in Dynamic Games," *Working paper.*

Reny, P. J. (1992): "Backward Induction, Normal Form Perfection and Explicable Equilibria," *Econometrica*, 60, 627–649.

Reny, P. J. and A. J. Robson (2004): "Reinterpreting Mixed Strategy Equilibria: A Unification of the Classical and Bayesian Views," *Games and Economic Behavior*, 48, 355–384.

Rényi, A. (1955): "On a New Axiomatic Theory of Probability," *Acta Mathematica Academiae Scientiarum Hungaricae*, 6, 285–335.

Robson, A. J. (1994): "An "Informationally Robust Equilibrium" for Two-Person Nonzero-Sum Games," *Games and Economic Behavior*, 7, 233–245.

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

Solan, E. and L. Yariv (2004): "Games with Espionage," *Games and Economic Behavior*, 47, 172–199.

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

——— (2011): "Sensitivity of Equilibrium Behavior to Higher-Order Beliefs in Nice Games," *Games and Economic Behavior*, 72, 288–300.

——— (2013): "Robust Predictions in Infinite-Horizon Games—an Unrefinable Folk Theorem," *Review of Economic Studies*, 80, 365–394.

Ziegler, G. (2022): "Informational Robustness of Common Belief in Rationality," *Games and Economic Behavior*, 132, 592–597.

Zuazo-Garin, P. (2017): "Uncertain Information Structures and Backward Induction," *Journal of Mathematical Economics*, 71, 135–150.