Chen, Y-C., Di Tillio, A., Faingold, E., & Xiong, S. (2017). Characterizing the Strategic Impact of Misspecified Beliefs. *Review of Economic Studies*, *84*(4), 1424-1471. https://doi.org/10.1093/restud/rdw061

Characterizing the Strategic Impact of Misspecified Beliefs*

Yi-Chun Chen[†] Alfredo Di Tillio[‡] Eduardo Faingold[§] Siyang Xiong[¶]

August 11, 2016

Abstract

Previous research has established that the predictions of game theory are quite sensitive to the assumptions made about the players' beliefs. We evaluate the severity of this robustness problem by characterizing conditions on the primitives of the model—the players' beliefs and higher-order beliefs about the payoff-relevant parameters—for the behavior of a given Harsanyi type to be approximated by the behavior of (a sequence of) perturbed types. This amounts to providing belief-based characterizations of the strategic topologies of Dekel, Fudenberg, and Morris (2006). We apply our characterizations to a variety of questions concerning robustness to perturbations of higher-order beliefs, including genericity of types that are consistent with a common prior, and we investigate the connections between our notions of robustness and the notion of common p-belief of Monderer and Samet (1989).

Keywords: Games with incomplete information, rationalizability, higher-order beliefs, robustness

JEL Classification: C70, C72.

*We thank the editor and three anonymous referees for their comments and suggestions, which greatly improved the paper. We are especially grateful to Eddie Dekel, Jeffrey Ely, Stephen Morris and Marcin Pęski for numerous stimulating discussions and insightful comments. We also thank Pierpaolo Battigalli, Drew Fudenberg, Atsushi Kajii, Barton Lipman, George Mailath, Hervé Moulin, Daisuke Oyama, Dov Samet, Muhamet Yildiz, and participants at various seminars and conferences for helpful comments. Eduardo Faingold and Siyang Xiong thank the National Science Foundation (grants SES-1227593 and SES-1227620) for financial support.

[†]Department of Economics, National University of Singapore, ecsycc@nus.edu.sg

[‡]Department of Economics and IGIER, Università Luigi Bocconi, alfredo.ditillio@unibocconi.it

[§]Department of Economics and Cowles Foundation, Yale University, eduardo.faingold@yale.edu

[¶]Department of Economics, University of Bristol, siyang.xiong@bristol.ac.uk

1 Introduction

A major concern with non-cooperative game theory is its reliance on details. The formal description of a strategic situation as a game requires informational assumptions that are often not verifiable in full detail by the analyst in real life, such as the players' beliefs about the precise order of moves, the actions available to the players when they move, and their exact payoff functions. Unfortunately, game theoretic solutions are known to depend sensitively on those assumptions. For example, in an exchange economy with uncertainty, where agents share a common prior on the underlying state of the world, there is no feasible trade that is commonly known to be mutually acceptable (Milgrom and Stokey, 1982), whereas such a trade can exist if the common prior and common knowledge assumptions are slightly violated (Morris, 1994). In auction environments, bidders with independent values retain information rents (Myerson, 1981), while the auctioneer can fully extract their surplus when their values are slightly correlated (Crémer and McLean, 1988).[1] In the alternating-offers bargaining model of Rubinstein (1982), players reach an immediate agreement in the unique (subgame perfect) equilibrium, but delay to agreement is possible in equilibrium when the players have heterogeneous beliefs about who will make an offer in each period (Yildiz, 2003). In all these disparate and prominent economic scenarios, a model that makes unwarranted informational assumptions may deliver predictions that are not robust.

What form of misspecification error can be allowed in a model of the players' beliefs, to ensure the model will deliver robust predictions across a wide range of economic situations? Focusing on strategic-form games with incomplete information, we take the point of view of the analyst who posits a type space (Harsanyi, 1967-68) to model the players' uncertainty, and recognizes that his model may be misspecified. For example, he may assume that there is common knowledge of the true payoff-relevant parameter, but understands that such common knowledge assumption can be at best an approximation of reality. Or, as is often the case in practice, he may posit a finite type space, or a type space with a common prior, but realizes that the true type space may be larger, or that the true common prior distribution may be slightly different from the one assumed, or even that the players may have slightly different priors. To analyze the impact of such kinds of misspecifications, we study the tail properties of the hierarchies of beliefs encoded in the Harsanyi types—a player's beliefs about the payoff-relevant parameters, his beliefs about the other players' beliefs about the payoff-relevant parameters, and so on, ad infinitum—and their implications for behavior. Our main finding is an exact characterization of what it takes for a pair of types to display similar strategic behaviors. Thus, we measure the minimum level of precision of the analyst's information model that is required for accurate predictions of strategic play.

To explain our results we first need to be precise about what we mean by "strategic behavior." Our behavioral assumption is that players play a Bayesian equilibrium on a type space (possibly

[1] Full surplus extraction has also been itself the subject of a debate concerning whether, and in what sense, it is a robust result. See Heifetz and Neeman (2006) and Chen and Xiong (2013).

without a common prior). Thus, from the perspective of the analyst, who does not know the true type space of the players and has a concern for robustness, the relevant solution concept is (interim correlated) rationalizability (Dekel, Fudenberg, and Morris, 2006). Indeed, the set of actions that are rationalizable for a type t coincides with the set of actions that can be played in some Bayesian equilibrium on some type space, by some type that has the same hierarchy of beliefs as t (Dekel, Fudenberg, and Morris, 2007, Remark 2). A similar perspective is taken by Bergemann and Morris (2009) in the context of robust mechanism design. See also Aumann (1987) and Brandenburger and Dekel (1987) for early papers pioneering this approach.

Formally, our main results are characterizations of the *strategic topology* and the *uniform strategic topology* of Dekel, Fudenberg, and Morris (2006). The former is the coarsest topology on the universal type space (Mertens and Zamir, 1985)—the space of all hierarchies of beliefs—under which the correspondence that maps each type of a player into his set of rationalizable actions displays the same kind of continuity properties that the best-reply, Nash equilibrium and rationalizability correspondences exhibit in complete information games.[2] Thus, for any player, a sequence of types t^n converges in the strategic topology to a type t if and only if, for every finite game and every action a of the player in the game, the following conditions are equivalent: (a) action a is rationalizable for type t; (b) for every $\varepsilon > 0$ and sufficiently large n, action a is ε-rationalizable for type t^n, where ε is a size of sub-optimization allowed in the incentive constraints. Convergence in the uniform strategic topology adds the requirement that the rate of convergence in (b) be uniform across all finite games (with uniformly bounded payoffs).

As shown by Dekel, Fudenberg, and Morris (2006), a sequence of types converges in the strategic topology only if it converges in the *product topology*: for every integer $k \geqslant 1$, the sequence of k-order beliefs must converge weakly. However, the Electronic Mail game of Rubinstein (1989) and, more generally, the structure theorem of Weinstein and Yildiz (2007), show that convergence in the product topology does not imply strategic convergence. Our characterizations are based on a strengthening of product convergence that requires k-order beliefs to converge at a rate that is uniform in k.

We first explain the characterization of the uniform strategic topology, as it is simpler to state and can serve as a benchmark for the other characterization result. For each k, endow the space of k-order beliefs with the Prohorov distance, a standard distance that metrizes the topology of weak convergence of probability measures (Billingsley, 1999). Say that a sequence of types t^n converges *uniform-weakly* to a type t if the k-order belief of t^n converges to the k-order belief of t and the rate of convergence is uniform in k. Our first main result, Theorem 1, states that uniform strategic convergence is equivalent to uniform weak convergence.[3] To interpret, suppose that the analyst would like to make predictions with some minimal level of accuracy, and moreover he

[2] See the introduction of Dekel, Fudenberg, and Morris (2006) for a precise analogy.

[3] The partial result that uniform weak convergence implies uniform strategic convergence was proved in Chen, Di Tillio, Faingold, and Xiong (2010). The reverse implication is new to the present paper.

wants to achieve this level of accuracy *uniformly* across all strategic situations that the players might face. A necessary and sufficient condition for such uniformly robust prediction is that the analyst's model of the players' beliefs and higher-order beliefs be sufficiently precise, with the required degree of precision, *as measured by the Prohorov distance*, binding uniformly over all levels of the belief hierarchy.

The content of Theorem 1 can be dissected in two parts. First, the theorem underscores the role of uniform convergence of hierarchies of beliefs as a requirement for robustness. In light of the structure theorem of Weinstein and Yildiz (2007), which shows that the tails of the hierarchies of beliefs can have a large impact on strategic behavior, the role of uniform convergence should not come as a surprise. Second, the theorem *quantifies* the impact of a misspecification at each order of the hierarchy by the Prohorov distance. We view this as a nontrivial part of the theorem. Indeed, the Prohorov distance, on which the notion of uniform weak convergence is based, is but one of many equivalent distances that metrize the topology of weak convergence of probability measures. For any such distance, one can consider the associated uniform distance over infinite hierarchies of beliefs. It turns out that these distances may generate different topologies over infinite hierarchies, even though the induced topologies over k-order beliefs coincide for each finite k.[4] Theorem 1 identifies one of these uniform distances that ultimately characterizes the uniform-strategic topology.

The characterization of the strategic topology (Theorem 2) is also based on uniform convergence and the Prohorov metric, but is more subtle. The relevant class of events for uniform weak convergence, and a fortiori, uniform strategic convergence, is the entire Borel σ-algebra of the universal type space. By contrast, our characterization of the strategic topology highlights the role of coarser information structures called *frames*. A frame is a profile of finite partitions of the universal type space—one partition for each player—that satisfies a measurability condition: each player's belief concerning the events in the frame must pin down a unique atom of that player's partition. (We discuss the meaning of this condition below.) For any frame $\mathcal{P}$ and any positive integer k, we define a distance over types, $d^k_{\mathcal{P}}$, that is analogous to the Prohorov distance over k-order beliefs, but restricts the events for which the proximity is measured to those in the frame $\mathcal{P}$. Say that a sequence of types t^n converges to a type t *uniform-weakly on* $\mathcal{P}$ if, for every positive integer k, t^n converges to t under $d^k_{\mathcal{P}}$ and the rate of convergence is uniform in k. Our second main result, Theorem 2, states that a sequence of types converges strategically if and only if it converges uniform-weakly on every frame.

A frame can be interpreted as a "self-contained" coarsening of the canonical information structure of the universal type space. Each event in the partition of player i must be measurable with respect to the partition generated by identifying types that share the same beliefs about the param-

[4]In Chen, Di Tillio, Faingold, and Xiong (2010) we report an example of a sequence of types that converges uniform weakly but fails to converge in the uniform topology associated with a distance (different from Prohorov) that metrizes the topology of weak convergence of probability measures.

eter and the events in the partition of $-i$. In other words, each event in the partition of player i can be interpreted as a statement regarding the belief of player i about the parameters and the events in the partition of player $-i$, which can in turn be interpreted as statements about player $-i$'s beliefs about the parameters and the events in player i's partition, and so on. If the strategic behavior of player i in a game is measurable with respect to player i's partition, it is intuitive that the strategic behavior of player $-i$ must in turn be measurable with respect to player $-i$'s partition. This is the fundamental intuition for why strategic convergence implies uniform weak convergence only on frames and not on all information structures.

To shed further light on the impact of higher-order beliefs, we use our characterization to investigate the connection between strategic convergence and a natural notion of uniform convergence based on *common p-beliefs* (Monderer and Samet, 1989).[5] Say that a sequence of types t^n *converges in common beliefs* to a type t if t^n converges to t in the product topology and, for every event E and every $p \geqslant 0$, the following conditions are equivalent: (i) E is common p-belief for type t; (ii) for every $\varepsilon > 0$, $k \geqslant 1$ and sufficiently large n, type t^n has common $(p - \varepsilon)$-belief on the event that the players have k-order beliefs that are ε-close to those from E. This is the interim analogue of the ex ante notion of convergence based on common p-beliefs that the seminal papers of Monderer and Samet (1996) and Kajii and Morris (1998) have shown to characterize the ex ante strategic topology for Bayesian equilibrium on countable common prior type spaces. We establish, as a corollary of our Theorem 2, that strategic convergence implies convergence in common beliefs (Corollary 1). However, somewhat surprisingly, we find that the converse fails. These results highlight a fundamental difference between the common prior, equilibrium, ex ante framework of the early literature and our non-common prior, non-equilibrium, interim framework. Nonetheless, when the limit is a *finite type*—a type that belongs to a finite type space—we show that convergence in common beliefs is equivalent to uniform weak convergence, and hence, a fortiori, to both uniform strategic and strategic convergence (Theorem 3).

In the last part of the paper we use our characterizations to examine the robustness of the most widely used models in applied game theory, namely those assuming types consistent with a *common prior*. Lipman (2003) shows that such types are dense in the universal type space under the product topology. We revisit, and reverse, this result in two ways. First, we show that the set of non-common prior types contains a set that is open and dense under the strategic topology (Theorem 4), that is, common prior types are nowhere dense under the strategic topology. Second, turning to an ex ante perspective, we show that when the set of all *type spaces* is endowed with the Hausdorff topology based on the *product* topology on types, there is an open and dense set of non-common prior type spaces. Moreover, the analogous statement holds when the set of all type spaces is endowed with the Hausdorff topology based on the strategic topology (Theorem 5).

To interpret these genericity results and illustrate their implications for economic modeling,

[5] An event E is common p-belief for a given type if that type assigns probability at least p to E, assigns probability at least p to the event that E obtains and the other players assign probability at least p to E, and so forth, ad infinitum.

consider an analyst who has imperfect knowledge of the players' hierarchy of beliefs, and posits a common prior type space. Given his limited knowledge, the analyst must deem it possible that the players' actual beliefs violate the common prior assumption, by the denseness of non-common prior types. In this case, openness implies that, in some strategic contexts, every model that makes sufficiently accurate predictions must be a non-common prior model. Thus, there exists no general principle on which the analyst can rely, that guarantees that predictions made under the common prior assumption will approximate the players' actual behavior. While there are strategic contexts where a common prior model does provide a good approximation, this needs to be assessed on a case-by-case basis.

The rest of the paper is organized as follows. Section 2 sets up the incomplete information model and reviews the basic definitions and properties of type spaces, hierarchies of beliefs, common p-beliefs and the solution concept of interim correlated rationalizability. Section 3 recalls the definitions of the strategic topology and the uniform strategy topology of Dekel, Fudenberg, and Morris (2006) and, after a review of the complete information benchmark, presents their characterizations in terms of beliefs. Section 3 studies the relationship of our characterizations with the notion of common p-belief of Monderer and Samet (1989). Section 4 examines the implications of our characterization theorems on finite and common prior models. Section 5 concludes with further discussion of the results and extensions.

2 Basic Setup

Given a measurable space X, we write $\Delta(X)$ for the space of probability measures on X, equipped with the σ-algebra generated by the sets of the form $\{\nu \in \Delta(X) : \nu(E) \geqslant p\}$, where $E \subseteq X$ is a measurable set and $p \in (0, 1]$. Unless otherwise stated, product spaces are endowed with the product σ-algebra, subspaces with the relative σ-algebra, and finite spaces with the discrete σ-algebra. With slight abuse of notation, we often denote one-point sets $\{x\}$ simply by x, and cylinders $E \times Y$ in product spaces $X \times Y$ by their bases $E \subseteq X$.

2.1 Games with Incomplete Information

In a game with incomplete information, payoffs depend on the profile of actions as well as on an exogenous parameter that may not be common knowledge among the players. The Bayesian approach to modeling games with incomplete information, due to Harsanyi (1967-68), requires the specification of a *type space* to model this possible lack of common knowledge. Such spaces describe the players' uncertainty about the payoff-relevant parameter, but also their uncertainty about each other's uncertainty about the parameter, and any higher-order uncertainty.

We consider two-player games with incomplete information and denote the set of players by

$I = \{1, 2\}$.[6] The space Θ of payoff-relevant parameters is assumed to be finite and to contain at least two points. A *game* is a profile $G = (A_i, g_i)_{i \in I}$ where A_i is a finite set of actions of player i and $g_i : A_i \times A_{-i} \times \Theta \to [-M, M]$ is his payoff function, where $M > 0$ is a uniform bound on payoffs that is fixed throughout.[7] A *type space* is a profile $T = (T_i, \mu_i)_{i \in I}$ where T_i is a measurable space of types of player i and $\mu_i : T_i \to \Delta(\Theta \times T_{-i})$ is a measurable function that associates, with each type t_i of player i, his belief $\mu_i(t_i)$ about the payoff-relevant parameters and the types of player $-i$. The details of how type spaces encode higher-order uncertainty are reviewed in section 2.3 below.

At this point, it is worthwhile digressing to discuss how the general formulation of type spaces above relates to the one commonly used in applications. Readers who are familiar with our formulation may skip the next two paragraphs without incurring any loss.

In most applications of games with incomplete information, uncertainty is modeled through the *standard interdependent values model*: the space of payoff-relevant parameters is assumed to have a product structure $\Theta = \times_i \Theta_i$; there is a probability distribution $f \in \Delta(\Theta)$ according to which the profile $\theta = (\theta_i)_{i \in I}$ is drawn; upon the realization of θ each player i learns his coordinate θ_i and forms a belief over Θ_{-i} by updating the prior distribution f via Bayes' formula. This model can be translated into our formulation as follows: for each $i \in I$, we have $T_i = \Theta_i$ and $\mu_i : T_i \to \Delta(\Theta \times T_{-i})$ is given by

$$\mu_i(t_i)(\theta, t_{-i}) = \begin{cases} f(t_i, t_{-i}) \Big/ \sum_{\theta_{-i} \in \Theta_{-i}} f(t_i, \theta_{-i}) & \text{if} \quad \theta = (t_1, t_2) \\ 0 & \text{if} \quad \theta \neq (t_1, t_2). \end{cases}$$

The standard interdependent values model is flexible in some dimensions: it nests both the private values model and the pure common values model as particular cases, and it allows types to be either independent or correlated. Moreover, the model can be easily generalized to one in which players have different prior distributions $f_i \in \Delta(\Theta)$, an extension that is desirable in some applications. Most of the mechanism design literature is built on the standard interdependent values model.

However, the standard interdependent values model is restrictive for the purpose of our robustness exercise. It embodies a strong common knowledge assumption: the belief of a type concerning the payoff relevant parameter only, called his *first-order belief*, completely pins down his joint belief concerning the parameter and the type of the other player. As a consequence, the *second-order belief* of a type, defined as his joint belief concerning the parameter and the first-order belief of the other player, is completely pinned down by his first-order belief, and likewise for all higher-order beliefs. In other words, all higher-order uncertainty is trivial in the standard interdependent values model. But, as discussed in the introduction, in this paper we are interested in quantifying the behavioral impact of perturbing the players' types, so it makes sense to allow

[6] We focus on two-player games for ease of notation. All our results extend to the general N-player case.

[7] Following standard notation, for each player i we let $-i$ designate the other player in I.

perturbations to be as general as possible, rather than constrained by common knowledge assumptions. In other words, while in many cases it might make perfect sense to adopt the interdependent values model for the unperturbed model, it is desirable to allow the perturbations to relax the common knowledge assumptions assumed in the unperturbed model, as the very nature of the exercise is to quantify the degree with which strategic behavior is sensitive to such common knowledge assumptions. For this reason, we adopt the general formulation of type spaces above.

2.2 Solution Concept

Our analysis is based on the solution concept of *interim correlated rationalizability* due to Dekel, Fudenberg, and Morris (2007), which is an extension of the complete information rationalizability of Bernheim (1984) and Pearce (1984) to games with incomplete information. The definition follows a recursive process of elimination of non-best-replies in which the elimination is performed on a type-by-type basis. An action is rationalizable for a type if it survives infinitely many rounds of elimination of non-best-replies; at each round an action is declared a non-best-reply for a type, if it is not a best-reply to any conjecture that is consistent with the type's beliefs and assigns probability one to each type of the other player choosing actions that have not been eliminated thus far. The formal definition below allows for ε-best replies, which leads to the concept of ε-rationalizability.

Let a game $G = (A_i, g_i)_{i \in I}$, a type space $T = (T_i, \mu_i)_{i \in I}$ and $\varepsilon \geqslant 0$ be given. For every $i \in I$ and $t_i \in T_i$ define the set of *0-order ε-rationalizable actions* as $R_i^0(t_i, G, T, \varepsilon) = A_i$. Proceeding recursively, for $k \geqslant 1$, define the set of *k-order ε-rationalizable actions* of type t_i, written $R_i^k(t_i, G, T, \varepsilon)$, as the set of actions $a_i \in A_i$ for which there exists a conjecture $\nu \in \Delta(\Theta \times T_{-i} \times A_{-i})$ satisfying the following three properties:

(a) Action a_i is an ε-best-reply to the conjecture: for every $a_i' \in A_i$,

$$\sum_{(\theta, a_{-i}) \in \Theta \times A_{-i}} \left[g_i(\theta, a_i, a_{-i}) - g_i(\theta, a_i', a_{-i}) \right] \nu\big(\theta \times T_{-i} \times a_{-i}\big) \geqslant -\varepsilon.$$

(b) The conjecture assigns probability one to the event that each type of player $-i$ plays an action that is $(k-1)$-order ε-rationalizable:

$$\nu\big(\Theta \times \big\{(t_{-i}, a_{-i}) : a_{-i} \in R_{-i}^{k-1}(t_{-i}, G, T, \varepsilon)\big\}\big) = 1.$$

(c) The conjecture is consistent with the belief of type t_i: for every $\theta \in \Theta$ and measurable $E_{-i} \subseteq T_{-i}$,

$$\mu_i(t_i)(\theta \times E_{-i}) = \nu(\theta \times E_{-i} \times A_{-i}).$$

The set of *ε-rationalizable actions* of type t_i, written $R_i(t_i, G, T, \varepsilon)$, is the set of actions that are k-order ε-rationalizable for every k:

$$R_i(t_i, G, T, \varepsilon) = \bigcap_{k \geqslant 1} R_i^k(t_i, G, T, \varepsilon).$$

2.3 Beliefs

Type spaces contain implicit descriptions of each player's beliefs about the payoff-relevant parameters, his beliefs about the other player's beliefs about the payoff-relevant parameters, and so on. Any such higher-order uncertainty can be modeled in a canonical type space, called *universal type space*, whose construction, due to Mertens and Zamir (1985), we now review. An equivalent formulation is found in Brandenburger and Dekel (1993).

A *first-order belief* of player i is a probability distribution over payoff-relevant parameters, that is, an element of $T_{i,1} = \Delta(\Theta)$. Recursively, for $k \geqslant 2$, a *k-order belief* of player i is a joint probability distribution over payoff-relevant parameters and $(k-1)$-order beliefs of the other player, that is, an element of $T_{i,k} = \Delta(\Theta \times T_{-i,k-1})$. A *hierarchy of beliefs* of player i is an infinite sequence in $T_{i,1} \times T_{i,2} \times \cdots$ satisfying the condition that higher-order beliefs are compatible with lower-order beliefs, that is, an element of the set

$$T_i^* := \left\{ (t_{i,1}, t_{i,2}, \ldots) \in \bigtimes_{k=1}^{\infty} T_{i,k} \ : \ \rho_{i,k}(t_{i,k+1}) = t_{i,k} \quad \forall k \geqslant 1 \right\},$$

where $\rho_{i,k}$ is the function that associates each $(k+1)$-order belief with its induced k-order belief.[8]

Given a type space $(T_i, \mu_i)_{i \in I}$, the hierarchy of beliefs $\tau_i(t_i) = (\tau_{i,1}(t_i), \tau_{i,2}(t_i), \ldots)$ induced by a type $t_i \in T_i$ is defined as follows. The first-order belief $\tau_{i,1}(t_i)$ is the marginal of $\mu_i(t_i)$ on Θ, that is, $\tau_{i,1}(t_i)(\theta) = \mu_i(t_i)(\theta \times T_{-i})$ for every $\theta \in \Theta$. For $k \geqslant 2$, recursively, the k-order belief $\tau_{i,k}(t_i)$ is defined by letting

$$\tau_{i,k}(t_i)\big(\theta \times E\big) = \mu_i(t_i)\big(\theta \times \tau_{-i,k-1}^{-1}(E)\big)$$

for every $\theta \in \Theta$ and measurable $E \subseteq T_{-i,k-1}$. Thus, every type from an arbitrary type space is naturally mapped to a hierarchy of beliefs.

Conversely, every hierarchy of beliefs can be viewed as a type: there exists a unique measurable function

$$\mu_i^* \colon T_i^* \to \Delta(\Theta \times T_{-i}^*)$$

[8] The first-order belief $\rho_{i,1}(t_{i,2})$ induced by a second-order belief $t_{i,2}$ is the marginal of $t_{i,2}$ on Θ. Proceeding recursively, the k-order belief induced by a $(k+1)$-order belief $t_{i,k+1}$ is the probability distribution $\rho_{i,k}(t_{i,k+1})$ on $\Theta \times T_{-i,k-1}$ such that $\rho_{i,k}(t_{i,k+1})(\theta \times E) = t_{i,k+1}(\theta \times \rho_{-i,k-1}^{-1}(E))$ for every $\theta \in \Theta$ and measurable $E \subseteq T_{-i,k-1}$.

with the property that the probability measure $\mu_i^*(t_{i,1}, t_{i,2}, \ldots)$ extends each of the measures $t_{i,1}, t_{i,2}, \ldots$ Moreover, μ_i^* is an isomorphism.[9] The type space $(T_i^*, \mu_i^*)_{i \in I}$ is called the *universal type space*.[10] To ease notation, for each event $E \subseteq \Theta \times T_{-i}$ we shall often write $\mu_i^*(E \mid t_i)$ instead of the more cumbersome $\mu_i^*(t_i)(E)$.

Dekel, Fudenberg, and Morris (2007) show that the set of ε-rationalizable actions of any type t_i from an arbitrary type space T is pinned down by the hierarchy of beliefs of t_i, that is,

$$R_i(t_i, G, T, \varepsilon) = R_i\big(\tau_i(t_i), G, T^*, \varepsilon\big).$$

This fact has an important consequence for the framing of our robustness exercise. It implies that we can, and henceforth will, identify types with their induced hierarchy of beliefs. We therefore drop the reference to the type space from the notation for ε-rationalizability, and write $R_i(t_i, G, \varepsilon)$ for the set of ε-rationalizable actions of type t_i. To further ease notation, for $\varepsilon = 0$ we denote $R_i(t_i, G, 0)$ simply by $R_i(t_i, G)$, and call it the set of *rationalizable actions* of type t_i.

2.4 Common Belief

The notion of *common belief*, due to Monderer and Samet (1989), is a useful tool to study the robustness of strategic behavior in many settings (Monderer and Samet, 1996; Kajii and Morris, 1997, 1998; Ely and Pęski, 2011). To define common belief in our context, we consider events that are subsets of the space $\Omega = \Theta \times T_1^* \times T_2^*$. Each point in Ω is a complete description of the "state of the world," in that it corresponds to a complete resolution of all the relevant uncertainty: the payoff-relevant parameter and the type of each player.

For each measurable set $E \subseteq \Omega$ and each type $t_i \in T_i^*$, let E_{t_i} designate the section of E over t_i:

$$E_{t_i} = \big\{(\theta, t_{-i}) \,:\, (\theta, t_1, t_2) \in E\big\},$$

which is a measurable set by standard arguments. For each $p \in [0, 1]$, the event that player i assigns probability at least p to E is

$$B_i^p(E) = \big\{t_i \in T_i^* \,:\, \mu_i^*(E_{t_i} \mid t_i) \geqslant p\big\},$$

which is also a measurable set.[11] Then, for each $\mathbf{p} = (p_1, p_2) \in [0, 1]^2$, define the event that E *is*

[9] That is, a measurable bijection with measurable inverse.

[10] The term universal is borrowed from category theory. It means that every abstract type space $(T_i, \mu_i)_{i \in I}$ is mapped into the universal type space (T_i^*, μ_i^*) by a unique *belief-preserving morphism*. In general, a belief-preserving morphism between arbitrary type spaces (T_i, μ_i) and (T_i', μ_i') is defined as a profile of measurable functions $\varphi_i : T_i \to T_i'$ such that $\mu_i'(\varphi(t_i))(\theta \times E) = \mu_i(t_i)(\theta \times \varphi_{-i}^{-1}(E))$ for every $\theta \in \Theta$, measurable $E \subseteq T_{-i}'$ and $t_i \in T_i$. The unique belief-preserving morphism that maps an arbitrary type space (T_i, μ_i) into the universal type space (T_i^*, μ_i^*) is the profile of mappings τ_i defined above.

[11] The measurability of $B_i^p(E)$ follows from the measurability of the map $t_i \mapsto \mu_i(E_{t_i} \mid t_i)$. The class of events

$$\big\{E \subseteq \Omega \,:\, E \text{ is a measurable set such that } t_i \mapsto \mu_i(E_{t_i} \mid t_i) \text{ is measurable}\big\}$$

mutual $\mathbf{p}$*-belief* as

$$B^{\mathbf{p}}(E) = \Theta \times B_1^{p_1}(E) \times B_2^{p_2}(E),$$

and the event that E *is common* $\mathbf{p}$*-belief* as

$$C^{\mathbf{p}}(E) \;=\; B^{\mathbf{p}}(E) \;\cap\; B^{\mathbf{p}}\big(E \cap B^{\mathbf{p}}(E)\big) \;\cap\; B^{\mathbf{p}}\big(E \cap B^{\mathbf{p}}\big(E \cap B^{\mathbf{p}}(E)\big)\big) \;\cap\; \cdots .$$

Then, define the event that E *is common* $\mathbf{p}$*-belief for player* i, written $C_i^{\mathbf{p}}(E)$, as the projection of $C^{\mathbf{p}}(E)$ on T_i^*, which is a measurable set because $C^{\mathbf{p}}(E)$ is a rectangle.[12] Thus, we have

$$C_i^{\mathbf{p}}(E) \;=\; B_i^{p_i}(E) \;\cap\; B_i^{p_i}\big(E \cap B_{-i}^{p_{-i}}(E)\big) \;\cap\; B_i^{p_i}\big(E \cap B_{-i}^{p_{-i}}\big(E \cap B_i^{p_i}(E)\big)\big) \;\cap\; \cdots .$$ [13]

Note that for $p_{-i} = 0$ we have $C_i^{\mathbf{p}}(E) = B_i^{p_i}(E)$, while for $\mathbf{p} = (1,1)$ the set $C_i^{(1,1)}(\theta) := C_i^{(1,1)}(\theta \times T_1^* \times T_2^*)$ contains only one type of player i.

Finally, common belief has the following well known fixed-point characterization:

$$C^{\mathbf{p}}(E) = B^{\mathbf{p}}(E \cap C^{\mathbf{p}}(E)) \qquad \text{and} \qquad C_i^{\mathbf{p}}(E) = B_i^{p_i}(E \cap C_{-i}^{\mathbf{p}}(E)). \tag{1}$$

3 Characterization of Robustness

To what kinds of misspecification of the players' uncertainty are the predictions of rationalizability robust? That is, given a player i and a type t_i, what sequences of perturbed types t_i^n converge to t_i in the sense that the strategic behaviors of t_i^n approach those of t_i? To formalize this question, we use the notions of strategic convergence proposed by Dekel, Fudenberg, and Morris (2006), which we now recall.

Definition 1 (Strategic topology)**.** A sequence of types t_i^n *converges strategically* to a type t_i if for every game G and every action a_i of player i in G, the following conditions are equivalent:

(a) a_i is rationalizable for t_i in G;

(b) for every $\varepsilon > 0$ there exists N such that for every $n \geqslant N$, a_i is ε-rationalizable for t_i^n in G.

The *strategic topology* is the topology of strategic convergence on T_i^*.[14]

---

can be readily verified to be a monotone class containing the algebra of finite disjoint unions of measurable rectangles, which generates the product σ-algebra on Ω. It follows that the map $t_i \mapsto \mu_i(E_{t_i} | t_i)$ is measurable for every measurable E.

[12] The definition of $C^{\mathbf{p}}(E)$ is analogous to the *common repeated belief* of Monderer and Samet (1996), which differs from the original definition of Monderer and Samet (1989). A similar definition appears in Ely and Pęski (2011) for the case where E is a rectangle. We allow the event E to be any measurable set.

[13] Recall that for notational convenience we often denote cylinders by their bases. Thus, here we identify $B_i^{p_i}(E)$ and $C_i^{\mathbf{p}}(E)$ with $\Theta \times B_i^{p_i}(E) \times T_{-i}^*$ and $\Theta \times C_i^{\mathbf{p}}(E) \times T_{-i}^*$, respectively.

[14] This definition follows Ely and Pęski (2011). The original definition of Dekel, Fudenberg, and Morris (2006) is stated a bit differently, but the two are equivalent, as shown in the working paper version of Ely and Pęski (2011).

The implication from (b) to (a) is a form of upper hemicontinuity of the rationalizable correspondence: if for every $\varepsilon > 0$ an action is ε-rationalizable for every type in the tail of the sequence, then the action is also rationalizable for the limit type. In particular, every action that is rationalizable for every type in the tail of the sequence remains rationalizable for the limit type. In non-technical terms, this means that the unperturbed model never fails to predict a behavior that is predicted by the perturbed model. It is not a strong requirement: Dekel, Fudenberg, and Morris (2006, Theorems 1 and 2) show that the implication from (b) to (a) (holding for every game and action) is equivalent to convergence in the *product topology*, that is, weak convergence of k-order beliefs for every $k \geqslant 1$.[15]

Similarly, the implication from (a) to (b) is a form of lower hemicontinuity: if an action is rationalizable for the limit type, then for every $\varepsilon > 0$ the action is ε-rationalizable for every type in the tail of the sequence. To interpret, it is easier to examine an equivalent condition that does not involve ε-rationalizability: every action that is *strictly* rationalizable for the limit type remains rationalizable for all types in the tail of the sequence.[16] In other words, the unperturbed model never makes predictions that are not valid for the perturbed model, unless those predictions rely on non-strict incentives (i.e. indifferences).[17] As the electronic mail game example of Rubinstein (1989) demonstrates, this condition is *not* implied by convergence in the product topology. Finally, if the implication from (a) to (b) holds for every action in every game, then the implication from (b) to (a) also holds (Dekel, Fudenberg, and Morris, 2006, Corollary 1).

The definition of strategic convergence allows the rate of convergence N to depend on the game G. An alternative, stronger notion of convergence, similar to the one adopted in the early literature on robustness (Monderer and Samet, 1989, 1996; Kajii and Morris, 1997, 1998), requires the rate of convergence N to be independent of G (across all games with uniformly bounded payoffs), which leads to the following definition.

Definition 2 (Uniform strategic topology)**.** A sequence of types t_i^n *converges uniform-strategically* to a type t_i if for every $\varepsilon > 0$ there exists N such that for every game G and every

[15] The definition of the product topology is recursive: endow $T_{i,1}$ with the Euclidean topology, and for each $k \geqslant 2$, endow $T_{i,k} = \Delta(\Theta \times T_{-i,k-1})$ with the topology of weak convergence of probability measures relative to the product topology on $\Theta \times T_{-i,k-1}$. A sequence of types t_i^n converges to t_i in the product topology if for every $k \geqslant 1$, the k-order belief of t_i^n converges to the k-order belief of t_i. Note that, since Θ is finite, the Borel σ-algebra of the product topology on T_i^* coincides with the σ-algebra assumed in the topology-free formulation of section 2.3. Likewise, the Borel σ-algebra on $\Delta(\Theta \times T_{-i}^*)$ generated by the topology of weak convergence of probability measures (induced by the product topology on T_{-i}^*) coincides with the σ-algebra of our topology-free formulation. Finally, if we endow each T_i^* with the product topology and $\Delta(\Theta \times T_{-i}^*)$ with the topology of weak convergence of probability measures, then T_i^* and $\Delta(\Theta \times T_{-i}^*)$ are compact metrizable, and μ_i^* is a homeomorphism (Mertens and Zamir, 1985).

[16] An action is strictly rationalizable if it is ε-rationaliable for some *negative* ε. In other words, all incentive constraints hold with a slack that is bounded away from zero. The working paper version of Ely and Pęski (2011) shows that the two definitions are equivalent.

[17] Imposing the condition also for rationalizable actions that rely on indifferences is way too strong, as it would lead to the discrete topology.

action a_i of player i in G, the following conditions are equivalent:

(a) a_i is rationalizable for t_i in G;

(b) for every $\varepsilon > 0$ and $n \geqslant N$, a_i is ε-rationalizable for t_i^n in G.

The *uniform strategic topology* is the topology of uniform strategic convergence on T_i^*.

The main question we address in the paper is:

Under what conditions on primitives (i.e., beliefs and higher-order beliefs) does a sequence of types t_i^n converge strategically, or uniform strategically, to a type t_i?

3.1 Complete Information Benchmark

Before turning to our general analysis, it is worth reviewing a particular case for which the robustness question above has a well-known, and simple, answer: the case of *complete information*, where some payoff-relevant parameter is common knowledge for the unperturbed type t_i. In this case, we have the following well known result:[18]

Proposition 1 (Convergence to complete information)**.** *Let $\theta \in \Theta$ and $t_i = C_i^{(1,1)}(\theta)$. For every sequence of types t_i^n, the following statements are equivalent:*

(i) for every $p \in (0,1)$ there exists N such that for every $n \geqslant N$, $t_i^n \in C_i^{(p,p)}(\theta)$;

(ii) t_i^n converges uniform-strategically to t_i;

(iii) t_i^n converges strategically to t_i.

The two canonical examples below illustrate the application of condition (i) as a criterion for strategic (and uniform strategic) approximation of common knowledge. In both examples, we assume that $\Theta = \{\theta_0, \theta_1\}$.

Example 1 (Public announcements with mistakes, Monderer and Samet, 1989)**.** One situation in which common knowledge of a payoff-relevant parameter arises is when there is an informed third party (assumed to be non-strategic) who publicly announces the true value of the parameter before a game is played. Moving slightly away from common knowledge, suppose that each player (independently) hears an erroneous announcement with a small probability $\varepsilon > 0$. Suppose also that both players assign prior probability $\frac{1}{2}$ to each parameter value. We thus have a type space with two types, $t_{i,0}^{\varepsilon}$ and $t_{i,1}^{\varepsilon}$, for each player i. Type $t_{i,0}^{\varepsilon}$ heard announcement "the parameter is

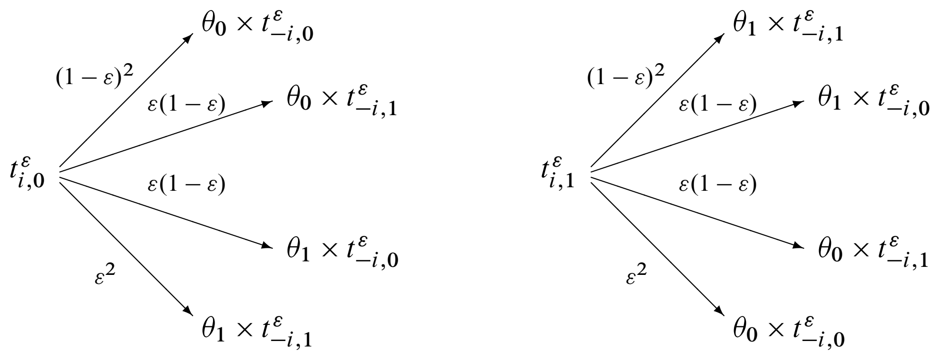

Figure 1: Type space for Example 1

θ_0," while type $t^{\varepsilon}_{i,1}$ heard announcement "the parameter is θ_1." Their beliefs, derived from Bayes' rule, are represented in the diagrams of Figure 1.

Thus, each type assigns probability $p_\varepsilon = (1-\varepsilon)^2$ to the event that both his own announcement and the other player's announcement are correct. As a result, for each player i we have $t^{\varepsilon}_{i,0} \in C_i^{(p_\varepsilon,p_\varepsilon)}(\theta_0)$ and $t^{\varepsilon}_{i,1} \in C_i^{(p_\varepsilon,p_\varepsilon)}(\theta_1)$. By Proposition 1, as $\varepsilon \to 0$, types $t^{\varepsilon}_{i,0}$ and $t^{\varepsilon}_{i,1}$ converge strategically to $t_{i,0} = C_i^{(1,1)}(\theta_0)$ and $t_{i,1} = C_i^{(1,1)}(\theta_1)$, respectively. This implies that, *irrespective of the payoff structure of the game*, the predictions of the model under common knowledge must be robust to the small amounts of incomplete information that arise in the public announcements protocol when players are subject to a small probability of hearing mistakes. ◆

Example 2 (Faulty message exchange, Rubinstein, 1989)**.** The players assign prior probability $\frac{1}{2}$ to θ_1. Player 1 then learns the true value of the parameter. If and only if it is θ_1, the players engage in an automated exchange. Player 1 sends a message to player 2. If the message arrives, player 2 automatically replies with a confirmation to player 1. If player 1 receives it, he automatically replies with a new message, and so on and so forth. Each time a message or confirmation is sent, it gets lost with probability $q > 0$, independently of all other events. The protocol gives rise to a countably infinite set of types $\{t_i^0, t_i^1, t_i^2, \ldots\}$ for each player i. Type t_1^0 of player 1 knows the parameter is θ_0 and, hence, no message was exchanged. Type t_2^0 of player 2 received no messages and is thus uncertain whether the parameter is θ_0 or θ_1, although he knows that in the first case no message was ever sent, while in the second exactly one was. For each $m \geqslant 1$, type t_1^m of player 1 knows the parameter is θ_1 and knows that exactly m messages were sent, so he is uncertain whether player 2 sent $m-1$ or m confirmations, while type t_2^m of player 2 knows the parameter is θ_1 and knows that exactly m confirmations were sent, so he is uncertain whether player 1 sent m or $m+1$ messages. The diagrams in Figure 2, where $q' = q/(1+q)$ and $q'' = 1/(2-q)$, illustrate the beliefs of the various types, which are derived from Bayes' rule.

[18] This equivalence is the interim analogue of the classic result of Monderer and Samet (1989), who characterize an ex ante notion of strategic approximation of common knowledge. The result is a corollary of Theorem 3 below.

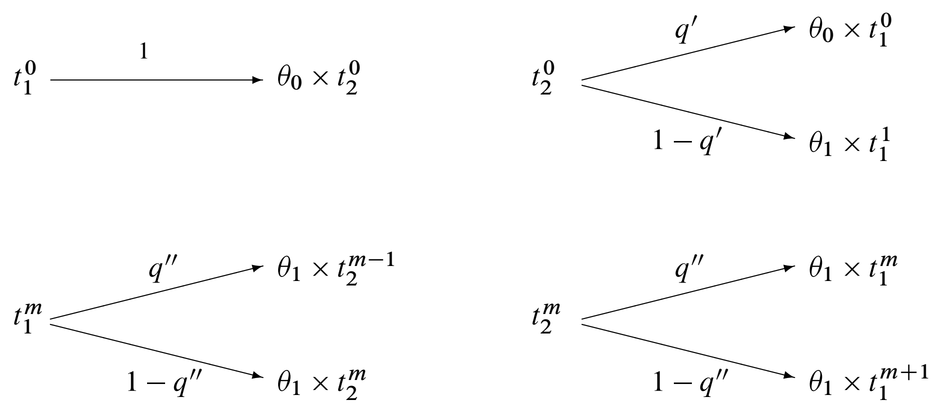

Figure 2: Type space for Example 2

As $m \to \infty$, the sequence of types t_i^m converges to $t_i = C^{(1,1)}(\theta_1)$ in the product topology, since for every $m \geqslant 1$ the $(2m-2+i)$-order beliefs of t_i^m exactly match those of t_i. That is, the first order belief of t_i^m assigns probability 1 to the event that the parameter is θ_1, the second-order belief of t_i^m assigns probability 1 to the event that the parameter is θ_1 *and* player $-i$ assigns probability 1 to θ_1, and so on, up to order $(2m-2+i)$. However, $t_i^m \notin C_i^{(p,p)}(\theta_1)$ for every $m \geqslant 0$ and $p > \max\{1-q', 1-q''\}$.[19] It follows, by Proposition 1, that t_i^m *does not* converge strategically to $t_i = C^{(1,1)}(\theta_1)$. The conclusion is that, for *some* payoff structures, the predictions under common knowledge of θ_1 are not robust to the perturbation given by the sequence t_i^m, despite the fact that the perturbation preserves arbitrarily high orders of mutual knowledge of θ_1.[20] ♦

Finally, as we will see later, strategic and uniform-strategic convergence are no longer equivalent when the unperturbed type t_i is not a complete information type. In the remainder of this section, we provide two generalizations of convergence criterion (i), which characterize strategic and uniform-strategic convergence in the general case. We begin with uniform-strategic convergence, as its characterization takes a simpler form.

3.2 Characterization of the Uniform Strategic Topology

Perturbations of the tails of belief hierarchies can have a large impact on strategic behavior. In the complete information case, the characterization of Proposition 1 quantifies the impact of such per-

[19] Type t_2^0 assigns probability $1-q' < p$ to θ_1, while t_1^1 assigns probability $q'' > 1-p$ to t_2^0. Thus, $t_2^0 \notin C_2^{(p,p)}(\theta_1)$ and, by (1), $t_1^1 \notin C_1^{(p,p)}(\theta_1)$. Proceeding by induction, suppose that $t_2^{m-1} \notin C_2^{(p,p)}(\theta_1)$ and $t_1^m \notin C_1^{(p,p)}(\theta_1)$ for some $m \geqslant 1$. Then, since t_2^m assigns probability q'' to t_1^m and t_1^{m+1} assigns probability q'' to t_2^m, again by (1) we have $t_2^m \notin C_2^{(1-q,1-q)}(\theta_1)$ and $t_1^{m+1} \notin C_1^{(1-q,1-q)}(\theta_1)$.

[20] The payoff structure for which the lack of robustness obtains is a version of the well known *coordinated attack* problem.

turbations by identifying a condition on the tails of the perturbed hierarchies—common p-belief, for p converging to 1—that is just enough to ensure the robustness of rationalizable behavior. This suggests that in order to characterize strategic (and uniform strategic) convergence in the general case, the necessary condition that k-order beliefs converge for every k (convergence in the product topology) should be strengthened by a requirement that the rate of convergence be uniform in k.

In order to define uniformity across the orders of the belief hierarchy, we first need to fix a distance on the space of k-order beliefs that metrizes the topology of weak convergence of probability measures.[21] The choice of distance matters: different distances, even if equivalent at every order k, may yield different uniformities. The characterization of uniform-strategic convergence that we provide is based on the Prohorov distance (see e.g. Billingsley, 1999, p. 72). The definition is recursive: for each integer $k \geqslant 1$, define the distance d_i^k on T_i^* as the Prohorov distance over k-order beliefs, assuming the space of $(k-1)$-order beliefs of player $-i$ is metrized by d_{-i}^{k-1}. Thus, for each player i, we set $d_i^0 \equiv 0$ and, for each integer $k \geqslant 0$ and types t_i and t_i',

$$d_i^{k+1}(t_i, t_i') = \inf\left\{\delta > 0 : \mu_i^*(E \mid t_i) \leqslant \mu_i^*\left(E^{\delta,k} \,\middle|\, t_i'\right) + \delta \;\text{ for every measurable } E \subseteq \Theta \times T_{-i}^*\right\},$$

where

$$E^{\delta,k} = \left\{(\theta, t_{-i}') \in \Theta \times T_{-i}^* \;:\; d_{-i}^k(t_{-i}, t_{-i}') < \delta \;\text{ for some } t_{-i} \text{ with } (\theta, t_{-i}) \in E\right\}.$$ [22]

That is, types t_i and t_i' are at a distance at most δ at order $k+1$ if for every event E concerning the parameter and the type of player $-i$, the probability that t_i assigns to E is at most δ above the probability that t_i' assigns to a δ-neighborhood of E at order k.

The following notion of convergence of types, introduced in Chen, Di Tillio, Faingold, and Xiong (2010), is the uniform counterpart of product convergence when the topology of weak convergence of k-order beliefs is metrized by the Prohorov distance.

Definition 3 (Uniform weak convergence)**.** A sequence of types t_i^n converges *uniform-weakly* to a type t_i if

$$d_i^{UW}(t_i, t_i^n) := \sup_{k \geqslant 1} d_i^k(t_i, t_i^n) \to 0 \quad \text{as} \quad n \to \infty.$$

Thus, under uniform-weak convergence, the tails of the belief hierarchies converge at the same rate as the lower-order beliefs, where the rate of convergence is measured relative to the Prohorov distance at each order. The examples below illustrate the definition.

Example 3 (Convex combinations with convergent weights)**.** Consider a sequence of convex combinations of a fixed pair of types in which the weight on one of the types converges to one. That

[21] Recall that a necessary condition for strategic convergence is that k-order beliefs converge weakly.

[22] Viewed as a distance on T_i^*, d_i^k is only a pseudo-distance—as opposed to a standard distance—, since there exist distinct types with the same k-order beliefs (and hence different ℓ-order beliefs, for some $\ell > k$).

is, given a pair of types $t_i, t_i' \in T_i^*$, for each $n \geqslant 1$, let t_i^n be the type whose belief is:

$$\mu_i^*(\cdot|t_i^n) = \Big(1 - \frac{1}{n}\Big)\mu_i^*(\cdot|t_i) + \frac{1}{n}\mu_i^*(\cdot|t_i').$$

Thus, for each $n \geqslant 1$, $k \geqslant 1$ and measurable $E \subseteq \Theta \times T_{-i}^*$,

$$\mu_i^*(E^{1/n,k}|t_i^n) \geqslant \mu_i^*(E|t_i^n) \geqslant \Big(1 - \frac{1}{n}\Big)\mu_i^*(E|t_i) \geqslant \mu_i^*(E|t_i) - 1/n,$$

and hence $d_i^k(t_i, t_i^n) \leqslant 1/n$. Therefore, $t_i^n \to t_i$ uniform-weakly. ♦

Example 4 (Faulty message exchange with convergent hazard rates)**.** Consider the type space $\{t_i^0, t_i^1, t_i^2 \ldots\}$ of Example 2 and, for each $n \geqslant 1$, let $\{t_i^{0,n}, t_i^{1,n}, t_i^{2,n} \ldots\}$ be the analogous type space where the probability of a message getting lost is q_n rather than q. We thus view each t_i^m as a type in the unperturbed model, and $t_i^{m,n}$ as a perturbation of type t_i^m. If the sequence q_n converges to q, then for each $m \geqslant 0$ the sequence $t_i^{m,n}$ converges to t_i^m uniform-weakly as $n \to \infty$. To see this, let $q_n' = q_n/(1+q_n)$, $q_n'' = 1/(2-q_n)$ and $\delta_n = \max\{|q_n' - q'|, |q_n'' - q''|\}$. Types t_1^0 and $t_1^{0,n}$ are both certain that the parameter is θ_0. Types t_2^0 and $t_2^{0,n}$ assign probabilities q' and q_n', respectively, to θ_0. For each player i and $m \geqslant 1$, types $t_i^{m,n}$ and t_i^m are both certain the parameter is θ_1. Thus, for each player i and $m \geqslant 0$, the first-order distance between $t_i^{m,n}$ and t_i^m is at most δ_n, that is, for every $\delta > \delta_n$, type $t_i^{m,n}$ is in the first-order δ-neighborhood $\{t_i^m\}^{\delta,1}$ of type t_i^m. Proceeding inductively, let $k \geqslant 1$ and suppose that $t_i^{m,n} \in \{t_i^m\}^{\delta,k}$ for each player i, $m \geqslant 0$ and $\delta > \delta_n$. This immediately implies that for each i, parameter θ and indices $\ell, m \geqslant 0$,

$$\mu_i^*\big(\theta \times t_{-i}^{\ell}\big|t_i^m\big) > 0 \quad \Rightarrow \quad \mu_i^*\big(\theta \times \{t_{-i}^{\ell}\}^{\delta,k}\big|t_i^{m,n}\big) \geqslant \mu_1^*\big(\theta \times t_{-i}^{\ell,n}\big|t_i^m\big) - \delta \qquad \forall \delta > \delta_n.$$

Thus, the $(k+1)$-order distance between t_i^m and $t_i^{m,n}$ is at most δ_n. Since $\delta_n \to 0$, it follows that $t_i^{m,n} \to t_i^m$ uniform-weakly for every player i and $m \geqslant 0$ ♦

In Chen, Di Tillio, Faingold, and Xiong (2010), we showed that uniform-weak convergence implies uniform-strategic convergence. Here we prove the reverse implication, thus establishing the equivalence between the two notions. The proof of the result is in appendix A.2.

Theorem 1 (Characterization of uniform strategic convergence)**.** *A sequence of types converges uniform-strategically if and only if it converges uniform-weakly.*

Ever since Rubinstein's (1989) seminal paper, misspecifications of higher-order beliefs have been recognized to have a potentially large impact on strategic predictions. The systematic treatment of Weinstein and Yildiz (2007) exposed the pervasiveness of this sensitivity by showing that the phenomenon is not peculiar to the Electronic Mail game, and hence advocated wider scrutiny of the assumptions one makes about the players' subjective beliefs. Theorem 1 quantifies the strategic impact of such assumptions (uniformly over games) by identifying an appropriate measure of proximity of hierarchies of beliefs. In effect, the role of the Prohorov distance in the definition of uniform weak convergence, and hence in the characterization result, is nontrivial. Given any distance that metrizes the topology of weak convergence of probability measures, one can always define an associated uniform distance over infinite hierarchies of beliefs. However, an example in Chen, Di Tillio, Faingold, and Xiong (2010, Section 5.2) shows that these distances may generate different topologies over infinite hierarchies, even though the induced topologies over k-order beliefs coincide for each finite k. Theorem 1 identifies one of these uniform distances that ultimately characterizes the uniform-strategic topology.

Theorem 1 serves as a benchmark for the characterization of the strategic topology in the next section, which is also based on uniform convergence in the orders of the hierarchy of beliefs. Uniform weak convergence requires convergence of beliefs concerning *all* events in the universal type space. Relaxing this requirement is the key to our characterization of the strategic topology below.

3.3 Characterization of the Strategic Topology

The definition of uniform weak convergence embodies two kinds of uniformity. The first kind, emphasized in the previous section, requires the rate of convergence of k-order beliefs to be uniform in k. The second kind, which applies to any fixed order k, is the uniformity in events that is implicit in the definition of the Prohorov distance: the k-order distance between types t_i and t_i' is less than δ if and only if

$$\sup_E \left[\mu_i^*(E|t_i) - \mu_i^*(E^{\delta,k-1}|t_i')\right] \leqslant \delta,$$

where the supremum ranges over *all* measurable subsets $E \subseteq \Theta \times T_{-i}^*$. Such uniformity in events is unnecessarily strong if the goal is to characterize the strategic topology, which does not require the rate of strategic convergence to be uniform in games. Indeed, given a fixed game, not all events concerning player $-i$ are strategically relevant for player i.[23] The main result of this section, Theorem 2, characterizes the strategic topology in terms of a notion of convergence of types that maintains the uniformity in the orders of the belief hierarchy, but only considers events from certain coarse information structures, which we now define.

Definition 4 (Frames)**.** A *frame* is a profile $\mathcal{P} = (\mathcal{P}_i)_{i \in I}$, where each $\mathcal{P}_i$ is a finite measurable partition of T_i^* such that for each $t_i, t_i' \in T_i^*$,

$$\mu_i^*(\theta \times E \,|\, t_i) = \mu_i^*(\theta \times E \,|\, t_i') \quad \forall \theta \in \Theta,\ \forall E \in \mathcal{P}_{-i} \quad \Longrightarrow \quad t_i' \in \mathcal{P}_i(t_i).$$

[23] For instance, consider a game where an action a_i^* is rationalizable for a type t_i if and only if he attaches probability greater than a threshold to player $-i$ playing a certain action a_{-i}^*. Then, in order for type t_i to verify whether he can rationalize a_i^*, he only needs to check whether his belief about the event that a_{-i}^* is rationalizable for player $-i$ lies above or below the threshold. In particular, the probability that type t_i assigns to any measurable proper subset of the latter event does not matter.

Thus, a partition profile is a frame if any two types of player i that agree on their beliefs concerning the parameter and the events in the partition of player $-i$ must belong to the same element of i's partition. Equivalently, each event in the partition of player i must be measurable with respect to the partition generated by identifying types that share the same beliefs about the parameter and the events in the partition of player $-i$. A frame is thus a coarse information model that is "self-contained:" each event in the partition of player i can be interpreted as a statement regarding the belief of player i *about the parameters and the events in the partition of player* $-i$, which can in turn be interpreted as statements about player $-i$'s beliefs *about the parameters and the events in player i's partition*, and so on.

As the notion of frames will play a central role in our characterization, it is helpful to go over a few examples to illustrate the definition.

Example 5 (Finite-order frames)**.** A *k-order frame* is a frame whose elements are k-order measurable events, that is, any two types with the same k-order beliefs must belong to the same partition element. Note that *any* profile of first-order measurable partitions is a (first-order) frame, as it can be readily verified from Definition 4. Examples of higher order frames can be constructed recursively, beginning with an arbitrary first-order frame. For instance, fix $\theta \in \Theta$ and, for each player i, fix a Borel measurable subset $B_i \subseteq [0,1]$. Then, partition T_i^* according to whether the probability that i assigns to θ lies in B_i. That is, consider the first-order frame $\mathcal{P}_i = \{F_i, T_i^* \backslash F_i\}$, where

$$F_i = \{t_i : \mu_i^*(\theta | t_i) \in B_i\}.$$

Next, fix a Borel measurable subset $B_i' \subseteq [0,1]$ for each player i and consider the partition $\mathcal{P}_i' = \{F_i', T_i^* \setminus F_i'\}$, where F_i' is the event that player i assigns a probability in B_i' to the event that the parameter is θ and player $-i$ assigns to θ a probability in B_{-i}. That is,

$$F_i' = \{t_i : \mu_i^*(\theta \times F_{-i} | t_i) \in B_i'\}.$$

Then, the join of the partitions $\mathcal{P}_i$ and $\mathcal{P}_i'$, which we denote by $\mathcal{P}_i''$, partitions T_i^* into four second-order measurable events, according to whether a type belongs to each of the events F_i and F_i'.[24] The profile $\mathcal{P}'' = (\mathcal{P}_i'')_{i \in I}$ is an example of a second-order frame. Examples of higher order frames can be constructed in a similar fashion. ⧫

Example 6 (Common belief frames)**.** Partitioning types according to whether or not a given subset of parameters is common belief gives rise to a frame. Formally, for each $E \subseteq \Theta$ and $\mathbf{p} \in [0,1]^2$, the profile of bi-partitions $\mathcal{P}_i = \{C_i^{\mathbf{p}}(E), T_i^* \setminus C_i^{\mathbf{p}}(E)\}$, $i \in I$, is a frame. Indeed, if any two types of player i agree on the probabilities assigned to the elements of $\Theta \times \mathcal{P}_{-i}$, then they must agree on the probability of the event $E \times C_{-i}^{\mathbf{p}}(E)$. Thus, by property (1) of common beliefs, either both types belong to $C_i^{\mathbf{p}}(E)$, or neither does. More generally, given a frame $\mathcal{P} = (\mathcal{P}_i)_{i \in I}$, a

[24] Recall that the join of a pair of partitions $\mathcal{P}_i$ and $\mathcal{P}_i'$, written $\mathcal{P}_i \vee \mathcal{P}_i'$ is the coarsest partition that is finer than both $\mathcal{P}_i$ and $\mathcal{P}_i'$.

collection of events $E_1, \dots, E_m \in 2^\Theta \otimes \mathcal{P}_1 \otimes \mathcal{P}_2$ and a collection of pairs $\mathbf{p}_1, \dots, \mathbf{p}_m \in [0,1]^2$, the profile of joined partitions

$$\mathcal{P}_i \vee \{C_i^{\mathbf{p}_1}(E_1), T_i^* \setminus C_i^{\mathbf{p}_1}(E_1)\} \vee \cdots \vee \{C_i^{\mathbf{p}_m}(E_m), T_i^* \setminus C_i^{\mathbf{p}_m}(E_m)\}, \qquad i \in I,$$

is a frame, called a *common belief frame*. ♦

Example 7 (Strategic frames)**.** Given a game $G = (A_i, g_i)_{i \in I}$, the *strategic frame* associated with G is the profile of partitions $\mathcal{P} = (\mathcal{P}_i)_{i \in I}$ where the elements of each $\mathcal{P}_i$ are the equivalence classes of types with the same rationalizable actions in G, that is,

$$\mathcal{P}_i = \Big\{\{t_i : R_i(t_i, G) = A_i'\} : \varnothing \neq A_i' \subseteq A_i\Big\}.$$

The proof that $\mathcal{P}$ is indeed a frame, given in Appendix A.3, is based on the observation that an action is rationalizable for a type of player i only if it can be rationalized (at each order) by a conjecture $\nu \in \Delta(\Theta \times T_{-i}^* \times A_{-i})$ that assumes that all types of player $-i$ that share the same rationalizable actions behave in the same way. That is, for each $A_{-i}' \subseteq A_{-i}$ and $a_{-i} \in A_{-i}'$, the ratio $\nu(\theta \times E_{-i} \times a_{-i})/\nu(\theta \times E_{-i} \times A_{-i}')$ is the same for each parameter θ and each event $E_{-i} \subseteq \{t_{-i} : R_{-i}(t_{-i}, G) = A_{-i}'\}$. This observation implies that, in order for a type of player i to rationalize an action in G, what matters is the probability assigned to events of the form $\theta \times \{t_{-i} : R_{-i}(t_{-i}, G) = A_{-i}'\}$, not how this probability is distributed within those events. As a result, all types of player i who agree on the probabilities assigned to the elements of $\Theta \times \mathcal{P}_{-i}$ must have the same rationalizable actions and thus belong to the same element of $\mathcal{P}_i$. ♦

We now define the notion of convergence that is used in the characterization of the strategic topology below. Given a frame $\mathcal{P} = (\mathcal{P}_i)_{i \in I}$, for each player i set $d_{i,\mathcal{P}}^0 \equiv 0$ and, recursively, for each integer $k \geqslant 0$ and types t_i and t_i', define

$$d_{i,\mathcal{P}}^{k+1}(t_i, t_i') = \inf\Big\{\delta > 0 \;:\; \mu_i^*(E \,|\, t_i) \leqslant \mu_i^*\big(E_{\mathcal{P}}^{\delta,k} \,\big|\, t_i'\big) + \delta \quad \forall E \in 2^\Theta \otimes \mathcal{P}_{-i}\Big\},$$ [25]

where $E_{\mathcal{P}}^{\delta,k}$ denotes the δ-neighborhood of event E at order k relative to $\mathcal{P}$, that is,

$$E_{\mathcal{P}}^{\delta,k} = \Big\{(\theta, t_{-i}') \in \Theta \times T_{-i}^* \;:\; d_{-i,\mathcal{P}}^k(t_{-i}, t_{-i}') < \delta \;\text{ for some } t_{-i} \text{ with } (\theta, t_{-i}) \in E\Big\}.$$

Thus, the definition of $d_{i,\mathcal{P}}^k$ is similar to that of the k-order Prohorov distance d_i^k, but restricts the events for which the proximity is measured to those in the frame $\mathcal{P}$.[26]

Definition 5 (Uniform weak convergence relative to a frame)**.** Let $\mathcal{P} = (\mathcal{P}_i)_{i \in I}$ be a frame. A sequence of types t_i^n converges to a type t_i *uniform-weakly on* $\mathcal{P}$ if

$$d_{i,\mathcal{P}}^{UW}(t_i, t_i^n) := \sup_{k \geqslant 1} d_{i,\mathcal{P}}^k(t_i, t_i^n) \to 0 \quad \text{as} \quad n \to \infty.$$

[25] Here, $2^\Theta \otimes \mathcal{P}_{-i}$ denotes the algebra of events in $\Theta \times T_{-i}^*$ generated by the sets of the form $\theta \times F_{-i}$, with $F_{-i} \in \mathcal{P}_{-i}$.

[26] This restriction makes $d_{i,\mathcal{P}}^k$ only a pre-distance, that is, it satisfies $d_{i,\mathcal{P}}^k(t_i, t_i') \geqslant 0$ and $d_{i,\mathcal{P}}^k(t_i, t_i) = 0$, but it may fail symmetry and the triangle inequality.

With this definition in place, we are ready to state our main result, whose proof is presented in Appendix A.4.

Theorem 2 (Characterization of strategic convergence)**.** *A sequence of types converges strategically if and only if it converges uniform-weakly on every frame.*

The theorem has both conceptual and practical implications. First, it deepens our understanding of the foundations of strategic robustness (or lack thereof) by qualifying the impact that the tails of the belief hierarchies have on the strategic behaviors of a type. As in Theorem 1, the result draws attention to a *particular form* of uniform convergence across the levels of the belief hierarchy as a condition for robustness. While it is expected that some form of uniform convergence should play a role, it is much less clear at the outset what kind of uniformity would ultimately lead to a characterization. The theorem highlights the role of frames as the coarse information structures that are relevant for strategic convergence.

Second, the characterization provides a *practical criterion* for robustness, similar to the role of common p-beliefs in the complete information case. From the applied perspective, the *if* direction of the theorem is more useful. Indeed, it is often the case that the analyst is interested in a particular game, or a small parametrized family of games. If the perturbations that the analyst contemplates pass the robustness criterion provided by the theorem, then the analyst is reassured that the predictions of the unperturbed model are robust in the game(s) of interest. However, if the perturbation fails the criterion, then the theorem only implies that there is *some* payoff structure for which the predictions of the unperturbed model fail to be robust. Failure of the criterion thus "raises a flag" to the analyst, but is ultimately inconclusive: further examination is required to determine whether the predictions of interest are robust. Nonetheless, the *only if* direction of the result is still useful, as it ensures the criterion is *sharp*: among all sufficient conditions for robustness that do not make use of the payoff structure, the criterion of Theorem 2 is one that minimizes the instances when the "flag is raised" unnecessarily.

The applied perspective put forward above leads naturally to a question: How easy is it to verify that a sequence of perturbations passes the robustness criterion of Theorem 2? This is ultimately a complexity question and it is hard to give a precise answer. Instead, we provide an example below to illustrate a relatively simple application of the criterion in a nontrivial case where the stronger (but easier-to-check) condition of uniform weak convergence fails. The economics of the example is interesting in its own right: the unperturbed type space is the one generated by the automated message exchange protocol of Rubinstein (1989), while the perturbed type space is generated by a variation of the same protocol in which the players, whom we interpret as boundedly rational, become "overwhelmed by evidence" once the number of messages exchanged reaches an arbitrarily large (but finite) threshold, in which case they believe the parameter is common knowledge.

Example 8 (Faulty message exchange with "boundedly rational" players)**.** Suppose the unperturbed model is the message exchange type space $\{t_i^0, t_i^1, t_i^2, \ldots\}$ of Example 2. Recall that in that type space there is no type that has common knowledge of θ_1, regardless of the number of messages exchanged. The perturbed model is a variation of this type space in which the players mistakenly believe there is common knowledge of θ_1 as soon as they have received n messages, where $n \geqslant 1$ is interpreted as the level of sophistication of the players.[27] Formally, for each $n \geqslant 1$ we consider a finite type space $\{s_i^{0,n}, s_i^{1,n}, \ldots, s_i^{n,n}\}$ where $s_i^{n,n}$ has common knowledge of θ_1, and the beliefs of the remaining types mimic those of the type space in Example 2, as illustrated in Figure 3. For each player i and $m \geqslant 0$, does the sequence $s_i^{m,n}$ converge strategically to t_i^m as $n \to \infty$?

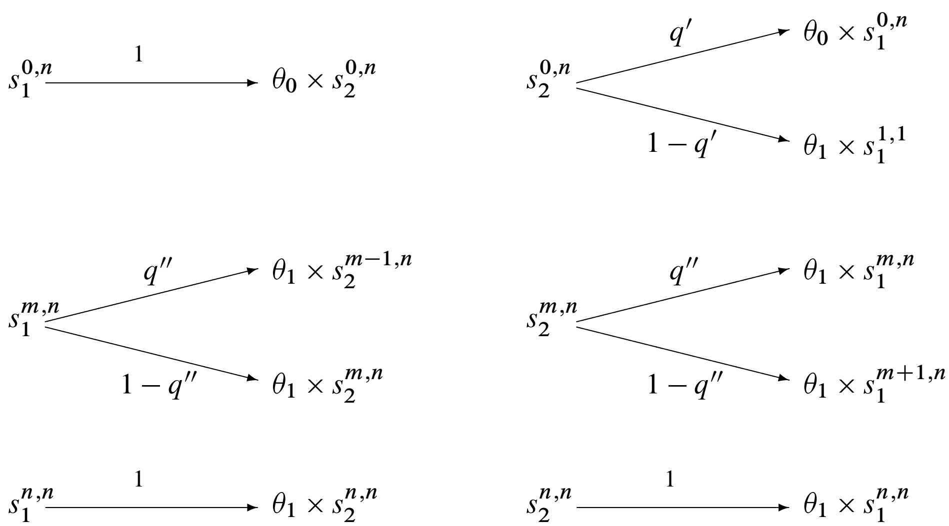

Figure 3: Type space for Example 8

The sequence $s_i^{m,n}$ fails the uniform weak converge criterion. Indeed, it can be readily verified that the $(2n-1+i)$-order Prohorov distance between $s_i^{n,n}$ and t_i^n must exceed $\min\{q', q''\}$ and hence, by an induction argument, for each $0 \leqslant m \leqslant n-1$ the $(4n-2m-i+1)$-order Prohorov distance between $s_i^{m,n}$ and t_i^m is also at least $\min\{q', q''\}$.

Nonetheless, we show that the sequence converges strategically by verifying the criterion of Theorem 2: the sequence $s_i^{m,n}$ converges to t_i^m uniform-weakly on every frame. In effect, for every frame $\mathcal{P} = (\mathcal{P}_i)_{i \in I}$ we show that $d_{i,\mathcal{P}}^{UW}(s_i^{m,n}, t_i^m) = 0$ for every n sufficiently large. To see why, first note that, since each $\mathcal{P}_i$ is finite, there exists N large enough that, for each $n \geqslant N-1$ and player i, the intersection $\mathcal{P}_i(t_i^n) \cap \{t_i^0, t_i^1, t_i^2, \ldots\}$ is infinite.[28] This implies that for every

[27] Here the words "mistakenly" and "sophistication" should be interpreted with care. To present the perturbed model we have used a bounded rationality story: unlike in the unperturbed model, the players are not able to apply Bayes' rule exactly; once the number of messages exchanged exceeds a finite threshold, they believe there is common knowledge of θ_1, rather than just mutual knowledge up to a finite order. However, bounded rationality is only a story here, as the perturbed model is still a standard type space.

[28] Following standard notation, $\mathcal{P}_i(t_i)$ designates the element of $\mathcal{P}_i$ containing t_i.

$i \in I$, $n \geqslant N$, $k \geqslant 1$ and $\delta > 0$, type $s_{-i}^{n,n} = C_{-i}^{(1,1)}(\theta_1)$ belongs to the k-order Prohorov δ-neighborhood of each of the frame elements

$$\mathcal{P}_{-i}(t_{-i}^{n-1}), \quad \mathcal{P}_{-i}(t_{-i}^{n}) \quad \text{and} \quad \mathcal{P}_{-i}(t_{-i}^{n+1}),$$

as each of these sets, having an infinite intersection with $\{t_{-i}^0, t_{-i}^1, t_{-i}^2, \ldots\}$, must contain some type t_{-i}^m that has the same k-order beliefs as $s_{-i}^{n,n}$.[29] As the support of $\mu_i^*(\cdot|t_i^n)$ is contained in $\theta_1 \times \{t_{-i}^{n-1}, t_{-i}^n, t_{-i}^{n+1}\}$, it follows that

$$\mu_i^*(E_{\mathcal{P}}^{\delta,k}|s_i^{n,n}) \geqslant \mu_i^*(E^{\delta,k}|s_i^{n,n}) \geqslant \mu_i^*(E|t_i^n) - \delta \quad \forall E \in 2^\Theta \otimes \mathcal{P}_1 \otimes \mathcal{P}_2,$$

and since $\delta > 0$ is arbitrary, we have $d_{i,\mathcal{P}}^k(t_i^n, s_i^{n,n}) = 0$ for every $k \geqslant 1$ and $n \geqslant N$. Next, since the beliefs of $s_i^{0,n}, \ldots, s_i^{n-1,n}$ mimic those of $t_i^0, \ldots, t_i^{n-1}$ (compare Figures 2 and 3), for every $0 \leqslant m \leqslant n-1$ we have $d_{i,\mathcal{P}}^1(t_i^m, s_i^{m,n}) = 0$, and for every $k \geqslant 1$,

$$d_{i,\mathcal{P}}^{k+2}(t_i^0, s_i^{0,n}) \leqslant \max\left\{d_{i,\mathcal{P}}^k(t_i^1, s_i^{1,n}), d_{i,\mathcal{P}}^k(t_i^0, s_i^{0,n})\right\},$$

and for every $1 \leqslant m \leqslant n-1$,

$$d_{i,\mathcal{P}}^{k+2}(t_i^m, s_i^{m,n}) \leqslant \max\left\{d_{i,\mathcal{P}}^k(t_i^{m+1}, s_i^{m+1,n}), d_{i,\mathcal{P}}^k(t_i^m, s_i^{m,n}), d_{i,\mathcal{P}}^k(t_i^{m-1}, s_i^{m-1,n})\right\}.$$

These conditions, combined with the fact that $d_{i,\mathcal{P}}^k(t_i^n, s_i^{n,n}) = 0$ for every $k \geqslant 1$ and $n \geqslant N$ (demonstrated above), imply that $d_{i,\mathcal{P}}^k(t_i^m, s_i^{m,n}) = 0$ for every $k \geqslant 1$, $n \geqslant N$ and $0 \leqslant m \leqslant n$. Therefore, $d_{i,\mathcal{P}}^{UW}(t_i^m, s_i^{m,n}) = 0$ for every $n \geqslant N$ and $0 \leqslant m \leqslant n$. Since the frame $\mathcal{P}$ is arbitrary, we conclude that $t_i^{m,n}$ converges to t_i^m uniform-weakly on every frame. ♦

It is worthwhile remarking that the proof of the theorem, in Appendix A.4, actually demonstrates a stronger, context-dependent version of the *if* direction:

Given a game G and $\varepsilon > 0$, any action that is rationalizable for a type t_i in G must remain rationalizable for any type sufficiently far in the tail of any given sequence converging to t_i uniform-weakly on the strategic frame of G.

To the extent that uniform weak convergence on a given frame (strategic or otherwise) is a less demanding condition than uniform weak convergence on all frames, such a context-dependent criterion for robustness is potentially useful. Whether it will actually be useful in practice will depend on the possibility of exploiting properties of the particular payoff functions of interest (beyond finiteness of action sets) to obtain a tractable sufficient condition for uniform weak convergence on the corresponding strategic frame (weaker than uniform weak convergence on all frames). We return to this issue in section 5, where we discuss the possibility of characterizions of strategic convergence for *classes* of games (as opposed to all games).

[29] Recall from Example 2 that $s_i^{n,n}$ and t_i^m are at zero $(2m-2+i)$-order Prohorov distance.

However, the context-dependent version of the *only if* direction fails to hold: convergence of rationalizable behavior in a given game G does *not* imply uniform weak convergence on the strategic frame of G. To see why, pick any game G in which the payoff of each player depends on his own action and on the parameter only, so that the strategic frame of G is a nontrivial first-order frame. Then, any sequence of types whose first order beliefs converge must also converge strategically in game G. But this clearly does not imply that the sequence will converge uniform weakly on the strategic frame of G, as it does not even imply convergence at second order.

A related comment is that uniform weak convergence on all strategic frames is equivalent to strategic convergence. This follows directly from the stronger version of the *if* direction discussed above, and the *only if* direction of Theorem 2. However, we find such a characterization less interesting than the one stated in the theorem, as we do not have a characterization of strategic frames that does not make explicit reference to the rationalizable correspondence, hence we regard strategic frames as "endogenous" objects. By contrast, general frames are defined in terms of a purely epistemic condition. In fact, *not every frame is a strategic frame*. In a strategic frame, each player's partition must contain an element that is open in the product topology, namely, any element consisting of types whose set of rationalizable actions is minimal (w.r.t. set inclusion).[30] General frames, however, need not have elements that are open sets. Letting B_i be the set of rational numbers in the first-order frame of Example 5 illustrates this fact.

Finally, Theorem 2 enables a connection with the early literature on robustness and, in particular, with the notion of common p-beliefs. This connection is the subject of the next section.

3.4 Connection to Common Beliefs

Previous work on robustness of Bayesian equilibrium in common prior settings (Monderer and Samet, 1989, 1996; Kajii and Morris, 1998) has proved equivalences between ex ante notions of strategic convergence and convergence of common beliefs. In this section, we reexamine this connection in our interim, non-common prior framework.

Definition 6 (Common belief convergence)**.** A sequence of types t_i^n *converges in common beliefs* to a type t_i if for every $\delta > 0$ and $k \geqslant 1$ there exists $N \geqslant 1$ such that for every $n \geqslant N$, $\mathbf{p} \in [0,1]^2$ and measurable $E \subseteq \Omega$,

$$t_i \in C_i^{\mathbf{p}}(E) \quad \Longrightarrow \quad t_i^n \in C_i^{\mathbf{p}-(\delta,\delta)}\big(E^{\delta,k}\big).$$ [31]

Thus convergence in common beliefs requires each event that is common $\mathbf{p}$-belief for the limit

[30] That such an element is product-open follows from minimality and the upper hemi-coninuity of the rationalizable correspondence in the product topology.

[31] If we modify the definition by restricting the class of events E to be the class of sets that are closed in the product topology, then the notion of convergence remains the same. Indeed, letting $\bar{E}$ denote the product-topology closure of E, we have $C_i^{\mathbf{p}}(E) \subseteq C_i^{\mathbf{p}}(\bar{E})$ and $E^{\delta,k} = (\bar{E})^{\delta,k}$.

type to have, at each order k, a δ-neighborhood that is common-($\mathbf{p} - (\delta, \delta)$)-belief for all types that lie sufficiently far in the tail of the sequence, where "sufficiently far" depends on δ and on the order k, but is independent of the event E and the degree of common belief $\mathbf{p}$.

To shed light on the definition, it is useful to begin with a formal analogy with product convergence. Recall that a sequence of types t_i^n converges to a type t_i in the product topology if and only if for every $\delta > 0$ and $k \geqslant 1$ there exists $N \geqslant 1$ such that for every $n \geqslant N$, $p \in [0,1]$ and measurable set $E \subseteq \Omega$,[32]

$$t_i \in B_i^p(E) \implies t_i^n \in B_i^{p-\delta}(E^{\delta,k}).$$

Thus, the definition of convergence in common beliefs formally replaces the belief operator B_i^p in the characterization of product convergence above, by the *common* belief operator $C_i^{\mathbf{p}}$. This strengthens product convergence by imposing conditions on the tails of the belief hierarchy, which are irrelevant for product convergence. Indeed, convergence in common beliefs implies product convergence, because setting $\mathbf{p} = (p, 0)$ yields $B_i^p = C_i^{\mathbf{p}}$. As for the connection with strategic convergence, Theorem 2 yields the following corollary, proved in appendix A.5. The result follows from the fact that uniform weak convergence in every common belief frame (Example 6) implies convergence in common beliefs, which we establish in appendix A.5 through a sequence of lemmas.

Corollary 1 (Common belief convergence)**.** *Strategic convergence implies convergence in common beliefs.*

However, *the converse does not hold*, as we show with a striking example presented in appendix A.6. The corollary and the counterexample shed further light on the robustness problem by providing a necessary condition for strategic convergence that, albeit not sufficient, is easier to interpret than the more abstract notion of uniform weak convergence on frames. Furthermore, they clarify the connection between the strategic topology for rationalizability on the universal type space and the ex ante strategic topologies for Bayesian equilibrium in common prior type spaces of the early literature. In particular, the example shows that the equivalence between strategic robustness and common p-beliefs breaks down when we adopt the interim approach.

Finally, our analysis of convergence in common beliefs enables a connection with the notion of *critical* and *regular* types, due to Ely and Pęski (2011). Critical types are those types to which product convergence fails to imply strategic convergence, while regular types are the non-critical ones.[33] Indeed, Ely and Pęski (2011, Theorem 1) show that a type $t_i \in T_i^*$ is critical if and only if $t_i \in C_i^{p\mathbf{1}}(E)$ for some $p > 0$ and some product-closed, proper subset $E \subset \Omega$. That paper

[32] This follows directly from the following two facts: (i) the Mertens-Zamir isomorphism $\mu_i : T_i^* \to \Delta(\Theta \times T_{-i}^*)$ becomes a homeomorphism when each T_j^* is endowed with the product topology and $\Delta(\Theta \times T_{-i}^*)$ is endowed with the topology of weak convergence (as remarked in footnote 15); (ii) the Prohorov metric on $\Delta(\Theta \times T_{-i}^*)$ metrizes the topology of weak convergence.

[33] The analogous definition for uniform-strategic convergence is: t_i is *uniformly critical* if there is some sequence

highlights a tension concerning critical types. On one hand, they are pervasive in applications—for instance, types in finite type spaces and common prior type spaces (both of which we discuss in the next section) are examples of critical types. On the other hand, they are topologically small: Ely and Pęski (2011, Theorem 2) show that the set of regular types is a residual subset of the universal type space under the product topology. Our next result, proved in appendix A.7, shows that this tension disappears when one considers strategic genericity.

Corollary 2 (Genericity of Critical Types)**.** *The set of critical types is open and dense in the universal type space under the strategic topology.*

At first sight, it might seem appropriate for an analyst with limited knowledge of the players' higher-order beliefs to hedge against that uncertainty by adopting regular types as his model, for regular types exhibit continuous rationalizable behavior (by definition). Strikingly, Corollary 2 shows that this is not a robust modeling choice. Suppose that the analyst only knows the true type's k-order belief. No matter how large k is, modeling the unknown true type by means of a regular type with that same k-order belief is always a feasible choice for the analyst, as Theorem 2 in Ely and Pęski (2011) implies that the set of regular types is dense in the product topology. But, since critical types are also dense in the product topology, the analyst cannot rule out the possibility that the true type is critical. Moreover, since the set of critical types is open in the strategic topology, if the true type is indeed critical and the desired level of prediction accuracy small, the assumed regular type must deliver inaccurate predictions in some game. To put it differently, while regularity does shield the analyst against misspecified beliefs of sufficiently high order, it may fail to deliver robust predictions, because "sufficiently high" is necessarily higher than k when the true type is critical, a possibility that the analyst cannot exclude.

4 Implications for Finite Models and Common Priors

In this section we examine some implications of our results for two commonly used classes of models, namely finite type spaces and common prior type spaces.

4.1 Finite Types

It is common, in both theoretical and applied work, to model incomplete information scenarios by positing a finite type space. In addition to advantages in terms of tractability, this simplifying assumption has received some justification since Dekel, Fudenberg, and Morris (2006) show that *finite types*, those belonging to a finite type space, are dense in the universal type space under the

that converges to t_i in the product topology but fails to converge uniform-strategically. An immediate implication of our Theorem 1 is that *all* types in the universal type space are uniformly critical, since for every type t_i there is always a sequence that converges to t_i in the product topology but does not converge uniform-weakly.

strategic topology. That is, the strategic behaviors of any type can be approximated by the strategic behaviors of finite types.[34]

As we have seen above, strategic convergence is a *strictly* stronger condition than convergence in common beliefs. However, when the limit type is a finite type, we show that common belief convergence implies uniform weak convergence. We thus have the following theorem, proved in appendix A.8.

Theorem 3. *Given a finite type t_i and a sequence of (possibly infinite) types t_i^n, the following statements are equivalent:*

(a) $t_i^n \to t_i$ uniform-weakly;

(b) $t_i^n \to t_i$ uniform-strategically;

(c) $t_i^n \to t_i$ strategically;

(d) $t_i^n \to t_i$ uniform-weakly on every frame;

(e) $t_i^n \to t_i$ in common beliefs.

Thus, when the unperturbed model is finite, any distinction between the uniform strategic and the strategic topology disappears. Furthermore, we obtain equivalence with common belief convergence, closing the gap with the literature on ex ante robustness in the case of perturbations of finite types. In other words, by assuming finite types, the analyst can guarantee robustness of his predictions to perturbations, provided the perturbations that he contemplates as possible preserve approximate common beliefs in the sense of Definition 6.

4.2 Common Prior Types

The common prior assumption, according to which the beliefs of all players are derived from the same prior by conditioning on their private information, is a cornerstone of virtually all models of information economics. However, its interpretation and methodological justification are controversial—see Morris (1995), Gul (1998) and Aumann (1998)—, and its positive implications often puzzling, as witnessed by the celebrated agreement theorem (Aumann, 1976) and the no-trade theorem that followed it (Milgrom and Stokey, 1982). Taking a more fundamental approach to the issue, Lipman (2003) investigates the extent to which models satisfying the assumption approximate models where it is violated, focusing on finite-order beliefs approximations. In this section we adopt a similar perspective, and apply our characterizations to revisit the problem, considering also *strategic* approximations.

[34] Chen, Di Tillio, Faingold, and Xiong (2010, Theorem 3) show, however, that they are nowhere dense in the uniform strategic topology.

We recall the following definition from Mertens and Zamir (1985).

Definition 7. A *common prior* is a probability measure $\pi \in \Delta(\Theta \times T_1^* \times T_2^*)$ such that, for each player i and measurable sets $E \subseteq T_i^*$ and $F \subseteq \Theta \times T_{-i}^*$,

$$\pi(E \times F) = \int_E \mu_i^*(F|t_i)\pi(dt_i).$$

Thus, π is a common prior if, for each player i, the function $t_i \mapsto \mu_i^*(F|t_i)$ is (a version of) the conditional probability under π of the event F, given the type of player i. When π assigns positive probability to a particular type t_i, Bayes rule applies given the event $E = \{t_i\}$, and the equation above can be stated in terms of the usual formula for conditional probability, that is,

$$\mu_i^*(F|t_i) = \pi(F|t_i) := \frac{\pi(t_i \times F)}{\pi(t_i)}.$$

As noted in Lipman (2003), there are two standard, related approaches to common priors in the literature. The first follows an *ex ante* perspective, as it formalizes the common prior assumption as a property of a *type space*. Let $\mathcal{T}$ denote the family of all type spaces $(T_i)_{i\in I}$ such that each set T_i is closed in the product topology on T_i^*.

Definition 8. A type space $(T_i)_{i\in I}$ is a *common prior type space* if there exists a common prior π such that, for each player i, the set T_i is the support of the marginal of π on T_i^*.[35] Every type space in $\mathcal{T}$ that is not a common prior type space is a *non-common prior type space*.

The support condition formalizes in a meaningful way the requirement that the beliefs of the types in T_i are consistent with a common prior. Indeed, as shown in our online appendix Chen, Di Tillio, Faingold, and Xiong (2016), there exist common priors whose marginals on T_i^* have the whole set T_i^* as their supports, hence without the support condition *every* type space is a common prior type space.

The second approach, adopted by Lipman (2003), follows the analogous *interim* perspective and applies to *types*.

Definition 9. A type $t_i \in T_i^*$ is a *common prior type* if there is a common prior π with $\pi(t_i) > 0$. Every type in T_i^* that is not a common prior type is a *non-common prior type*.

The positivity condition is needed for the definition of a common prior type to have a minimal bite—without the condition, *every* type is a common prior type (see again our online appendix Chen, Di Tillio, Faingold, and Xiong, 2016). Observe that a finite type space $T = (T_i)_{i\in I}$ is a

[35]This means that T_i is the smallest subset of T_i^* that is closed in the product topology on T_i^* and has probability 1 under π. The faulty message exchange type space of Example 2 becomes a common prior type space, once for each player i we add the type $C_i^{(1,1)}(\theta_1)$ to the set of types of player i.

common prior type space if and only if, for each player i, each type $t_i \in T_i$ is a common prior type. That is, in the case of a finite type space, the ex ante and the interim approach deliver the same formalization of the common prior assumption.

Lipman (2003) shows that the set of finite common prior types is dense in the product topology, but explicitly warns that this result should not be interpreted as providing a foundation for the common prior assumption: Although every type can be approximated by a (finite) common prior type in the product topology, the strategic behavior of that type can be very different from the strategic behavior of any approximating common prior type. Our next theorem shows that this lack of robustness is, in effect, a generic phenomenon in the universal type space. In particular, the denseness result of Lipman (2003) is reversed once we consider the strategic rather than the product topology. The proof of the result is in appendix A.9.

Theorem 4 (Genericity of non-common prior types)**.** *The set of non-common prior types contains a set of types that is open and dense in T_i^* under the strategic topology.*

The fact that the set of non-common prior types contains a (strategically) open set means that no non-common prior type in that set can be approximated, in terms of strategic behaviors, by common prior types. Thus, if we cannot exclude that the true type lies in that set, then by assuming a common prior type we cannot guarantee in advance (i.e. whatever the game is) that our predictions will be close to those regarding the true type. Now, by denseness, *every* type is arbitrarily close (in the strategic topology) to a non-common prior type. Thus, unless we have sufficiently detailed knowledge of the true type's strategic behavior in each game, then indeed we cannot exclude that the true type is a non-common prior type. Strategic genericity of non-common prior types can then be stated in non-technical terms as follows: By assuming a common-prior type, a modeler has—from his own, imperfect view of the player's higher-order beliefs—no advance guarantee that his predictions will accurately describe the behavior of the true type that he is trying to approximate.

A caveat of Theorem 4 is that many economic models assume an uncountable common prior type space with a nonatomic prior, so that no type is a common prior type according to our definition. Since Theorem 4 does not apply to such cases, it is natural to wonder whether genericity of non-common priors critically depends on the interim perspective. To address this issue, we now turn to the ex ante perspective and consider topologies on *type spaces* rather than types. Specifically, let $T^* = \times_{i \in I} T_i^*$ and identify each type space $(T_i)_{i \in I}$ with the rectangle $\times_{i \in I} T_i \subseteq T^*$. Then $\mathcal{T}$ can be viewed as a family of (product-)closed subsets of T^*. Let d_i^p and d_i^s be any distances on T_i^* that metrize the product and the strategic topology, respectively.[36] For each

[36]Dekel, Fudenberg, and Morris (2006) show that the strategic topology on T_i^* is metrizable by the distance d_i^S defined as follows. For each game $G = (A_j, g_j)_{j \in I}$, $a_i \in A_i$ and $t_i \in T_i^*$, let $e_i(t_i, a_i, G) = \inf\{\varepsilon \geqslant 0 : a_i \in R_i(t_i, G, \varepsilon)\}$. For each $m \geqslant 1$, write $\mathcal{G}^m$ for the set of games where, for every $j \in I$, the set A_j has at most m elements. Then, for each $t_i, t_i' \in T_i^*$, $d_i^S(t_i, t_i') = \sum_{m \geqslant 1} 2^{-m} \sup_{G=(A_i,g_i)_{i \in I} \in \mathcal{G}^m} \max_{a_i \in A_i} \left| e_i(t_i, a_i, G) - e_i(t_i', a_i, G) \right|$.

$t = (t_i)_{i \in I}$ and $t' = (t_i)_{i \in I}$ in T^*, let

$$d^P(t, t') = \max_{i \in I} d_i^P(t_i, t_i') \qquad \text{and} \qquad d^S(t, t') = \max_{i \in I} d_i^S(t_i, t_i').$$

The topologies on $\mathcal{T}$ that we consider are the Hausdorff topologies on $\mathcal{T}$ corresponding to the distances d^P and d^S on T^*. For each $T, T' \in \mathcal{T}$ and $d \in \{d^P, d^S\}$, let

$$h_d(T, T') = \max\left\{ \sup_{t \in T} \inf_{t' \in T'} d(t, t'), \sup_{t' \in T'} \inf_{t \in T} d(t, t') \right\}.$$

Definition 10. The *product topology on type spaces* is the topology on $\mathcal{T}$ induced by the metric h_{d^P}. The *strategic topology on type spaces* is the topology on $\mathcal{T}$ induced by the metric h_{d^S}.

In other words, a sequence of type spaces T^n converges to a type space T in the product topology on type spaces if for every $\delta > 0$, $k \geqslant 1$ and sufficiently large n, the k-order belief of each type in T is δ-close to the k-order belief of some type in T^n, and vice versa, the k-order belief of each type in T^n is δ-close to the k-order belief of some type in T. Similarly, T^n converges to T in the strategic topology on type spaces if for every $\varepsilon > 0$, game G and sufficiently large n, the rationalizable actions of each type in T are ε-rationalizable for some type in T^n, and vice versa, the rationalizable actions of each type in T^n are ε-rationalizable for some type in T.

We are now ready to present our last theorem, which we prove in appendix A.9.

Theorem 5 (Genericity of non-common prior type spaces)**.** *The set of non-common prior type spaces contains a set of type spaces that is open and dense in $\mathcal{T}$ under both the product topology and the strategic topology on type spaces.*

The theorem highlights a fundamental difference between the ex ante and the interim approach to common priors. If we change the perspective and consider type spaces as opposed to types, we obtain again a reversal of Lipman's (2003) denseness result, even if we still take product convergence as the underlying notion of proximity of types. Moreover, the strategic genericity of non-common prior types extends to non-common prior type spaces. Thus, common prior models only serve to approximate a nongeneric set of incomplete information scenarios, whether we consider the approximation from the interim perspective (Theorem 4) or the ex ante perspective (Theorem 5). As a consequence of these results, we conclude that there is no general, principled approach to justify common priors based on strategic approximations.

5 Discussion: Interpretation of Results, and Extensions

5.1 Universal Quantifier over Payoffs

The definition of the strategic topology, on which our robustness exercise is based, has a universal quantifier over payoff structures. This has important implications for both the interpretation and the applicability of our results. As discussed in section 3.3, Theorem 2 provides a practical criterion for robustness that can be useful in situations where solving for the rationalizability correspondence globally in the universal type space is impractical. Although the condition is not necessary for robustness relative to a fixed payoff structure, the theorem shows that the condition is sharp: there is no way to improve it unless one looks for joint conditions on hierarchies of beliefs and payoffs. Of course, such a context-free approach comes with a caveat: when the analyst verifies that the condition fails, he needs to undertake a more in-depth examination of the robustness problem in the particular strategic context of interest.

In light of this caveat, it is natural to wonder whether it is possible to obtain characterizations that make more use of the payoff information, while maintaining a reasonable degree of tractability. First, consider the extreme opposite of our approach, namely to look for a characterization of robustness that has *complete* freedom to depend on the payoff functions of the players. As discussed in the paragraph following example 8, it may be possible to obtain *sufficient* conditions for convergence of rationalizable behavior that depend on the strategic context. As for context-dependent conditions that are both necessary and sufficient, we do not think they can be insightful or useful, as they must boil down to a reformulation of the definition of strategic convergence in the particular game. This issue is not peculiar to multi-agent settings. Already in a decision theoretic framework à la Savage (one-person game against nature), how does one characterize the robustness of a decision maker's choices to sequences of perturbations of his subjective beliefs? A sharp sufficient condition (when the state space is compact and utilities are continuous) is that the sequence of perturbations converges to the unperturbed belief of the decision maker in the topology of weak convergence of probability measures. But if the goal is to find a necessary and sufficient condition among all those that are allowed to depend on the decision maker's utility directly, then no such a characterization exists that will not amount to rephrasing the definition of robustness relative to the given utility function.

What does seem possible, and extremely desirable, is to develop characterizations of robustness for *classes* of games. That is, to pursue characterizations that lie between our context-free approach and the other extreme of full context dependence. One would fix a class of games satisfying a certain property, and ask what is the weakest sufficient condition (on sequences of perturbations) for robust predictions in all games satisfying the property. Thus, the characterization is allowed to rely on some property of payoffs, but not on further details of the payoff functions.[37]

While developing characterizations for classes of games is important, we anticipate some challenges. For illustration, consider the following single-agent problem. Assume that the agent has two actions to choose: *high* or *low*. The agent's payoffs depend on a fundamental variable θ that takes finitely many possible values in the real line. A modeler believes that a single-crossing condition holds: *high* is strictly better for the agent if θ is no less than a fixed cutoff $\overline{\theta}$, while *low* is strictly better if θ is below $\overline{\theta}$. Thus, only games that satisfy this single-crossing condition (relative

[37]Such an approach has been pursued recently by Morris, Shin, and Yildiz (2015) in the context of global games.

to a fixed order on Θ and a fixed threshold $\overline{\theta}$) would concern the modeler. Denote this class of games by $\mathcal{G}$. Say that a sequence of types/beliefs μ^n converges to μ $\mathcal{G}$-strategically if for every game $G \in \mathcal{G}$ and every action a in G, a is a best reply for μ in G if and only if for every $\varepsilon > 0$ there exists N such that for every $n \geqslant N$, a is an ε-best response for μ^n in G. For this class of decision problems, we can show that $\mathcal{G}$-strategic convergence is equivalent to the following condition: $\mu^n(\{\theta : \theta \geqslant \overline{\theta}\}) \to \mu(\{\theta : \theta \geqslant \overline{\theta}\})$ if $\mu(\{\theta : \theta \geqslant \overline{\theta}\}) = 0$ or 1, and $\mu^n(\theta) \to \mu(\theta)$ for every θ if $\mu(\{\theta : \theta \geqslant \overline{\theta}\}) \in (0, 1)$.[38] While this suggests that characterizing robustness in terms of beliefs is indeed possible for some classes of games, we note that even for this very simple class of one-player games satisfying a single-crossing condition the characterization is somewhat unwieldy, in that it rests sometimes on beliefs over coarse events, and sometimes on beliefs over fine events, depending on properties of the limit belief. In general, we do not know what properties of classes of games are likely to lead to tractable characterizations of robustness in terms of beliefs. This is an important direction for future research.

It is worth noting that the tension between the sharpness of a sufficient condition and the complexity of its verification process is rather common in economic theory. A notable example, which serves as a good analogy to our exercise, is Blackwell's informativeness theorem on the comparison of statistical experiments (Blackwell, 1951, 1953). If two experiments, A and B, are ranked by Blackwell's criterion so that B is a garbling of A, then the decision maker must earn a higher expected utility from experiment A than from B. But what if B is not a garbling of A? Blackwell's theorem only asserts that there will be *some* payoff function under which the decision maker's payoff from B is higher than from A. Nonetheless, Blackwell's theorem is generally regarded as useful, as the garbling condition is both sharp and easy to check.[39] The literature on comparison of statistical experiments has also found it useful to complement Blackwell's theorem with analogous results that focus on subclasses of payoffs functions, as in Lehmann (1988) and, more recently, Quah and Strulovici (2009).

Finally, we remark that there are some settings, such as mechanism design, where it is desirable to achieve robustness of strategic behavior for *all* payoff functions (possibly, in a class).

[38] A proof is available upon request. Note that belief convergence, i.e., $\mu^n(\theta)$ converges to $\mu(\theta)$ for every θ, implies strategic convergence but it is unnecessarily strong, whereas convergence of $\mu^n(\{\theta : \theta \geq \overline{\theta}\})$ to $\mu(\{\theta : \theta \geq \overline{\theta}\})$ and $\mu^n(\{\theta : \theta < \overline{\theta}\})$ to $\mu(\{\theta : \theta < \overline{\theta}\})$ is a necessary requirement, but it is too weak to ensure strategic convergence. To see the former, note that for θ' and θ'' larger than $\overline{\theta}$, the sequence of point masses at θ'' converges to the point mass at θ' $\mathcal{G}$-strategically but not in beliefs; to see the latter, consider $\Theta = \{-1, 0, 1\}$ and $\overline{\theta} = 0$ such that $u(a, 1) - u(a', 1) = 2$, $u(a, 0) - u(a', 0) = 1$, and $u(a, -1) - u(a', -1) = -2$. Let $\mu''(\theta = 0)$, $\mu'(\theta = 1)$, $\mu''(\theta = -1)$, and $\mu'(\theta = -1)$ be all equal to $1/2$. Then, μ'' and μ' assign the same probability on the events $\{\theta \geq \overline{\theta}\}$ and $\{\theta < \overline{\theta}\}$; however, the sequence (μ^n) with $\mu_n = \mu''$ for every n does not converge to μ' $\mathcal{G}$-strategically.

[39] Other examples of similar approaches in economic theory are stochastic orders, such as first-order and second-order stochastic dominance. Also, the revealed preference exercise in decision theory is predicated on similar ideas: although the analyst might be interested on a small class of decision problems, the axioms typically assume knowledge of the decision maker's behavior in a universal domain of decisions that might go well beyond the context of interest.

5.2 Robustness under Non-Strict Incentives

The definition of strategic convergence, on which our robustness exercise is based, requires that in every game, every action that is rationalizable for a type remains, for $\varepsilon > 0$ arbitrarily small, ε-rationalizable for any type sufficiently far in the tail of the sequence of perturbations. This condition is equivalent to requiring *strictly* rationalizable actions for a type to remain rationalizable for small perturbations. It is thus natural to wonder what can be said regarding stronger notions of robustness that require rationalizable actions that rely on indifferences to also be robust. The answer will depend on the notion of robustness. For instance, if we ask what sequences of types converge in the sense that every action is robustly rationalizable in every game, then nothing interesting can be said: only constant sequences converge, that is, we get the discrete topology.

Nonetheless, we do think it would be fruitful to study weaker notions of robustness of non-strictly rationalizable actions, and we believe that the strategic topology should play a central role in such analysis. What we have in mind is an analogy with the refinements literature on Nash equilibrium, which aims at characterizing strategies (or connected components of strategies) that are robust to small perturbations of behavior (the "trembles"), or equivalently, perturbations of payoffs (Kohlberg and Mertens, 1986). In the refinements literature, the space of perturbations is finite-dimensional, so perturbations simply mean convergence in the Euclidean topology, which also turns out to be the weak topology generated by continuity of the *strict* Nash equilibrium correspondence. In other words, in the refinements literature, strict equilibria are, by definition, robust. And the only question is: in a given game, which non-strict equilibria are robust to perturbations in the topology that renders strict equilibria robust in every game (i.e. the Euclidean topology)? A similar question can be asked in the context of our exercise, which considers perturbations of the players' beliefs and higher-order beliefs. By analogy, it seems natural to expect that in order to make progress in understanding which (possibly non-strictly) rationalizable actions are robust in a given game, we first need to understand the weak topology on types generated by the continuity of the correspondence of strictly rationalizable actions, which is the analogue of the Euclidean topology used in the refinements literature.

5.3 Extension to Compact-Continuous Games

Interim correlated rationalizability is well defined when the space Θ and the action sets A_i are compact metrizable and the payoff functions are continuous: the recursive and fixed-point definitions coincide, and are non-empty valued.

The characterization of the strategic topology in terms of uniform weak convergence on frames (Theorem 2) continues to hold in the compact-continuous case. The characterization of uniform strategic convergence (Theorem 1) also remains valid, provided the definition of uniform strategic convergence is modified to require the rate of convergence to be uniform on all classes of compact continuous games whose payoffs are not only uniformly bounded, but also equicontinuous in θ.

A Appendix

In the appendix we use the following notations. Given two measurable spaces X, Y and a probability measure $\mu \in \Delta(X \times Y)$, we write $\operatorname{marg}_X \mu$ for the marginal of the distribution μ on X, that is, for every measurable $E \subseteq X$, $(\operatorname{marg}_X \mu)(E) = \mu(E \times Y)$. Given a measurable correspondence $\varsigma : X \rightrightarrows Y$, we define graph $\varsigma = \{(x, y) \in X \times Y : y \in \varsigma(x)\}$. If X is finite and ν is a probability measure on X, we write supp ν to denote the support of ν, that is, $\operatorname{supp} \nu = \{x \in X : \nu(x) > 0\}$. Finally, given a game $G = (A_i, g_i)_{i \in I}$, a probability distribution $\beta \in \Delta(\Theta \times A_{-i})$ and $\varepsilon \geqslant 0$, we write $BR_i(\beta, G, \varepsilon)$ for the set of ε-best replies to β, that is, the set of all $a_i \in A_i$ such that $g_i(a_i, \beta) \geqslant g_i(a_i', \beta) - \varepsilon$ for all $a_i' \in A_i$.

A.1 Properties of rationalizability

Our proofs below use the following properties of interim correlated rationalizability. First, in the definition of the k-order ε-rationalizable actions of a type t_i we can replace the conjectures of the form $\nu \in \Delta(\Theta \times T_{-i}^* \times A_{-i})$ with those of the form $\sigma_{-i} : \Theta \times T_{-i}^* \to \Delta(A_{-i})$. The two kinds of conjectures are related by the disintegration formula:

$$\nu(\theta \times E_{-i} \times a_{-i}) = \int_{E_{-i}} \sigma_{-i}(\theta, t_{-i})[a_{-i}]\, \mu_i^*(\theta \times dt_{-i} | t_i),$$

for every $\theta \in \Theta$, $a_{-i} \in A_{-i}$ and measurable subset $E_{-i} \subseteq T_{-i}^*$.[40] Thus, $R_i^k(t_i, G, \varepsilon)$ is the set of all $a_i \in A_i$ for which there is a measurable function $\sigma_{-i} : \Theta \times T_{-i}^* \to \Delta(A_{-i})$ that satisfies:

$$\operatorname{supp} \sigma_{-i}(\theta, t_{-i}) \subseteq R_{-i}^{k-1}(t_{-i}, G, \varepsilon) \qquad \forall (\theta, t_{-i}) \in \Theta \times T_{-i}^*, \tag{2}$$
$$\int_{\Theta \times T_{-i}^*} \left[g_i(a_i, \sigma_{-i}(\theta, t_{-i}), \theta) - g_i(a_i', \sigma_{-i}(\theta, t_{-i}), \theta) \right] d\mu_i^*(\theta, t_{-i} | t_i) \geqslant -\varepsilon \qquad \forall a_i' \in A_i.$$

(See Corollary 1 in Dekel, Fudenberg, and Morris, 2007.) For future reference, a measurable function σ_{-i} that satisfies (2) is called a $(k-1)$-*order ε-rationalizable conjecture*; a measurable function σ_{-i} such that $\operatorname{supp} \sigma_{-i}(\theta, t_{-i}) \subseteq R_{-i}(t_{-i}, G, \varepsilon)$ for every $(\theta, t_{-i}) \in \Theta \times T_{-i}^*$ is called an *ε-rationalizable conjecture*.

Second, we can define rationalizability in terms of iterated eliminations of strongly dominated actions. For each $i \in I$ and $t_i \in T_i^*$, let $S_i^0(t_i, G, \varepsilon) = A_i$ and, recursively for $k \geqslant 1$, define $S_i^k(t_i, G, \varepsilon)$ as the set of all $a_i \in A_i$ such that for every *mixed* deviation $\alpha_i \in \Delta(A_i)$ there is a

[40] Given a ν that satisfies $\operatorname{marg}_{\Theta \times T_{-i}^*} \nu = \mu_i^*(t_i)$, the disintegration fomula only pins down σ_{-i} up to a set of $\mu_i^*(t_i)$-probability zero. But, outside this null set, we can set σ_{-i} equal to a measurable selection from the correspondence $R_{-i}^{k-1}(\cdot, G, \varepsilon)$, thus ensuring that (2) is satisfied everywhere provided $\nu(\Theta \times \operatorname{graph} R_{-i}^{k-1}) = 1$ (as opposed to almost everywhere). The fact that such a measurable selection exists follows from the upper hemi-continuity of $R_{-i}^{k-1}(\cdot, G, \varepsilon)$ (in the product topology) and the Kuratowsky-Nyll-Nardzewski Selection Theorem.

conjecture $\sigma_{-i} : \Theta \times T^*_{-i} \to \Delta(A_{-i})$ that satisfies:

$$\operatorname{supp} \sigma_{-i}(\theta, t_{-i}) \subseteq S^{k-1}_{-i}(t_{-i}, G, \varepsilon) \qquad \forall(\theta, t_{-i}) \in \Theta \times T^*_{-i}, \quad \text{and}$$
$$\int_{\Theta \times T^*_{-i}} \left[g_i(a_i, \sigma_{-i}(\theta, t_{-i}), \theta) - g_i(\alpha_i, \sigma_{-i}(\theta, t_{-i}), \theta)\right] d\mu^*_i(\theta, t_{-i}|t_i) \geqslant -\varepsilon. \tag{3}$$

Then,

$$R^k_i(t_i, G, \varepsilon) = S^k_i(t_i, G, \varepsilon),$$

and thus,

$$R_i(t_i, G, \varepsilon) \;=\; \bigcap_{k \geqslant 1} S^k_i(t_i, G, \varepsilon).$$

(See Chen, Di Tillio, Faingold, and Xiong (2010), Proposition 1.) Likewise, $a_i \in R_i(t_i, G, \varepsilon)$ if and only if, for every $\alpha_i \in \Delta(A_i)$, there is an ε-rationalizable conjecture $\sigma_{-i} : \Theta \times T^*_{-i} \to \Delta(A_i)$ that satisfies (3).

Finally, rationalizability has a characterization in terms of best-reply sets. Say a profile of correspondences $\varsigma_i : T^*_i \rightrightarrows A_i$, $i \in I$, has the *ε-best-reply property* if for each $i \in I$, $t_i \in T^*_i$ and $a_i \in \varsigma_i(t_i)$ there exists $\nu \in \Delta(\Theta \times T^*_{-i} \times A_{-i})$ such that the following conditions hold:

$$\operatorname{marg}_{\Theta \times T^*_{-i}} \nu = \mu^*_i(t_i), \quad \left(\operatorname{marg}_{T^*_{-i} \times A_{-i}} \nu\right)\left[\operatorname{graph} \varsigma_{-i}\right] = 1,^{41} \quad a_i \in BR_i\left(\operatorname{marg}_{A_{-i} \times \Theta} \nu, G, \varepsilon\right).$$

Then $R_i(\cdot, G, \varepsilon)$ is the greatest (w.r.t. pointwise set inclusion) profile of correspondences with the ε-best-reply property.[42] (See Claim 2 in Dekel, Fudenberg, and Morris, 2007.)

A.2 Proof of Theorem 1

The *if* direction is proved in Chen, Di Tillio, Faingold, and Xiong (2010). The proof of the *only if* direction relies on Lemma 1, Corollary 3 and Lemma 2 below.

We begin with some useful definitions and notations. Given a game $G = (A_i, g_i)_{i \in I}$, say that an action $a^0_i \in A_i$ is a *zero action* for player i if $g_i\left(a^0_i, a_{-i}, \theta\right) = 0$ for all θ and a_{-i}. An action profile $a^c = (a^c_i, a^c_{-i})$ is a *coordination pair* if $g_i\left(a^c_i, a^c_{-i}, \theta\right) = \max g_i$ for all θ and i. (In particular, any action of player i that is part of a coordination pair is rationalizable for any type of player i.) Finally, let $\beta_{t_i, \sigma_{-i}} \in \Delta\left(A_{-i} \times \Theta\right)$ denote the belief of type t_i over actions of player $-i$ and payoff-relevant parameters, when he has conjecture $\sigma_{-i} : \Theta \times T^*_{-i} \to \Delta(A_{-i})$, that is,

$$\beta_{t_i, \sigma_{-i}}\left(a_{-i}, \theta\right) := \int_{T^*_{-i}} \sigma_{-i}(\theta, t_{-i})\left[a_{-i}\right] \mu^*_i(\theta \times dt_{-i}|t_i) \quad \forall(a_{-i}, \theta) \in A_{-i} \times \Theta.$$

[41] When graph ς_i is not measurable, this expression is taken to mean that $\operatorname{marg}_{T_{-i} \times A_{-i}} \nu_i$ assigns probability one to a measurable subset of graph ς_i.

[42] Such greatest profile of correspondences is well defined, because the pointwise union of any family of profiles of correspondences with the ε-best-reply property must also have the ε-best-reply property.

The only if part of the theorem is a direct implication of Lemma 2 below. Lemma 1 and Corollary 3 are intermediate results.

Lemma 1. *For every $\varepsilon > 0$, integer $k \geqslant 1$, player j and finite set of finite types $\{t_{j,1}, t_{j,2}, \ldots, t_{j,N}\} \subset T_j^*$, there is a game $G = (A_i, g_i)_{i \in I}$ with payoffs in the interval $[-5, 3]$, and a set of actions $\{a_{j,1}^*, a_{j,2}^*, \ldots, a_{j,N}^*\} \subset A_j$, such that:*

(i) every player i has a zero action $a_i^0 \in A_i$;

(ii) there is a coordination pair $a^c \in A_1 \times A_2$ such that $a_j^c \notin \left\{a_{j,1}^*, a_{j,2}^*, ..., a_{j,N}^*\right\}$ and $g_i\left(a_i, a_{-i}^c, \theta\right) \geqslant -2$ for every $a_i \in A_i$, $\theta \in \Theta$ and $i \in I$.

(iii) $R_i(\cdot, G, \gamma) = R_i^k(\cdot, G, \gamma)$ for every $i \in I$ and $\gamma \in \left[0, \frac{1}{2}\right)$;

(iv) for every $1 \leqslant n \leqslant N$, $a_{j,n}^* \in R_j\left(t_{j,n}, G\right)$;

(v) for every $1 \leqslant n \leqslant N$ and $s_j \in T_j^*$ with $d_j^k\left(t_{j,n}, s_j\right) > \varepsilon$, $a_{j,n}^* \notin R_j\left(s_j, G, \frac{\varepsilon}{2}\right)$;

Proof. The proof is by induction on k. Consider first $k = 1$ and $j = 1$ ($j = 2$ can be similarly proved). Fix a finite set of player 1's types $\left\{t_{1,1}, t_{1,2}, ..., t_{1,N}\right\}$. Enumerate the nonempty subsets of Θ as E_1, E_2,..., E_L. For each $(n, \ell) \in \{1, 2, \ldots, N\} \times \{0, 1, \ldots, L\}$ consider the function $\phi_{n,\ell} : \Theta \to [-1, 1]$ described in the following table:

	$\theta \in E_\ell$	$\theta \notin E_\ell$
$\ell = 0$	0	0
$\ell \geqslant 1$	$-\left(1 - \mu_1\left(E_\ell \vert t_{1,n}\right)\right)$	$\mu_1\left(E_\ell \vert t_{1,n}\right)$

Thus, the functions $\phi_{n,\ell}$ define an auxiliary game between player 1 and Nature, where $\ell = 0$ is a safe bet for player 1, and $\ell \geqslant 1$ is a risky bet on the event $\theta \notin E_\ell$. The rewards of the risky bets are such that:

- any type that has the same first-order beliefs as type $t_{1,n}$ is exactly indifferent between $\ell = 0$ and any $\ell \geqslant 1$;
- any type whose first-order belief is different from that of $t_{1,n}$ strictly prefers some risky bet $\ell \geqslant 1$ than the safe bet $\ell = 0$.

We use the functions $\phi_{n,\ell}$ to construct a game $G = (A_i, g_i)_{i \in I}$ to prove our claim for $k = 1$. In this game,

$$A_1 = \left(\{1, 2, \ldots, N\} \times \{0, 1, \ldots, L\}\right) \dot{\cup} \left\{a_1^0, a_1^c\right\} \text{ and } A_2 = \{1, 2, \ldots, N\} \dot{\cup} \left\{a_2^0, a_2^c\right\}.$$

Player 1's payoffs are specified as follows:

- a_1^0 is a zero action for player 1;
- if player 1 chooses a_1^c, she gets 3 if player 2 chooses a_2^c, and gets 0 otherwise (regardless of θ);
- if player 1 chooses $(n, \ell) \in \{1, \dots, N\} \times \{0, \dots, L\}$ and the state is θ, she gets $\phi_{n,\ell}(\theta)$ if player 2 chooses n, and she gets $\phi_{n,\ell}(\theta) - 1$ if player 2 chooses any action different from n.

Player 2's payoffs are specified as follows:

- Player 2 gets 3 if (a_1^c, a_2^c) is chosen (regardless of θ), and gets 0 otherwise.

Thus, when the state is θ, we can draw the payoff matrix $g(\cdot, \cdot, \theta)$ as follows:

	1	2	$\cdots$	N	a_2^0	a_2^c
$(1,\ell)$	$\phi_{1,\ell}(\theta), 0$	$\phi_{1,\ell}(\theta) - 1, 0$	$\cdots$	$\phi_{1,\ell}(\theta) - 1, 0$	$\phi_{1,\ell}(\theta) - 1, 0$	$\phi_{1,\ell}(\theta) - 1, 0$
$(2,\ell)$	$\phi_{2,\ell}(\theta) - 1, 0$	$\phi_{2,\ell}(\theta), 0$	$\cdots$	$\phi_{2,\ell}(\theta) - 1, 0$	$\phi_{2,\ell}(\theta) - 1, 0$	$\phi_{2,\ell}(\theta) - 1, 0$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
(N,ℓ)	$\phi_{N,\ell}(\theta) - 1, 0$	$\phi_{N,\ell}(\theta) - 1, 0$	$\cdots$	$\phi_{N,\ell}(\theta), 0$	$\phi_{N,\ell}(\theta) - 1, 0$	$\phi_{N,\ell}(\theta) - 1, 0$
a_1^0	$0, 0$	$0, 0$	$\cdots$	$0, 0$	$0, 0$	$0, 0$
a_1^c	$0, 0$	$0, 0$	$\cdots$	$0, 0$	$0, 0$	$3, 3$

Since $\phi_{n,\ell}(\theta) \in [-1, 1]$ for all n, ℓ and θ, game G has payoffs bounded between -2 and 3. We claim that G, along with the actions $a_{1,n}^* := (n, 0)$, $n = 1, \dots, N$, satisfy properties (i)–(v) for $k = 1$. Properties (i) and (ii) clearly hold. To prove (iii), first note that $R_2(\cdot, G, \gamma) = R_2^1(\cdot, G, \gamma) = A_2$ for all $\gamma \geqslant 0$. Indeed, the profile of correspondences $(\varsigma_i)_{i \in I}$, where $\varsigma_2(t_2) = A_2$ and $\varsigma_1(t_1) = \{a_1^0, a_1^c\}$ for all t_1 and t_2, has the γ-best reply property. It follows that $R_1(\cdot, G, \gamma) = R_1^1(\cdot, G, \gamma)$ for all $\gamma \geqslant 0$. Hence, (iii) holds.

It remains to prove (iv) and (v). First, $(n, 0)$ is rationalizable for $t_{1,n}$, since given the conjecture that player 2 plays n, type $t_{1,n}$ gets 0 by playing (n, ℓ) for any ℓ:

$$\mu_1(E_\ell | t_{1,n}) \cdot \left(-\left(1 - \mu_1(E_\ell | t_{1,n})\right)\right) + \left(1 - \mu_1(E_\ell | t_{1,n})\right) \cdot \mu_1(E_\ell | t_{1,n}) = 0,$$

and gets at most 0 by playing any action not in $\{(n, 0), \dots, (n, L)\}$. Thus, (iv) holds for $a_{1,n}^* = (n, 0)$. Second, consider any type s_1 with $d_1^1(t_{1,n}, s_1) > \varepsilon$. Then, there exists some $1 \leqslant \ell \leqslant L$ such that $\mu_1(E_\ell | t_{1,n}) > \mu_1(E_\ell^{\varepsilon,0} | s_1) + \varepsilon = \mu_1(E_\ell | s_1) + \varepsilon$.[43] Then, given any conjecture about the behavior of player 2, the difference in expected payoffs between (n, ℓ) and $(n, 0)$ for type s_1 is

$$\mu_1(E_\ell | s_1) \cdot \left(-\left(1 - \mu_1(E_\ell | t_{1,n})\right)\right) + \left(1 - \mu_1(E_\ell | s_1)\right) \cdot \mu_1(E_\ell | t_{1,n}) = \mu_1(E_\ell | t_{1,n}) - \mu_1(E_\ell | s_1) > \varepsilon.$$

[43] The equality follows because $d^0 \equiv 0$ and Θ is endowed with the discrete metric.

Hence, $a^*_{1,n} = (n, 0)$ is not ε-rationalizable for type s_1, which proves (v).

We now prove our claim for $k+1$ assuming that it holds for k. Again, we assume $j = 1$, and the proof for $j = 2$ is similar. Let $t_{1,1}, \ldots, t_{1,N}$ be arbitrary finite types of player 1. Consider the finite set $T_2 = \{t_{2,1}, t_{2,2}, \ldots, t_{2,N'}\}$ of all types of player 2 that are assigned positive probability by some $t_{1,n}$, for $n = 1, \ldots, N$. By the induction hypothesis, we can find a game $G = (A_i, g_i)_{i \in I}$, a set of actions $\{a^*_{2,1}, a^*_{2,2}, \ldots, a^*_{2,N'}\} \subset A_2$, and action profiles a^0 and a^c in G, that satisfy properties (i)–(v) relative to the finite set of finite types T_2.

Let $T_2^k = \{t^k_{2,1}, t^k_{2,2}, \ldots, t^k_{2,N'}\}$ be the set of k-order beliefs of types in T_2. Enumerate the nonempty subsets of $\Theta \times T_2^k$ as E_1, E_2,..., E_L. For each $1 \leqslant \ell \leqslant L$, define

$$F_\ell = \left\{\left(a^*_{2,n'}, \theta\right) : 1 \leqslant n' \leqslant N', \left(\theta, t^k_{2,n'}\right) \in E_\ell\right\}.$$

For each $(n, \ell) \in \{1, \ldots, N\} \times \{0, \ldots, L\}$, define a function $\phi_{n,\ell} : A_2 \times \Theta \to [-1, 1]$ as in the following table:

	$(a_2, \theta) \in F_\ell$	$(a_2, \theta) \notin F_\ell$
$\ell = 0$	0	0
$\ell \geqslant 1$	$-\left(1 - \mu_1(E_\ell \mid t_{1,n})\right)$	$\mu_1(E_\ell \mid t_{1,n})$

We use the game G and the functions $\phi_{n,\ell}$ to define a new game $\bar{G} = \left(\bar{A}_i, \bar{g}_i\right)$ to prove our claim for $k+1$. In this game,

$$\bar{A}_1 = A_1 \dot{\cup} \left(\{1, \ldots, N\} \times \{0, 1, \ldots, L\}\right) \dot{\cup} \left\{\bar{a}^0_1\right\} \text{ and } \bar{A}_2 = A_2 \times \{0, 1, \ldots, N\}.$$

Player 1's payoffs are specified as follows (see also the following table):

- $\bar{a}^0_1$ is a zero action for player 1;
- if player 1 chooses $a_1 \in A_1$ and the state is θ, he gets $g_1(a_1, a_2, \theta)$ if player 2 chooses $(a_2, 0) \in A_2 \times \{0\}$, and gets $g_1\left(a_1, a^c_2, \theta\right) - 3$ otherwise.
- if player 1 chooses (n, ℓ) and the state is θ, he gets $\frac{\phi_{n,\ell}(a^c_2, \theta)}{2} - 1$ if player 2 chooses $(a_2, 0) \in A_2 \times \{0\}$; he gets $\phi_{n,\ell}(a_2, \theta)$ if player 2 chooses $(a_2, n) \in A_2 \times \{n\}$; and he gets $\phi_{n,\ell}\left(a^c_2, \theta\right) - 1$ if player 2 chooses $(a_2, m) \in A_2 \times \{m\}$ with $m \neq n$ and $m \neq 0$.

	$A_2 \times \{0\}$	$A_2 \times \{1\}$	$A_2 \times \{2\}$	$\cdots$	$A_2 \times \{N\}$
A_1	$g_1(a_1, a_2, \theta)$	$g_1\left(a_1, a^c_2, \theta\right) - 3$	$g_1\left(a_1, a^c_2, \theta\right) - 3$	$\cdots$	$g_1\left(a_1, a^c_2, \theta\right) - 3$
$(1, \ell)$	$\frac{\phi_{1,\ell}(a^c_2, \theta)}{2} - 1$	$\phi_{1,\ell}(a_2, \theta)$	$\phi_{1,\ell}\left(a^c_2, \theta\right) - 1$	$\cdots$	$\phi_{1,\ell}\left(a^c_2, \theta\right) - 1$
$(2, \ell)$	$\frac{\phi_{2,\ell}(a^c_2, \theta)}{2} - 1$	$\phi_{2,\ell}\left(a^c_2, \theta\right) - 1$	$\phi_{2,\ell}(a_2, \theta)$	$\cdots$	$\phi_{2,\ell}\left(a^c_2, \theta\right) - 1$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
(N, ℓ)	$\frac{\phi_{N,\ell}(a^c_2, \theta)}{2} - 1$	$\phi_{N,\ell}\left(a^c_2, \theta\right) - 1$	$\phi_{N,\ell}\left(a^c_2, \theta\right) - 1$	$\cdots$	$\phi_{N,\ell}(a_2, \theta)$
$\bar{a}^0_1$	0	0	0	$\cdots$	0

Player 2's payoffs are specified as follows:

- If player 2 chooses (a_2, m), he gets $g_2(a_2, a_1, \theta)$ if player 1 chooses $a_1 \in A_1$, and he gets $g_2\left(a_2, a_1^c, \theta\right)$ otherwise.

By the induction hypothesis, game G satisfies property (ii) and has payoffs in the interval $[-5, 3]$; it follows that game $\bar{G}$ also has payoffs in the interval $[-5, 3]$.

We now prove that game $\bar{G}$, along with the actions $a_{1,n}^* := (n, 0)$, $n = 1, \ldots, N$, satisfy properties (i)–(v) for $k+1$. First, (i) follows because $\bar{a}_1^0$ and (a_2^0, n) are zero actions for players 1 and 2, respectively. Second, (ii) is satisfied by the coordination pair $\left(a_1^c, \left(a_2^c, 0\right)\right)$: for any $\bar{a}_1 \in \bar{A}_1$ and $\theta \in \Theta$,

$$\bar{g}_1\left(\bar{a}_1, \left(a_2^c, 0\right), \theta\right) \geqslant \min\left\{g_1\left(a_1, a_2^c, \theta\right), \min_{n,\ell} \phi_{n,\ell}\left(a_2^c, \theta\right)/2 - 1, 0\right\} \geqslant -2;$$

moreover, $\bar{g}_2\left(\left(a_2, n\right), a_1^c, \theta\right) = g_2\left(a_2, a_1^c, \theta\right) \geqslant -2$ for any $(a_2, n) \in \bar{A}_2$ and $\theta \in \Theta$, and $\bar{g}_1\left(a_1^c, \left(a_2^c, 0\right), \theta\right) = \bar{g}_2\left(\left(a_2^c, 0\right), a_1^c, \theta\right) = 3$ for any $\theta \in \Theta$.

The proof of (iii)–(v) relies on the following claim, whose proof is postponed.

Claim 1. *For every integer $r \geqslant 0$ and $\gamma \in [0, 1/2)$,*

1. $R_1^r\left(\cdot, \bar{G}, \gamma\right) \cap A_1 = R_1^r\left(\cdot, G, \gamma\right)$*;*
2. $R_2^r\left(\cdot, \bar{G}, \gamma\right) = R_2^r\left(\cdot, G, \gamma\right) \times \{0, 1, 2, \ldots, N\}$.

We now prove that Claim 1 implies properties (iii)–(v).

(iii): By the induction hypothesis, $R_2\left(\cdot, G, \gamma\right) = R_2^k\left(\cdot, G, \gamma\right)$. Thus, Claim 1 implies that $R_2\left(\cdot, \bar{G}, \gamma\right) = R_2^k\left(\cdot, \bar{G}, \gamma\right)$. This, in turn, implies $R_1\left(\cdot, \bar{G}, \gamma\right) = R_1^{k+1}\left(\cdot, \bar{G}, \gamma\right)$ and hence (iii).

(iv): Given any $1 \leqslant n \leqslant N$, consider the conjecture $\sigma_2 : \Theta \times T_2 \to \Delta(\bar{A}_2)$ such that $\sigma_2\left(\theta, t_{2,n'}\right)\left[a_{2,n'}^*, n\right] = 1$ for each $n' = 1, \ldots N'$ and $\theta \in \Theta$. By part 1 of Claim 1 and the fact that G satisfies property (iv) (by the induction hypothesis), σ_2 is a rationalizable conjecture in $\bar{G}$. Moreover, given such a conjecture, $t_{1,n}$ gets an expected payoff of 0 by playing (n, ℓ) for any ℓ, and gets at most 0 by playing any action in $\bar{A}_1 \setminus \{(n, 0), \ldots, (n, L)\}$:

$$\begin{aligned}&\beta_{t_{1,n},\sigma_2}(F_\ell) \cdot \left(-\left(1 - \mu_1\left(E_\ell | t_{1,n}\right)\right)\right) + \left(1 - \beta_{t_{1,n},\sigma_2}(F_\ell)\right) \cdot \mu_1\left(E_\ell | t_{1,n}\right) \\ &\qquad = \mu_1\left(E_\ell | t_{1,n}\right) \cdot \left(-\left(1 - \mu_1\left(E_\ell | t_{1,n}\right)\right)\right) + \left(1 - \mu_1\left(E_\ell | t_{1,n}\right)\right) \cdot \mu_1\left(E_\ell | t_{1,n}\right) = 0.\end{aligned}$$

In particular, $a_{1,n}^* = (n, 0)$ is a best reply for $t_{1,n}$.

(v): Fix $1 \leqslant n \leqslant N$ and consider any type s_1 with $d_1^{k+1}\left(t_{1,n}, s_1\right) > \varepsilon$. Then, there exists some $1 \leqslant \ell \leqslant L$ such that

$$\mu_1\left(E_\ell | t_{1,n}\right) > \mu_1\left(\left(E_\ell\right)^{\varepsilon,k} \middle| s_1\right) + \varepsilon.$$

It follows that, given any $\frac{\varepsilon}{2}$-rationalizable conjecture $\sigma_2 : \Theta \times T_2^* \to \Delta(\bar{A}_2)$, the difference in expected payoffs between actions (n, ℓ) and $(n, 0)$ for type s_1 is at least $\varepsilon/2$. To prove this, we consider two cases separately: when player 2 chooses $m = n$; and when player 2 chooses $m \neq n$.

First, conditional on player 2 choosing n, the expected payoff difference between actions (n, ℓ) and $(n, 0)$ for type s_1, given an arbitrary $\frac{\varepsilon}{2}$-rationalizable conjecture σ_2, is

$$\sum_{a_2,\theta} \phi_{n,\ell}(a_2, \theta)\beta_{s_1,\sigma_2}(a_2, \theta|n)$$
$$= \beta_{s_1,\sigma_2}(F_\ell|n) \cdot (-(1 - \mu_1(E_\ell|t_{1,n}))) + (1 - \beta_{s_1,\sigma_2}(F_\ell|n)) \cdot \mu_1(E_\ell|t_{1,n}).$$

But, for any $(\theta, s_2) \in \Theta \times T_2^*$ and $n' = 1, \ldots, N'$,

$$\begin{aligned} \sigma_2(\theta, s_2)[a^*_{2,n'}, n] > 0 \quad &\Rightarrow \quad (a^*_{2,n'}, n) \in R_2\left(s_2, \bar{G}, \frac{\varepsilon}{2}\right) && \text{(since } \sigma_2 \text{ is } \tfrac{\varepsilon}{2}\text{-rationalizable)} \\ &\Rightarrow \quad a^*_{2,n'} \in R_2\left(s_2, G, \frac{\varepsilon}{2}\right) && \text{(by Claim 1)} \\ &\Rightarrow \quad d_2^k(t_{2,n'}, s_2) \leqslant \varepsilon, && \text{(by the induction hypothesis)} \end{aligned}$$

and thus, if $\beta_{s_1,\sigma_2}(n) > 0$,

$$\beta_{s_1,\sigma_2}(F_\ell|n) \leqslant \mu_1\left((E_\ell)^{\varepsilon,k}|s_1\right),$$

which implies

$$\begin{aligned} &\sum_{a_2,\theta} \phi_{n,\ell}(a_2, \theta)\beta_{s_1,\sigma_2}(a_2, \theta|n) \\ &\geqslant \mu_1\left((E_\ell)^{\varepsilon,k}|s_1\right) \cdot (-(1 - \mu_1(E_\ell|t_{1,n}))) + \left(1 - \mu_1\left((E_\ell)^{\varepsilon,k}|s_1\right)\right) \cdot \mu_1(E_\ell|t_{1,n}) \\ &= \mu_1(E_\ell|t_{1,n}) - \mu_1\left((E_\ell)^{\varepsilon,k}|s_1\right) > \varepsilon. \end{aligned}$$

Second, conditional on player 2 choosing $m \neq n$, the expected payoff difference between actions (n, ℓ) and $(n, 0)$ for type s_1 (given any conjecture) is at least $\phi_{n,\ell}(a_2^c, \theta)/2 = \mu_1(E_\ell|t_{1,n})/2 > \varepsilon/2$. (This is because $a_2^c \neq a^*_{2,n'}$ for all n', and hence $(a_2^c, \theta) \notin F_\ell$ for every θ.) We have thus shown that, given any $\varepsilon/2$-rationalizable conjecture, and conditional on any choice of $m = 0, \ldots, N$ by player 2 with $\beta_{s_1,\sigma_2}(m) > 0$, type s_1 gains at least $\varepsilon/2$ by deviating from $(n, 0)$ to (n, ℓ). Thus, he also gains $\varepsilon/2$ unconditionally on m, and hence property (v) follows.

To conclude the proof, it remains to prove Claim 1. We prove it by induction on $r \geqslant 0$. First, the claim is trivially true for $r = 0$. We now consider $r \geqslant 1$, assume that the claim holds for any $0 \leqslant r' < r$, and prove that it also holds for r.

$\underline{R_1^r(t_1, G, \gamma) \supset R_1^r(t_1, \bar{G}, \gamma) \cap A_1}$: Let $a_1 \in R_1^r(t_1, \bar{G}, \gamma) \cap A_1$. Then, there is an $(r-1)$-order γ-rationalizable conjecture $\bar{\sigma}_2 : \Theta \times T_2^* \to \Delta(\bar{A}_2)$ in $\bar{G}$ such that for any $a_1' \in A_1$,

$$\int_{\Theta \times T_2^*} d\mu_1(\theta, t_2|t_1) \sum_{(a_2,n) \in \bar{A}_2} \left[\bar{g}_1(a_1, (a_2, n), \theta) - \bar{g}_1(a_1', (a_2, n), \theta)\right] \bar{\sigma}_2(\theta, t_2)[(a_2, n)] \geqslant -\gamma. \tag{4}$$

Consider the mapping $\varphi_2 : \bar{A}_2 \to A_2$,

$$\varphi_2(a_2, n) = \begin{cases} a_2, & \text{if } n = 0; \\ a_2^c, & \text{if } n \neq 0. \end{cases}$$

Define σ_2 as the conjecture in G such that $\sigma_2(\theta, t_2)[a_2] = \bar{\sigma}_2(\theta, t_2)\left[\varphi_2^{-1}(a_2)\right]$ for each $(\theta, t_2) \in \Theta \times T_2^*$ and $a_2 \in A_2$. Since $\bar{\sigma}_2$ is $(r-1)$-order γ-rationalizable in $\bar{G}$, $\bar{\sigma}_2(\theta, t_2)[(a_2, 0)] > 0$ implies $(a_2, 0) \in R_2^{r-1}\left(t_2, \bar{G}, \gamma\right)$, and by the induction hypothesis, $a_2 \in R_2^{r-1}(t_2, G, \gamma)$. Moreover, a_2^c is part of a coordination pair in G, hence a_2^c is rationalizable in G for any type. Thus, σ_2 is an $(r-1)$-order γ-rationalizable conjecture in G. Moreover, for each $a_1' \in A_1$,

$$\begin{aligned}
&\int_{\Theta \times T_2^*} d\mu_1(\theta, t_2|t_1) \sum_{(a_2,n) \in \bar{A}_2} \left[\bar{g}_1(a_1, (a_2, n), \theta) - \bar{g}_1\left(a_1', (a_2, n), \theta\right)\right] \bar{\sigma}_2(\theta, t_2)[(a_2, n)] \\
&= \int_{\Theta \times T_2^*} d\mu_1(\theta, t_2|t_1) \Bigg(\sum_{a_2 \in A_2} \left[g_1(a_1, a_2, \theta) - g_1\left(a_1', a_2, \theta\right)\right] \bar{\sigma}_2(\theta, t_2)[(a_2, 0)] \\
&\qquad + \left[\left(g_1\left(a_1, a_2^c, \theta\right) - 3\right) - \left(g_1\left(a_1', a_2^c, \theta\right) - 3\right)\right] \bar{\sigma}_2(\theta, t_2)\left[\left\{(a_2, n) \in \bar{A}_2 : n > 0\right\}\right] \Bigg) \\
&= \int_{\Theta \times T_2^*} d\mu_1(\theta, t_2|t_1) \sum_{a_2 \in A_2} \left[g_1(a_1, a_2, \theta) - g_1\left(a_1', a_2, \theta\right)\right] \sigma_2(\theta, t_2)[a_2] \qquad (5)
\end{aligned}$$

Then, (4) and (5) imply $a_1 \in R_1^r(t_1, G, \gamma)$.

$R_1^r(t_1, G, \gamma) \subset R_1^r\left(t_1, \bar{G}, \gamma\right) \cap A_1$: Let $a_1 \in R_1^r(t_1, G, \gamma)$. Then, there exists an $(r-1)$-order γ-rationalizable conjecture $\sigma_2 : \Theta \times T_2^* \to \Delta(A_2)$ in G such that for each $a_1' \in A_1$,

$$\int_{\Theta \times T_2^*} d\mu_1(\theta, t_2|t_1) \sum_{a_2 \in A_2} \left[g_1(a_1, a_2, \theta) - g_1\left(a_1', a_2, \theta\right)\right] \sigma_2(\theta, t_2)[a_2] \geqslant -\gamma. \qquad (6)$$

Define $\bar{\sigma}_2$ as the conjecture in $\bar{G}$ such that $\bar{\sigma}_2(\theta, t_2)[(a_2, 0)] = \sigma_2(\theta, t_2)[a_2]$ for each $(\theta, t_2) \in \Theta \times T_2^*$ and $a_2 \in A_2$ (and thus $\bar{\sigma}_2(\theta, t_2)[(a_2, n)] = 0$ for each $n > 0$). Since σ_2 is $(r-1)$-order γ-rationalizable in G , $\sigma_2(\theta, t_2)[a_2] > 0$ implies $a_2 \in R_2^{r-1}(t_2, G, \gamma)$, and by the induction hypothesis, $(a_2, 0) \in R_2^{r-1}\left(t_2, \bar{G}, \gamma\right)$. Hence, $\bar{\sigma}_2$ is $(r-1)$-order γ-rationalizable in G. We will now show that a_1 is a γ-best reply to $\bar{\sigma}_2$ for t_1 in $\bar{G}$. First, by (6) and the definition of $\bar{\sigma}_2$,

$$\int_{\Theta \times T_2^*} d\mu_1(\theta, t_2|t_1) \sum_{a_2 \in A_2} \left[g_1(a_1, a_2, \theta) - g_1\left(a_1', a_2, \theta\right)\right] \bar{\sigma}_2(\theta, t_2)[(a_2, 0)] \geqslant -\gamma \quad \forall a_1' \in A_1. \qquad (7)$$

Second, setting $a_1' = a_1^0$ in (6) and recalling that $\gamma < 1/2$,

$$\int_{\Theta \times T_2^*} d\mu_1(\theta, t_2|t_1) \sum_{a_2 \in A_2} g_1(a_1, a_2, \theta)\, \sigma_2(\theta, t_2)[a_2] \geqslant -\gamma > -1/2.$$

Since $\phi_{n,\ell}\left(a_2^c, \theta\right)/2 - 1 \leqslant -1/2$ for all n and ℓ,

$$\int_{\Theta \times T_2^*} d\mu_1(\theta, t_2|t_1) \sum_{a_2 \in A_2} \left[\bar{g}_1(a_1, a_2, \theta) - \bar{g}_1((n, \ell), a_2, \theta)\right] \bar{\sigma}_2(\theta, t_2)[(a_2, 0)] \geqslant 0 \quad \forall n, \ell. \qquad (8)$$

By (7) and (8), a_1 is a γ-best reply to $\bar{\sigma}_2$ for t_1 in $\bar{G}$, and hence $a_1 \in R_1^r\left(t_1, \bar{G}, \gamma\right)$.

$\underline{R_2^r\left(t_2, \bar{G}, \gamma\right) \subset R_2^r\left(t_2, G, \gamma\right) \times \{0, 1, \ldots, N\}}$: Suppose $(a_2, m) \in R_2^r\left(t_2, \bar{G}, \gamma\right)$. Then, there is an $(r-1)$-order γ-rationalizable conjecture $\bar{\sigma}_1 : \Theta \times T_1^* \to \Delta\left(\bar{A}_1\right)$ in $\bar{G}$ such that for each $\left(a_2', m'\right) \in \bar{A}_2$,

$$\int_{\Theta\times T_1^*} d\mu_2(\theta, t_1|t_2) \sum_{\bar{a}_1\in\bar{A}_1} \left[\bar{g}_2\left((a_2, m), \bar{a}_1, \theta\right) - \bar{g}_2\left(\left(a_2', m'\right), \bar{a}_1, \theta\right)\right] \bar{\sigma}_1(\theta, t_1)[\bar{a}_1] \geqslant -\gamma.$$

Consider the map $\varphi_1 : \bar{A}_1 \to A_1$ such that

$$\varphi_1(\bar{a}_1) = \begin{cases} \bar{a}_1 & \text{if } \bar{a}_1 \in A_1, \\ a_1^c & \text{if } \bar{a}_1 \notin A_1. \end{cases}$$

Let σ_1 be the conjecture in G such that $\sigma_1(\theta, t_1)[a_1] = \bar{\sigma}_1(\theta, t_1)\left[\varphi_1^{-1}(a_1)\right]$ for each $(\theta, t_1) \in \Theta \times T_1^*$ and $a_1 \in A_1$. For each $\bar{a}_1 \in A_1$, since $\bar{\sigma}_1$ is $(r-1)$-order γ-rationalizable in $\bar{G}$, $\bar{\sigma}_1(\theta, t_1)[\bar{a}_1] > 0$ implies $\bar{a}_1 \in R_1^{r-1}(t_1, \bar{G}, \gamma)$, and by the induction hypothesis, $\bar{a}_1 \in R_1^{r-1}(t_1, G, \gamma)$. Moreover, a_1^c is part of a coordination pair in G, hence it is rationalizable for any type. Thus, σ_1 is an $(r-1)$-order γ-rationalizable conjecture in G. Moreover, (4) implies that for each $a_2' \in A_2$,

$$\begin{aligned}
-\gamma &\leqslant \int_{\Theta\times T_1^*} d\mu_2(\theta, t_1|t_2) \sum_{\bar{a}_1\in\bar{A}_1} \left[\bar{g}_2\left((a_2, m), \bar{a}_1, \theta\right) - \bar{g}_2\left(\left(a_2', m'\right), \bar{a}_1, \theta\right)\right] \bar{\sigma}_1(\theta, t_1)[\bar{a}_1] \\
&= \int_{\Theta\times T_1^*} d\mu_2(\theta, t_1|t_2) \Bigg(\sum_{a_1\in A_1} \left[g_2\left(a_2, a_1, \theta\right) - g_2\left(a_2', a_1, \theta\right)\right] \bar{\sigma}_1(\theta, t_1)[a_1] \\
&\qquad\qquad + \left[g_2\left(a_2, a_1^c, \theta\right) - g_2\left(a_2', a_1^c, \theta\right)\right] \bar{\sigma}_1(\theta, t_1)\left[\bar{A}_1 \backslash A_1\right] \Bigg) \\
&= \int_{\Theta\times T_1} d\mu_2(\theta, t_1|t_2) \sum_{a_1\in A_1} \left[g_2\left(a_2, a_1, \theta\right) - g_2\left(a_2', a_1, \theta\right)\right] \sigma_1(\theta, t_1)[a_1]
\end{aligned}$$

Therefore, $a_2 \in R_2^r(t_2, G, \gamma)$.

$\underline{R_2^r\left(t_2, \bar{G}, \gamma\right) \supset R_2^r\left(t_2, G, \gamma\right) \times \{0, 1, \ldots, N\}}$: Let $(a_2, m) \in R_2^r(t_2, G, \gamma) \times \{0, 1, \ldots, N\}$. Then, there is an $(r-1)$-order γ-rationalizable conjecture $\sigma_1 : \Theta \times T_1^* \to \Delta(A_1)$ in G such that for each $a_2' \in A_2$,

$$\int_{\Theta\times T_1^*} d\mu_2(\theta, t_1|t_2) \sum_{a_1\in A_1} \left[g_2\left(a_2, a_1, \theta\right) - g_2\left(a_2', a_1, \theta\right)\right] \sigma_1(\theta, t_1)[a_1] \geqslant -\gamma. \tag{9}$$

Let $\bar{\sigma}_1$ be the conjecture in $\bar{G}$ such that $\bar{\sigma}_1(\theta, t_1)[a_1] = \sigma_1(\theta, t_1)[a_1]$ for each $(\theta, t_1) \in \Theta \times T_1^*$ and $a_1 \in A_1$ (and thus $\bar{\sigma}_1(\theta, t_1)[a_1] = 0$ for any $a_1 \notin A_1$). Since σ_1 is $(r-1)$-order γ-rationalizable in G, $\sigma_1(\theta, t_1)[a_1] > 0$ implies $a_1 \in R_1^{r-1}(t_1, G, \gamma)$, and by the induction hypothesis, $a_1 \in R_1^{r-1}\left(t_1, \bar{G}, \gamma\right)$. Hence, $\bar{\sigma}_1$ is $(r-1)$-order γ-rationalizable in $\bar{G}$. Since $\bar{g}_2((a_2, m), a_1, \theta) = \bar{g}_2((a_2, m'), a_1, \theta)$ for each m and m', it follows from (9) that $(a_2, m) \in R_2^r\left(t_2, \bar{G}, \gamma\right)$. ■

Corollary 3. *For every $\varepsilon > 0$, player i, positive integer k and finite type $t_i \in T_i^*$, there exists a game G and an action a_i of player i in G, such that:*

(i) $R_j(\cdot, G, \gamma) = R_j^k(\cdot, G, \gamma)$ *for every* $\gamma \in [0, M/10)$ *and* $j \in I$*;*

(ii) $a_i \in R_i(t_i, G)$*;*

(iii) $a_i \notin R_i^k(s_i, G, M\varepsilon/10)$ *for every* $s_i \in T_i^*$ *with* $d_i^k(s_i, t_i) > \varepsilon$.

Proof. Immediate implication of Lemma 1, upon rescaling the payoffs by a factor of $M/5$. ■

Lemma 2. *For every $\varepsilon > 0$ there exists $\delta > 0$ such that for every $i \in I$ and $t_i, t_i' \in T_i^*$ with $d_i^{UW}(t_i, t_i') > \varepsilon$ there is a game G such that $R_i(t_i, G) \not\subseteq R_i(t_i', G, \delta)$.*

Proof. Fix an $\varepsilon > 0$, a player i, an integer $k \geqslant 1$ and types $t_i, t_i' \in T_i^*$ with $d_i^k(t_i, t_i') > \varepsilon$. Fix $0 < \delta < M\varepsilon/10$ and choose $\rho > 0$ small enough that

$$\frac{M(\varepsilon - \rho)}{10} - 4M\rho > \delta. \tag{10}$$

Since finite types are dense in the product topology, there is a finite type t_i'' such that $d_i^k(t_i, t_i'') < \rho$. Then, $d_i^k(t_i'', t_i') > \varepsilon - \rho$. By Corollary 3, there is a game $G' = (A_i', g_i')_{i \in I}$ and some action a_i' of player i in G', such that $a_i' \in R_i\left(t_i'', G', 0\right)$ and $a_i' \notin R_i^k(t_i', G', M(\varepsilon - \rho)/10)$. By Proposition 2 in Chen, Di Tillio, Faingold, and Xiong (2010), $a_i' \in R_i^k(t_i, G', 4M\rho)$. Then, it follows from (i) of Corollary 3 that $a_i' \in R_i(t_i, G', 4M\rho)$ and $a_i' \notin R_i(t_i', G', M(\varepsilon - \rho)/10)$.

To conclude, consider the game $G = (A_j, g_j)_{j \in I}$, defined as follows:

$$A_i = A_i', \quad A_{-i} = A_{-i}' \times A_i',$$

$$g_i(a_i, (a_{-i}^{-i}, a_{-i}^{i}), \theta) = \begin{cases} g_i'(a_i, a_{-i}^{-i}, \theta) + 4M\rho & : \quad a_i^i = a_{-i}^i \\ g_i'(a_i, a_{-i}^{-i}, \theta) & : \quad \text{otherwise,} \end{cases}$$

and

$$g_{-i}((a_{-i}^{-i}, a_{-i}^{i}), a_i, \theta) = g_{-i}'(a_{-i}^{-i}, a_i, \theta).$$

In game G, player $-i$ is indifferent among all actions a_{-i}^i; moreover, player i gets an additional payoff of $4M\rho$ whenever his action matches player $-i$'s choice of a_{-i}^i. Therefore, $R_i(\cdot, G, \gamma) = R_i(\cdot, G', \gamma + 4M\rho)$ for every $\gamma \geqslant 0$. In particular, we have that $a_i' \in R_i(t_i, G, 0)$ and $a_i' \notin R_i(t_i', G, M(\varepsilon - \rho)/10 - 4M\rho) \supseteq R_i(t_i', G, \delta)$, where the inclusion follows from (10).[44] ■

The only if direction of Theorem 1 then follows directly from Lemma 2.

[44] If necessary, rescale the payoffs from G to ensure $|g_j| \leqslant M$ for every $j \in I$, and rescale δ by the same factor.

A.3 Strategic frames

For each player i and $\varnothing \neq A'_i \subseteq A_i$, let

$$[A'_i] = \{t_i : R_i(t_i, G) = A'_i\}.$$

Fix any two types $t_i, t'_i \in T^*_i$ and suppose that $\mu^*_i(\theta \times [A'_{-i}]|t_i) = \mu^*_i(\theta \times [A'_{-i}]|t'_i)$ for each $\theta \in \Theta$ and $\varnothing \neq A'_i \subseteq A_i$. We must show that $R_i(t_i, G) = R_i(t'_i, G)$. If $a_i \in R_i(t_i, G)$ then, by the characterization of rationalizability in terms of best-reply sets (Appendix A.1), there is a conjecture $\nu \in \Delta(\Theta \times T^*_{-i} \times A_{-i})$ satisfying $\mathrm{marg}_{\Theta \times T^*_{-i}} \nu = \mu^*_i(t_i)$, $\nu(\mathrm{graph}\, R_{-i}(\cdot, G)) = 1$ and $a_i \in BR_i(\mathrm{marg}_{A_{-i} \times \Theta} \nu, G)$. Then, define a conjecture $\nu' \in \Delta(\Theta \times T^*_{-i} \times A_{-i})$ for type t'_i as follows: for each $\theta \in \Theta$, $a_{-i} \in A_{-i}$ and measurable $E \subseteq T^*_{-i}$,

$$\nu'(\theta \times E \times a_{-i}) = \sum_{A'_{-i}} \frac{\nu(\theta \times [A'_{-i}] \times a_{-i})}{\nu(\theta \times [A'_{-i}] \times A_{-i})} \mu^*_i(\theta \times (E \cap [A'_{-i}])|t'_i),$$

where the summation ranges over all $A'_{-i} \subseteq A_{-i}$ such that $\mu^*_i(\theta \times (E \cap [A'_{-i}])|t'_i) > 0$. Note that ν' is well defined, because $\mu^*_i(\theta \times (E \cap [A'_{-i}])|t'_i) > 0$ implies $\nu(\theta \times [A'_{-i}] \times A_{-i}) > 0$, since

$$\nu(\theta \times [A'_{-i}] \times A_{-i}) = \mu^*_i(\theta \times [A'_{-i}]|t_i) = \mu^*_i(\theta \times [A'_{-i}]|t'_i). \tag{11}$$

By construction, we have $\mathrm{marg}_{\Theta \times T^*_{-i}} \nu' = \mu^*_i(t'_i)$. Also, the condition $\nu(\mathrm{graph}\, R_{-i}(\cdot, G)) = 1$ implies that for every $\theta \in \Theta$, $a_{-i} \in A_{-i}$ and $A'_{-i} \subseteq A_{-i}$,

$$\nu'(\theta \times [A'_{-i}] \times a_{-i}) > 0 \implies \nu(\theta \times [A'_{-i}] \times a_{-i}) > 0 \implies a_{-i} \in A'_{-i}.$$

Hence, $\nu'(\mathrm{graph}\, R_{-i}(\cdot, G)) = 1$. Finally, (11) above implies $\mathrm{marg}_{A_{-i} \times \Theta} \nu' = \mathrm{marg}_{A_{-i} \times \Theta} \nu$. Thus, $a_i \in BR_i(\mathrm{marg}_{A_{-i} \times \Theta} \nu', G)$, and therefore $a_i \in R_i(t'_i, G)$. We have thus shown that $R_i(t_i, G) \subseteq R_i(t'_i, G,)$, and the opposite inclusion follows by interchanging the roles of t_i and t'_i in the argument above. This proves that the profile of partitions is indeed a frame.

A.4 Proof of Theorem 2

We begin with the following auxiliary result about the structure of ε-rationalizability.

Lemma 3. *Fix a game $G = (A_i, g_i)_{i \in I}$ and $\varepsilon \geqslant 0$. For every integer $k \geqslant 1$, $i \in I$ and $t_i \in T^*_i$, we have $a_i \in R^k_i(t_i, G, \varepsilon)$ if and only if, for every $\alpha_i \in \Delta(A_i)$,*

$$\sum_{\theta \in \Theta, B \subseteq A_{-i}} \max_{a_{-i} \in B} \left[g_i(a_i, a_{-i}, \theta) - g_i(\alpha_i, a_{-i}, \theta) \right] \times \mu_i\left(\theta \times \left\{t_{-i} : R^{k-1}_{-i}(t_{-i}, G, \varepsilon) = B\right\} \Big| t_i\right) \geqslant -\varepsilon. \tag{12}$$

Likewise, $a_i \in R_i(t_i, G, \varepsilon)$ *if and only if, for every* $\alpha_i \in \Delta(A_i)$*,*

$$\sum_{\theta\in\Theta, B\subseteq A_{-i}} \max_{a_{-i}\in B} \Big[g_i(a_i, a_{-i}, \theta) - g_i(\alpha_i, a_{-i}, \theta) \Big] \\ \times\ \mu_i\Big(\theta \times \Big\{t_{-i} : R_{-i}(t_{-i}, G, \varepsilon) = B\Big\}\Big| t_i\Big) \geqslant -\varepsilon. \tag{13}$$

Proof. The first part is equivalent to the iterative definition of interim correlated rationalizability in terms of dominance, given by condition (3) in Appendix A.1. The second part follows directly from the first. ■

The if direction of the theorem is an immediate consequence of the following lemma:

Lemma 4. *Fix* $\delta > 0$ *and a game* $G = (A_i, g_i)_{i\in I}$ *and let* $\mathcal{P}$ *denote the strategic frame associated with* G*. For every integer* $k \geqslant 0$*,* $i \in I$ *and* $t_i, t_i' \in T_i^*$*,*

$$d^k_{i,\mathcal{P}}(t_i, t_i') < \delta \quad \Longrightarrow \quad R_i(t_i, G) \subseteq R^k_i(t_i', G, 4M\delta).$$

In particular, for every $i \in I$ *and* $t_i, t_i' \in T_i^*$*,*

$$d^{UW}_{i,\mathcal{P}}(t_i, t_i') < \delta \quad \Longrightarrow \quad R_i(t_i, G) \subseteq R_i(t_i', G, 4M\delta).$$

Proof. We need only prove the first result, as the second result is a straightforward implication of the first one. For $k = 0$ the result is trivially true, as $R_i^0 \equiv A_i$. Proceeding by induction, we assume the result is true for $k \geqslant 0$ and show that it remains true for $k+1$. Consider two types t_i, t_i' with $d^{k+1}_{i,\mathcal{P}}(t_i, t_i') < \delta$. Fix an arbitrary $a_i \in R_i(t_i, G)$ and let us show that $a_i \in R_i^{k+1}(t_i', G, 4M\delta)$. Given $\alpha_i \in \Delta(A_i)$, by Lemma 3 we have

$$\sum_{\theta\in\Theta, B\subseteq A_{-i}} \Delta g_i\big(\theta, B\big)\, \mu_i\big(\theta \times [B] \big| t_i\big) \geqslant 0, \tag{14}$$

where, for each $\theta \in \Theta$ and nonempty $B \subseteq A_{-i}$,

$$[B] = \big\{t_{-i} : R_{-i}(t_{-i}, G) = B\big\},$$
$$\Delta g_i(\theta, B) = \max_{a_{-i}\in B} \Big[g_i(a_i, a_{-i}, \theta) - g_i(\alpha_i, a_{-i}, \theta) \Big].$$

By Lemma 3, in order to prove that $a_i \in R_i^{k+1}(t_i', G, 4M\delta)$ we need only show that

$$\sum_{\theta\in\Theta, B\subseteq A_{-i}} \Delta g_i\big(\theta, B\big)\, \mu_i\big(\theta \times [[B]] \big| t_i'\big) \geqslant -4M\delta,$$

where

$$[[B]] = \big\{t_{-i} : R^k_{-i}(t_{-i}, G, 4M\delta) = B\big\} \qquad \forall B \subseteq A_{-i}.$$

To prove this, first note that the induction hypothesis implies

$$[B]^{\delta,k}_{\mathcal{P}} \subseteq \bigcup_{C \supseteq B} [[C]] \qquad \forall B \subseteq A_{-i}. \tag{15}$$

Second, enumerate the elements of the finite set $\Theta \times \{B : \varnothing \neq B \subseteq A_{-i}\}$ as $\{(\theta_n, B_n)\}_{n=1}^{N}$ ($N = |\Theta|(2^{|A_{-i}|} - 1)$) so that

$$\Delta g_i(\theta_n, B_n) \geqslant \Delta g_i(\theta_{n+1}, B_{n+1}) \qquad \forall n = 1, \ldots, N-1 \tag{16}$$

and

$$(\theta_m = \theta_n \quad \text{and} \quad B_m \supseteq B_n) \quad \Longrightarrow \quad m \leqslant n.^{45} \tag{17}$$

Then,

$$\begin{aligned}
&\sum_{\theta \in \Theta, B \subseteq A_{-i}} \Delta g_i(\theta, B)\, \mu_i\big(\theta \times [[B]] \big| t_i'\big) \\
&\geqslant \sum_{n=1}^{N} \Delta g_i(\theta_n, B_n) \Big(\mu_i\big(\theta_n \times [[B_n]] \big| t_i'\big) - \mu_i\big(\theta_n \times [B_n] \big| t_i\big)\Big) && (18)\\
&= \sum_{n=1}^{N-1} \Big(\Delta g_i(\theta_n, B_n) - \Delta g_i(\theta_{n+1}, B_{n+1})\Big) \\
&\qquad \times \sum_{m=1}^{n} \big(\mu_i\big(\theta_m \times [[B_m]] \big| t_i'\big) - \mu_i\big(\theta_m \times [B_m] \big| t_i\big)\big) && (19)\\
&= \sum_{n=1}^{N-1} \Big(\Delta g_i(\theta_n, B_n) - \Delta g_i(\theta_{n+1}, B_{n+1})\Big) \\
&\qquad \times \Bigg(\mu_i\Big(\bigcup_{m=1}^{n} \theta_m \times [[B_m]] \Big| t_i'\Big) - \mu_i\Big(\bigcup_{m=1}^{n} \theta_m \times [B_m] \Big| t_i\Big)\Bigg) && (20)\\
&\geqslant \sum_{n=1}^{N-1} \Big(\Delta g_i(\theta_n, B_n) - \Delta g_i(\theta_{n+1}, B_{n+1})\Big) \\
&\qquad \times \Bigg(\mu_i\Big(\Big(\bigcup_{m=1}^{n} \theta_m \times [B_m]\Big)^{\delta,k}_{\mathcal{P}} \Big| t_i'\Big) - \mu_i\Big(\bigcup_{m=1}^{n} \theta_m \times [B_m] \Big| t_i\Big)\Bigg) && (21)\\
&\geqslant \sum_{n=1}^{N-1} \big(\Delta g_i(\theta_n, B_n) - \Delta g_i(\theta_{n+1}, B_{n+1})\big)\big(-\delta\big) && (22)\\
&= -\big(\Delta g_i(\theta_1, B_1) - \Delta g_i(\theta_N, B_N)\big)\delta \\
&\geqslant -4M\delta, && (23)
\end{aligned}$$

[45] Such enumeration is possible because $\Delta g_i(\theta, B') \geqslant \Delta g_i(\theta, B)$ whenever $B' \supseteq B$.

where (18) follows from (14); (19) follows by a standard "integration by parts" argument; (20) follows from the fact that $\{(\theta_m, [B_m])\}_{m=1}^N$ and $\{(\theta_m, [[B_m]])\}_{m=1}^N$ are partitions of $\Theta \times T^*_{-i}$; (21) follows from (15), (16) and (17); (22) follows from the assumption that $d^{k+1}_{i,\mathcal{P}}(t_i, t'_i) \leqslant \delta$, the fact that $[B_m] \in \mathcal{P}_{-i}$ and (16); and (23) follows from $|\Delta g_i| \leqslant 2M$. ■

Before turning to the proof of the only if direction of Theorem 2, we introduce some notation. First, given a frame $\mathcal{P}$ and a type t_i, let $\mu^*_i(t_i)\big|_{\Theta\times\mathcal{P}_{-i}}$ denote the belief of type t_i over the events in $\Theta\times\mathcal{P}_{-i}$. We thus view $\mu^*_i(t_i)\big|_{\Theta\times\mathcal{P}_{-i}}$ as an element of the finite-dimensional simplex $\Delta(\Theta\times\mathcal{P}_{-i})$ (viewed as a subset of the Euclidean space $\mathbb{R}^{|\Theta|\cdot|\mathcal{P}_{-i}|}$). Second, let $\varphi_i : \Delta(\Theta \times \mathcal{P}_{-i}) \to \mathcal{P}_i$ designate the function that maps each $q \in \Delta(\Theta \times \mathcal{P}_{-i})$ into $\mathcal{P}_i(t_i)$, where t_i is some type with $\mu^*_i(t_i)\big|_{\Theta\times\mathcal{P}_{-i}} = q$. Since $\mathcal{P}$ is a frame, the definition of φ_i is independent of the choice of t_i. Third, given a game $G = (A_i, g_i)_{i\in I}$, for each $\beta_i \in \Delta(A_i \times \Theta)$ define a probability distribution $\beta^{\mathcal{P}}_i \in \Delta(\Theta \times \mathcal{P}_i)$ as follows:

$$\beta^{\mathcal{P}}_i(\theta, E) = \sum_{a_i\in\varphi_i^{-1}(E)} \beta_i(a_i, \theta) \qquad \forall(\theta, E) \in \Theta \times \mathcal{P}_i.$$

The proof of the only if direction of Theorem 2 relies on Lemmas 5 and 6 and Corollary 4 below.

Lemma 5. *Fix a frame $\mathcal{P}$. For each $\zeta > 0$ there exist $\varepsilon > 0$ and a game $G = (A_i, g_i)_{i\in I}$ such that for each $i \in I$,*

(i) A_i is a finite subset of $\Delta(\Theta \times \mathcal{P}_{-i})$ that is $\sqrt{\varepsilon}$-dense (relative to the Euclidean norm $\|\cdot\|$) in every element of the partition $\{\varphi_i^{-1}(E) : E \in \mathcal{P}_i\}$;

(ii) for each $a_{-i}, a'_{-i} \in A_{-i}$,

$$\varphi_{-i}(a_{-i}) = \varphi_{-i}(a'_{-i}) \implies g_i(a_i, a_{-i}, \theta) = g_i(a_i, a'_{-i}, \theta) \quad \forall\theta \in \Theta, a_i \in A_i;$$

(iii) for each $\beta_{-i} \in \Delta(A_{-i} \times \Theta)$,

$$BR_i(\beta_{-i}, G, \varepsilon) \supseteq \{a_i \in A_i : \|a_i - \beta^{\mathcal{P}}_{-i}\| < \sqrt{\varepsilon}\};$$

(iv) for each $\beta_{-i}, \beta'_{-i} \in \Delta(A_{-i} \times \Theta)$ with $\|\beta^{\mathcal{P}}_{-i} - \beta'^{\mathcal{P}}_{-i}\| > \zeta$,

$$BR_i(\beta'_{-i}, G, 2\varepsilon) \cap \{a_i \in A_i : \|a_i - \beta^{\mathcal{P}}_{-i}\| < \sqrt{\varepsilon}\} = \varnothing.$$

Proof. Fix $\zeta > 0$ and a frame $\mathcal{P}$. Let $0 < \varepsilon < \zeta^2/(1 + \sqrt{3})^2$. Cover the finite-dimensional simplex $\Delta(\Theta\times\mathcal{P}_{-i})$ by a finite union of open balls $B_1, \ldots, B_N$ of diameter $\sqrt{\varepsilon}$. Select one point from $B_n \cap \varphi_i^{-1}(E)$, for each $n = 1, \ldots, N$ and $E \in \mathcal{P}_i$, and let A_i denote the set of selected points. By construction, A_i satisfies (i).

Consider the quadratic score $s_i : \Delta(\Theta \times \mathcal{P}_{-i}) \times \Theta \times \mathcal{P}_{-i} \to \mathbb{R}$,

$$s_i(q, \theta, E) = 2q(\theta, E) - \|q\|^2 \qquad \forall (q, \theta, E) \in \Delta(\Theta \times \mathcal{P}_{-i}) \times \Theta \times \mathcal{P}_{-i},$$

which can be readily shown to satisfy

$$s_i(q, q) - s_i(q', q) = \|q' - q\|^2 \qquad q, q' \in \Delta(\Theta \times \mathcal{P}_{-i}). \tag{24}$$

Then, define $g_i : A_i \times A_{-i} \times \Theta \to \mathbb{R}$,

$$g_i(a_i, a_{-i}, \theta) = s_i(a_i, \theta, \varphi_{-i}(a_{-i})) \qquad \forall (a_i, a_{-i}, \theta) \in A_i \times A_{-i} \times \Theta,$$

which clearly satisfies (ii).

Given $\beta_{-i} \in \Delta(A_{-i} \times \Theta)$ and $a_i \in A_i$ with $\|a_i - \beta^{\mathcal{P}}_{-i}\| < \sqrt{\varepsilon}$, for any $a'_i \in A_i$ we have

$$\begin{aligned}
g_i(a_i, &\beta_{-i}) - g_i(a'_i, \beta_{-i}) \\
&= s_i(a_i, \beta^{\mathcal{P}}_{-i}) - s_i(a'_i, \beta^{\mathcal{P}}_{-i}) && \text{(by (ii))} \\
&= s_i(a_i, \beta^{\mathcal{P}}_{-i}) - s_i(\beta^{\mathcal{P}}_{-i}, \beta^{\mathcal{P}}_{-i}) + s_i(\beta^{\mathcal{P}}_{-i}, \beta^{\mathcal{P}}_{-i}) - s_i(a'_i, \beta^{\mathcal{P}}_{-i}) \\
&= -\|a_i - \beta^{\mathcal{P}}_{-i}\|^2 + \|a'_i - \beta^{\mathcal{P}}_{-i}\|^2 && \text{(by (24))} \\
&> -\varepsilon, && \text{(by } \|a_i - \beta^{\mathcal{P}}_{-i}\| < \sqrt{\varepsilon}\text{)}
\end{aligned}$$

hence $a_i \in BR_i(\beta_{-i}, G, \varepsilon)$, and this proves (iii).

Turning to (iv), let β_{-i}, $\beta'_{-i} \in \Delta(A_{-i} \times \Theta)$ with $\|\beta'^{\mathcal{P}}_{-i} - \beta^{\mathcal{P}}_{-i}\| > \zeta$ and $a_i \in A_i$ with $\|a_i - \beta^{\mathcal{P}}_{-i}\| < \sqrt{\varepsilon}$. Then, $\|a_i - \beta'^{\mathcal{P}}_{-i}\| > \zeta - \sqrt{\varepsilon} > \sqrt{3\varepsilon}$. By (i) we can find some $a'_i \in A_i$ with $\|a'_i - \beta'^{\mathcal{P}}_{-i}\| < \sqrt{\varepsilon}$. Thus,

$$\begin{aligned}
g_i(a_i, &\beta'_{-i}) - g_i(a'_i, \beta'_{-i}) \\
&= s_i(a_i, \beta'^{\mathcal{P}}_{-i}) - s_i(a'_i, \beta'^{\mathcal{P}}_{-i}) && \text{(by (ii))} \\
&= s_i(a_i, \beta'^{\mathcal{P}}_{-i}) - s_i(\beta'^{\mathcal{P}}_{-i}, \beta'^{\mathcal{P}}_{-i}) + s_i(\beta'^{\mathcal{P}}_{-i}, \beta'^{\mathcal{P}}_{-i}) - s_i(a'_i, \beta'^{\mathcal{P}}_{-i}) \\
&= -\|a_i - \beta'^{\mathcal{P}}_{-i}\|^2 + \|a'_i - \beta'^{\mathcal{P}}_{-i}\|^2 && \text{(by (24))} \\
&< -3\varepsilon + \varepsilon = -2\varepsilon, && \big(\text{by } \|a_i - \beta'^{\mathcal{P}}_{-i}\| > \sqrt{3\varepsilon} \\
& && \text{and } \|a'_i - \beta'^{\mathcal{P}}_{-i}\| < \sqrt{\varepsilon}\big)
\end{aligned}$$

and hence $a_i \notin BR_i(\beta'_{-i}, G, 2\varepsilon)$, as required. ■

Lemma 6. *Fix $\delta > 0$ and a frame $\mathcal{P}$. Let $\varepsilon > 0$ and $G = (A_i, g_i)_{i \in I}$ satisfy the properties (i)–(iv) of Lemma 5 relative to $\zeta = \delta / (|\Theta| \cdot \max_i |\mathcal{P}_i|)$. Then the following statements hold:*

(a) for every $i \in I$ and $t_i \in T^*_i$,

$$R(t_i, G, \varepsilon) \supseteq \left\{ a_i \in A_i \cap \varphi_i^{-1}\big(\mathcal{P}_i(t_i)\big) : \left\| a_i - \mu^*_i(t_i)\big|_{\Theta \times \mathcal{P}_{-i}} \right\| < \sqrt{\varepsilon} \right\};$$

(b) for every integer $k \geqslant 0$, $i \in I$ and $t_i, t_i' \in T_i^*$ with $d_{i,\mathcal{P}}^k(t_i, t_i') > \delta$,

$$R_i^k(t_i', G, 2\varepsilon) \cap \Big\{a_i \in A_i \cap \varphi_i^{-1}\big(\mathcal{P}_i(t_i)\big) : \big\|a_i - \mu_i^*(t_i)\big|_{\Theta\times\mathcal{P}_{-i}}\big\| < \sqrt{\varepsilon}\Big\} = \varnothing.$$

Proof. To prove (a), we show that the pair of correspondences $\varsigma_i : T_i^* \rightrightarrows A_i$, $i \in I$, defined by

$$\varsigma_i(t_i) = \Big\{a_i \in A_i \cap \varphi_i^{-1}\big(\mathcal{P}_i(t_i)\big) : \big\|a_i - \mu_i^*(t_i)\big|_{\Theta\times\mathcal{P}_{-i}}\big\| < \sqrt{\varepsilon}\Big\} \qquad \forall t_i \in T_i^*,$$

which is nonempty-valued by (i) of Lemma 5, has the ε-best-reply property. Indeed, given any $a_i \in \varsigma_i(t_i)$ and any conjecture $\sigma_{-i} : \Theta \times T_{-i}^* \to \Delta(A_{-i})$ with $\operatorname{supp}\sigma_{-i}(\theta, t_{-i}) \subseteq \varsigma_{-i}(t_{-i})$ for all $(\theta, t_{-i}) \in \Theta \times T_{-i}^*$, we must have

$$\beta_{t_i,\sigma_{-i}}\Big(\theta \times \big(A_{-i} \cap \varphi_{-i}^{-1}(E)\big)\Big) = \mu_i^*\big(\theta \times E|t_i\big) \qquad \forall \theta \in \Theta, E \in \mathcal{P}_{-i},$$

and hence, $\|a_i - \beta_{t_i,\sigma_{-i}}^{\mathcal{P}}\| = \big\|a_i - \mu_i^*(t_i)\big|_{\Theta\times\mathcal{P}_{-i}}\big\| < \sqrt{\varepsilon}$, by (i) of Lemma 5. It then follows by (iii) of Lemma 5 that $a_i \in BR_i(\beta_{t_i,\sigma_{-i}}, G, \varepsilon)$. We have thus shown that the profile of correspondences $(\varsigma_i)_{i\in I}$ has the ε-best-reply property.

Turning to (b), since $d_{i,\mathcal{P}}^0 \equiv 0$ the result is true for $k = 0$ (vacuously). Proceeding by induction, assume the result is true for $k \geqslant 0$ and let us show that it remains true for $k+1$. Fix $i \in I$ and $t_i, t_i' \in T_i^*$ with $d_{i,\mathcal{P}}^{k+1}(t_i, t_i') > \delta$. Then, there exist $\theta \in \Theta$ and $E \in \mathcal{P}_{-i}$ such that

$$\mu_i^*\big(\theta \times E_{\mathcal{P}}^{\delta,k}\big|t_i'\big) < \mu_i^*\big(\theta \times E\big|t_i\big) - \delta/(|\Theta| \cdot |\mathcal{P}_{-i}|). \tag{25}$$

Consider an arbitrary k-order 2ε-rationalizable conjecture $\sigma_{-i} : \Theta \times T_{-i}^* \to \Delta(A_{-i})$. By the induction hypothesis, for every $t_{-i}, t_{-i}' \in T_{-i}^*$, every $\theta \in \Theta$ and every $a_{-i} \in A_{-i}$ with $\big\|a_{-i} - \mu_{-i}^*(t_{-i})\big|_{\Theta\times\mathcal{P}_i}\big\| < \sqrt{\varepsilon}$, we can have $\sigma_{-i}(\theta, t_{-i}')[a_{-i}] > 0$ only if $d_{-i,\mathcal{P}}^k(t_{-i}, t_{-i}') \leqslant \delta$. In particular, for every $t_{-i}, t_{-i}' \in T_{-i}^*$, $\theta \in \Theta$ and $a_{-i} \in A_{-i}$,

$$\Big(\ \mu_{-i}^*(t_{-i})\big|_{\Theta\times\mathcal{P}_i} = a_{-i} \quad \text{and} \quad d_{-i,\mathcal{P}}^k(t_{-i}, t_{-i}') > \delta\ \Big) \quad \Longrightarrow \quad \sigma_{-i}(\theta, t_{-i}')[a_{-i}] = 0.$$

It follows that for every $(\theta, t_{-i}') \in \Theta \times T_{-i}^*$,

$$t_{-i}' \notin E_{\mathcal{P}}^{\delta,k} \quad \Longrightarrow \quad \sigma_{-i}(\theta, t_{-i}')\big[A_{-i} \cap \varphi_{-i}^{-1}(E)\big] = 0, \tag{26}$$

since for every $a_{-i} \in A_{-i} \cap \varphi_{-i}^{-1}(E)$ and every $t_{-i} \in T_{-i}^*$ with $\mu_{-i}^*(t_{-i})\big|_{\Theta\times\mathcal{P}_i} = a_{-i}$ we must have $t_{-i} \in E$. Hence, for every k-order 2ε-rationalizable conjecture σ_{-i} we have

$$\begin{aligned}
&\beta_{t_i',\sigma_{-i}}\Big(\theta \times \big(A_{-i} \cap \varphi_{-i}^{-1}(E)\big)\Big) \\
&\quad= \int_{T_{-i}^*} \sigma_{-i}(\theta, t_{-i}')\big[A_{-i} \cap \varphi_{-i}^{-1}(E)\big]\, \mu_i^*\big(\theta \times dt_{-i}'\big|t_i'\big) \\
&\quad= \int_{E_{\mathcal{P}}^{\delta,k}} \sigma_{-i}(\theta, t_{-i}')\big[A_{-i} \cap \varphi_{-i}^{-1}(E)\big]\, \mu_i^*\big(\theta \times dt_{-i}'\big|t_i'\big) \qquad \text{(by (26))} \\
&\quad\leqslant \mu_i^*\big(\theta \times E_{\mathcal{P}}^{\delta,k}\big|t_i'\big) \\
&\quad< \mu_i^*(\theta \times E|t_i) - \delta/(|\Theta| \cdot |\mathcal{P}_{-i}|), \qquad \text{(by (25))}
\end{aligned}$$

and thus, $\left\|\beta^{\mathcal{P}}_{t'_i,\sigma_{-i}} - \mu^*_i(t_i)\big|_{\Theta\times\mathcal{P}_{-i}}\right\| > \zeta$. It follows from part (iv) of Lemma 5 that

$$R^{k+1}_i(t'_i, G, 2\varepsilon) \cap \left\{a_i : \left\|a_i - \mu^*_i(t_i)\big|_{\Theta\times\mathcal{P}_{-i}}\right\| < \sqrt{\varepsilon}\right\} = \varnothing,$$

as was to be shown. ■

Corollary 4. *For every frame $\mathcal{P}$ and $\delta > 0$ there exists $\varepsilon > 0$ and a game G such that, for every $i \in I$ and $t_i, t'_i \in T^*_i$, if $d^{UW}_{i,\mathcal{P}}(t_i, t'_i) > \delta$ then $R_i(t_i, G) \not\subseteq R_i(t'_i, G, \varepsilon)$.*

Proof. A straightforward implication of Lemma 6 is that there exist $\varepsilon > 0$ and a game G' such that, for every $i \in I$ and $t_i, t'_i \in T^*_i$, if $d^{UW}_{i,\mathcal{P}}(t_i, t'_i) > \delta$ then $R_i(t_i, G', \varepsilon) \not\subseteq R_i(t'_i, G', 2\varepsilon)$. To conclude, we can use a construction similar to the last part of the proof of Lemma 2 to obtain a game G with $R_i(\cdot, G, \gamma) = R_i(\cdot, G', \gamma + \varepsilon)$ for all $\gamma \geqslant 0$. ■

The only if direction of Theorem 2 is an immediate implication of Corollary 4.

A.5 Proof of Corollary 1

We begin with the following definition. Given a frame $\mathcal{P}$ and a finite measurable partition Π_i of the finite-dimensional simplex $\Delta(\Theta \times \mathcal{P}_{-i})$, define the *partition on T^*_i induced by* Π_i, written T^*_i/Π_i, as follows: any two types of player i belong to the same element of T^*_i/Π_i if and only if their beliefs over $\Theta \times \mathcal{P}_{-i}$ belong to the same element of Π_i. The following lemma is straightforward from the definitions, and the proof is ommitted:

Lemma 7. *Each T^*_i/Π_i is a measurable partition of T^*_i, and the join $(\mathcal{P}_i \vee (T^*_i/\Pi_i))_{i\in I}$ is a frame.*[46]

We now need the following piece of notation. Given an integer $m \geqslant 1$, a measurable subset $E \subseteq \Omega$ and $\mathbf{p} \in [0,1]^2$, define the *event that E is m-order* $\mathbf{p}$*-belief* recursively as follows:

$$\left[B^{\mathbf{p}}\right]^m(E) \stackrel{\text{def}}{=} B^{\mathbf{p}}\Big(E \cap \left[B^{\mathbf{p}}\right]^{m-1}(E)\Big),$$

where $\left[B^{\mathbf{p}}\right]^0(E) \stackrel{\text{def}}{=} \Omega$. Then, the *event that E is m-order* $\mathbf{p}$*-belief for player i*, written $\left[B^{\mathbf{p}}_i\right]^m(E)$, is the projection of $\left[B^{\mathbf{p}}\right]^m(E)$ onto T^*_i. We these definitions in place, we have:

$$C^{\mathbf{p}}(E) = \bigcap_{m\geqslant 1} \left[B^{\mathbf{p}}\right]^m(E) \quad \text{and} \quad C^{\mathbf{p}}_i(E) = \bigcap_{m\geqslant 1} \left[B^{\mathbf{p}}_i\right]^m(E).$$

Lemma 8. *For every integer $k \geqslant 1$ and $\delta > 0$ there exists a k-order frame $\mathcal{P}$ such that, for every $i \in I$, every atom of $\mathcal{P}_i$ has d^k_i-diameter at most δ.*

[46] Recall that the join of a pair of partitions, denoted by the symbol $\vee$, is the coarsest partition that is finer than both partitions in the pair.

Proof. For $k = 1$ the result is trivial, as any profile of first-order measurable partitions is a frame. Proceeding by induction, consider $k \geqslant 1$, fix $\delta > 0$ and let $\mathcal{P}$ be a k-order measurable frame whose atoms all have d_i^k-diameter less than $\delta/2$. Let Π_i be a finite partition of the simplex $\Delta(\Theta \times \mathcal{P}_{-i})$ (viewed as a subset of the Euclidean space $\mathbb{R}^{|\Theta|\cdot|\mathcal{P}_{-i}|}$) into finitely many Borel measurable subsets with Euclidean diameter less than $\delta/\sqrt{|\Theta| \cdot |\mathcal{P}_{-i}|}$. By Lemma 7, the join $(\mathcal{P}_i \vee (\Pi_i/T_i^*))_{i\in I}$ is a $(k+1)$-order frame. We claim that every atom of $\mathcal{P}_i \vee (\Pi_i/T_i^*)$ has d_i^{k+1}-diameter less than δ. Let t_i and t_i' be two types with $t_i' \in \big(T_i^*/\Pi_i\big)(t_i)$, let E be a measurable subset of $\Theta \times T_{-i}^*$ and let us show that $\mu_i^*(E^{\delta,k}|t_i') \geqslant \mu_i^*(E|t_i) - \delta$. Since each atom of $\mathcal{P}_{-i}$ has d_i^k-diameter less than $\delta/2$, there is some $F \in 2^\Theta \otimes \mathcal{P}_{-i}$ with $E \subseteq F \subseteq E^{\delta/2,k}$. Then, $|\mu_i^*(F|t_i') - \mu_i^*(F|t_i)| < \delta$, because each atom of Π_i has Euclidean diameter less than $\delta/\sqrt{|\Theta| \cdot |\mathcal{P}_{-i}|}$ and we have $t_i' \in \big(T_i^*/\Pi_i\big)(t_i)$. Thus,

$$
\begin{aligned}
\mu_i^*(E^{\delta,k}|t_i') &\geqslant \mu_i^*(F^{\delta/2,k}|t_i') && (\text{by } F \subseteq E^{\delta/2,k}) \\
&\geqslant \mu_i^*(F|t_i') && (\text{by } F \subseteq F^{\delta/2,k}) \\
&\geqslant \mu_i^*(F|t_i) - \delta && (\text{by } |\mu_i^*(F|t_i') - \mu_i^*(F|t_i)| < \delta) \\
&\geqslant \mu_i^*(E|t_i) - \delta && (\text{by } E \subseteq F)
\end{aligned}
$$

as claimed. ■

Lemma 9. *Fix $\delta > 0$, an integer $k \geqslant 1$ and a k-order frame $\mathcal{P}$ whose atoms all have d_i^k-diameter less than δ for every $i \in I$. Then, for every $m = 0, \ldots, k$,*

$$d_i^m(t_i, t_i') \leqslant d_{i,\mathcal{P}}^m(t_i, t_i') + m\delta \qquad \forall i \in I,\ \forall t_i, t_i' \in T_i^*.$$

Proof. Fix $\delta > 0$, an integer $k \geqslant 1$ and a k-order frame $\mathcal{P}$ whose atoms all have d_i^k-diameter less than δ, for every $i \in I$. (Such a frame exists by Lemma 8.) For $m = 0$ the conclusion of the lemma is trivial, as $d_i^0 = d_{i,\mathcal{P}}^0 = 0$. Consider $1 \leqslant m \leqslant k$ and assume the conclusion of the lemma holds for $m-1$. Let $t_i, t_i' \in T_i^*$ and $\eta > d_{i,\mathcal{P}_i}^m(t_i, t_i')$ and let us show that $d_i^m(t_i, t_i') \leqslant \eta + m\delta$. Fix $E \in 2^\Theta \otimes \mathcal{P}_{-i}$. Since all the atoms of $\mathcal{P}_{-i}$ have d_{-i}^k-diameter less than δ, there is some $F \in 2^\Theta \otimes \mathcal{P}_{-i}$ with $E \subseteq F \subseteq E^{\delta,m-1}$. Then,

$$
\begin{aligned}
\mu_i^*(E^{\eta+m\delta,m-1}|t_i') &\geqslant \mu_i^*(F^{\eta+(m-1)\delta,m-1}|t_i') && (\text{by } F \subseteq E^{\delta,m-1}) \\
&\geqslant \mu_i^*(F_{\mathcal{P}}^{\eta,m-1}|t_i') && (\text{by } d_{-i}^{m-1} \leqslant d_{-i,\mathcal{P}}^{m-1} + (m-1)\delta) \\
&\geqslant \mu_i^*(F|t_i) - \eta && (\text{by } \eta > d_{i,\mathcal{P}_i}^m(t_i, t_i')) \\
&\geqslant \mu_i^*(E|t_i) - \eta - m\delta, && (\text{by } E \subseteq F)
\end{aligned}
$$

and hence, $d_i^m(t_i, t_i') \leqslant \eta + m\delta$. But since our choice of $\eta > d_{i,\mathcal{P}_i}^m(t_i, t_i')$ was arbitrary, we have shown that $d_i^m(t_i, t_i') \leqslant d_{i,\mathcal{P}_i}^m(t_i, t_i') + m\delta$, as required. ■

Lemma 10. *Fix $\delta > 0$, a frame $\widetilde{\mathcal{P}}$ and a finite set $P \subset [0,1]^2$. Let $\mathcal{P}$ denote the common belief frame*

$$\left(\tilde{\mathcal{P}}_i \vee \bigvee_{\mathbf{p},E} \{C_i^{\mathbf{p}}(E), T_i^* \setminus C_i^{\mathbf{p}}(E)\}\right)_{i \in I},$$

where the join ranges over all $\mathbf{p} \in P$ and $E \in 2^\Theta \otimes \widetilde{\mathcal{P}}_1 \otimes \widetilde{\mathcal{P}}_2$. Then, for every pair of integers $k \geqslant \ell \geqslant 1$, every $\mathbf{p} \in P$ and every $E \in 2^\Theta \otimes \widetilde{\mathcal{P}}_1 \otimes \widetilde{\mathcal{P}}_2$,

$$\left(C_i^{\mathbf{p}}(E)\right)_{\mathcal{P}}^{\delta,\ell} \subseteq \left[B_i^{\mathbf{p}-(\delta,\delta)}\right]^{\ell-k}\left(E_{\widetilde{\mathcal{P}}}^{\delta,k}\right) \qquad \forall i \in I.$$

Proof. We fix $k \geqslant 1$, $E \in 2^\Theta \otimes \widetilde{\mathcal{P}}_1 \otimes \widetilde{\mathcal{P}}_2$ and $\mathbf{p} \in P$ and prove the result by induction on $\ell \geqslant k$. First, the result is trivial for $\ell = k$, as $\left[B_i^{\mathbf{p}-(\delta,\delta)}\right]^0(\cdot) = T_i^*$ for every i. Next, suppose the result is true for $\ell \geqslant k$ and let us show that it remains true for $\ell + 1$. Pick $t_i \in C_i^{\mathbf{p}}(E)$ and $t_i' \in T_i^*$ with $d_{i,\mathcal{P}}^{\ell+1}(t_i, t_i') < \delta$ and let us show that

$$t_i' \in B_i^{p_i-\delta}\left(E_{\widetilde{\mathcal{P}}}^{\delta,k} \cap \left[B^{\mathbf{p}-(\delta,\delta)}\right]^{\ell-k}\left(E_{\widetilde{\mathcal{P}}}^{\delta,k}\right)\right) = \left[B_i^{\mathbf{p}-(\delta,\delta)}\right]^{\ell+1-k}\left(E_{\widetilde{\mathcal{P}}}^{\delta,k}\right). \tag{27}$$

Indeed,[47]

$$\begin{aligned}
\mu_i^*\left(\left(E_{\widetilde{\mathcal{P}}}^{\delta,k} \cap \left[B^{\mathbf{p}-(\delta,\delta)}\right]^{\ell-k}\left(E_{\widetilde{\mathcal{P}}}^{\delta,k}\right)\right)_{t_i'} \middle| t_i'\right) &\geqslant \mu_i^*\left(\left(\left(E \cap C^{\mathbf{p}}(E)\right)_{\mathcal{P}}^{\delta,\ell}\right)_{t_i'} \middle| t_i'\right) \\
&\geqslant \mu_i^*\left(\left(\left(E \cap C^{\mathbf{p}}(E)\right)_{t_i}\right)_{\mathcal{P}}^{\delta,\ell} \middle| t_i'\right) \\
&\geqslant \mu_i^*\left(\left(E \cap C^{\mathbf{p}}(E)\right)_{t_i} \middle| t_i\right) - \delta \\
&\geqslant p_i - \delta,
\end{aligned}$$

where the first inequality follows from the induction hypothesis and the fact that $d_{i,\widetilde{\mathcal{P}}}^k \leqslant d_{i,\mathcal{P}}^\ell$, the second from $d_{i,\mathcal{P}}^\ell(t_i, t_i') < \delta$, the third from $d_{i,\mathcal{P}}^{\ell+1}(t_i, t_i') < \delta$, and the last from $t_i \in C_i^{\mathbf{p}}(E) = B_i^{p_i}(E \cap C^{\mathbf{p}}(E))$. This proves (27). ■

Lemma 11. *For every $\delta > 0$ and integer $k \geqslant 1$ there exists a common belief frame $\mathcal{P}$ such that for every $t_i, t_i' \in T_i^*$, if $d_{i,\mathcal{P}}^{\text{UW}}(t_i, t_i') < \delta$, then for every $\mathbf{p} \in [0,1]^2$ and measurable $E \subseteq \Omega$,*

$$t_i \in C_i^{\mathbf{p}}(E) \quad \Longrightarrow \quad t_i' \in C_i^{\mathbf{p}-\delta(k+2)\mathbf{1}}\left(E^{\delta(k+2),k}\right).$$

Thus, uniform-weak convergence in every common belief frame implies convergence in common beliefs.

Proof. Fix $\delta > 0$, an integer $k \geqslant 1$ and a finite subset $P \subset [0,1]^2$ with the property that for every $\mathbf{p} \in [0,1]^2$ there exists $\mathbf{q} \in P$ with $\mathbf{p} \geqslant \mathbf{q} \geqslant \mathbf{p} - (\delta, \delta)$. Let $\widetilde{\mathcal{P}}$ be a k-order frame whose atoms

[47] Recall that, for any measurable subset $E \subseteq \Omega$ and any type t_i of player i, E_{t_i} denotes the section of E over t_i. See Section 2.4.

have d_i^k-diameter at most δ, for every player i. Such a frame exists by Lemma 8. Consider the associated common belief frame $\mathscr{P}$ constructed as in Lemma 10:

$$\Big(\tilde{\mathscr{P}}_i \vee \bigvee_{\mathbf{p},E} \{C_i^{\mathbf{p}}(E), T_i^* \setminus C_i^{\mathbf{p}}(E)\}\Big)_{i\in I},$$

where the join ranges over all $\mathbf{p} \in P$ and $E \in 2^\Theta \otimes \tilde{\mathscr{P}}_1 \otimes \tilde{\mathscr{P}}_2$.

Fix $i \in I$, a pair of types $t_i, t_i' \in T_i^*$ with $d_{i,\mathscr{P}}^{UW}(t_i, t_i') < \delta$, an event $E \subseteq \Omega$ and $\mathbf{p} \in [0,1]^2$ with $t_i \in C_i^{\mathbf{p}}(E)$. Pick $\mathbf{q} \in P$ with $\mathbf{p} \geqslant \mathbf{q} \geqslant \mathbf{p} - (\delta, \delta)$ and $F \in 2^\Theta \otimes \tilde{\mathscr{P}}_1 \otimes \tilde{\mathscr{P}}_2$ such that $E \subseteq F \subseteq E^{\delta,k}$. (Such an event F exists, as the atoms of $\tilde{\mathscr{P}}$ have d_i^k-diameter at most δ.) Then,

$$\begin{aligned}
t_i' \in \bigcap_{\ell \geqslant k} \big(C_i^{\mathbf{p}}(F)\big)_{\mathscr{P}}^{\delta,\ell} &\subseteq \bigcap_{m \geqslant 1} \Big[B_i^{\mathbf{q}-\delta\mathbf{1}}\Big]^m \big(F_{\tilde{\mathscr{P}}}^{\delta,k}\big) && \text{(by } d_{i,\mathscr{P}}^{UW}(t_i, t_i') < \delta, \mathbf{q} \leqslant \mathbf{p} \text{ and Lemma 10)} \\
&= C_i^{\mathbf{q}-\delta\mathbf{1}}\big(F_{\tilde{\mathscr{P}}}^{\delta,k}\big) && \text{(by the definition of common belief)} \\
&\subseteq C_i^{\mathbf{q}-\delta\mathbf{1}}\big(F^{(k+1)\delta,k}\big) && \text{(by Lemma 9)} \\
&\subseteq C_i^{\mathbf{p}-2\delta\mathbf{1}}\big(E^{(k+2)\delta,k}\big), && \text{(by } F \subseteq E^{\delta,k} \text{ and } \mathbf{q} \geqslant \mathbf{p} - (\delta,\delta))
\end{aligned}$$

and hence $t_i' \in C_i^{\mathbf{p}-(k+2)\delta\mathbf{1}}\big(E^{(k+2)\delta,k}\big)$. ■

It follows from Lemma 11 that uniform-weak convergence relative to every frame implies convergence in common beliefs. Thus, by Theorem 2, strategic convergence implies convergence in common beliefs, as was to be shown.

A.6 A sequence converging in common beliefs but not strategically

We construct an example of a sequence of types t_1^n that converges in common beliefs to a type t_1, but does not converge uniform-weakly on a frame and hence, by Theorem 2, does not converge strategically. To construct the sequence, fix $\theta_0 \in \Theta$ and $0 < p < q < 1$. For each player i, pick a type r_i that satisfies the following two conditions:

(i) r_i assigns probability zero to state θ_0;

(ii) for every product-closed proper subset $E \subset \Omega$ and $\eta > 0$, $r_i \notin C_i^{\eta\mathbf{1}}(E)$.[48]

Let s_i and t_i be the types who assign probability one to θ_0 and whose beliefs about the other player's types are as specified in Figure 4 below.

[48]Types that satisfy these conditions exist. Ely and Pęski (2011, Theorem 1) show that the types that satisfy property (ii) are precisely those types to which product convergence is equivalent to strategic convergence, called *regular types*. They show that the set of regular types is a residual subset of the universal type space (in the product topology), in particular a non-empty set. As for condition (i), note that any type of player i that assigns probability one to some (θ, u_{-i}), where u_{-i} is a type of player $-i$ that satisfies property (ii), must also satisfy property (ii). This implies the existence of types satisfying both (i) and (ii).

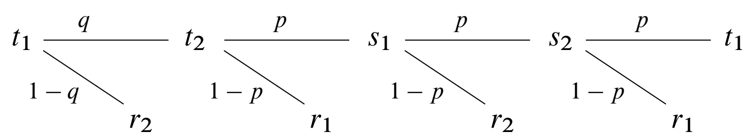

Figure 4: The types s_i and t_i.

Since r_1 and r_2 satisfy property (ii) above, for every $\mathbf{p} \in (0,1]^2$, every product-closed proper subset $E \subseteq \Omega$, and every $i \in I$,

$$t_i \in C_i^{\mathbf{p}}(E) \qquad \Longleftrightarrow \qquad \mathbf{p} \leqslant (p,p) \text{ and } E \supseteq \theta_0 \times \{s_1, t_1\} \times \{s_2, t_2\}. \tag{28}$$

In particular, no nontrivial event is common (q, p)-belief at t_1.

The construction of the sequence mimics the structure in Figure 4. Fix $0 < \Delta \leqslant q - p$ and for each player i define s_i^n and t_i^n as follows: let $t_1^1 = r_1$ and, for each $n \geqslant 1$, let s_2^n, s_1^n, t_2^n and t_1^{n+1} be the types who assign probability one to θ_0 and whose beliefs about the other player's types are as described in Figure 5 below.

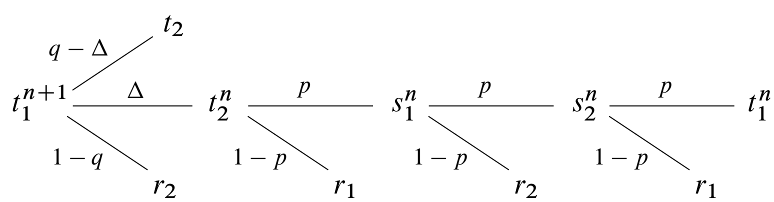

Figure 5: The sequences of types s_i^n and t_i^n.

The sequence t_1^n converges to t_1 in common beliefs. To see why, first note that $t_1^n \to t_1$ in the product topology: by the construction in Figures 4 and 5, s_2^1 has the same first-order belief as s_2, hence s_1^1 has the same second-order belief as s_1, which implies t_2^1 has the same third-order belief as t_2, and so forth. Second, given an arbitrary $\delta > 0$, by the construction in Figure 5 and the product-convergence $t_1^n \to t_1$, for each integer $k \geqslant 1$ we have $t_1^n \in B_1^{q-\Delta}\big(\theta_0 \times \{t_1\}^{\delta,k} \times t_2\big)$ for all n large enough. Since $t_2 \in C_2^{(p,p)}\big(\theta_0 \times \{s_1,t_1\} \times \{s_2,t_2\}\big)$ and $q - \Delta \geqslant p$, it follows that for all n large enough,

$$t_1^n \in B_1^p\Big(\theta_0 \times \{t_1\}^{\delta,k} \times C_2^{(p,p)}\big(\theta_0 \times \{s_1,t_1\} \times \{s_2,t_2\}\big)\Big) \subseteq C_1^{(p,p)}\Big(\big(\theta_0 \times \{s_1,t_1\} \times \{s_2,t_2\}\big)^{\delta,k}\Big).$$

It follows, by (28) and footnote 6, that $t_1^n \to t_1$ in common beliefs, as claimed.

It remains to exhibit a frame $\mathcal{P}$ such that $t_1^n \not\to t_1$ uniform-weakly on $\mathcal{P}$. For each $\mathbf{r} = (r_1, \ldots, r_4) \in (0,1]^4$, $i \in I$ and measurable subset $E \subseteq \Omega$, let $\big[M_i^{\mathbf{r}}\big]^0(E) = \Omega$, and for each $k \geqslant 1$, define recursively

$$\big[M_i^{\mathbf{r}}\big]^k(E) = B_i^{r_1}\big(E \cap B_{-i}^{r_2}\big(E \cap B_i^{r_3}\big(E \cap B_{-i}^{r_4}\big(E \cap \big[M_i^{\mathbf{r}}\big]^{k-1}(E)\big)\big)\big)\big).$$

Note that

$$\left[M_i^{\mathbf{r}}\right]^k(E) \subseteq B_i^{r_1}\Big(E \cap \Big[M_{-i}^{(r_2,r_3,r_4,r_1)}\Big]^{k-1}(E)\Big) \subseteq \left[M_i^{\mathbf{r}}\right]^{k-1}(E) \qquad \forall k \geqslant 1. \tag{29}$$

For each $k \geqslant 1$, let

$$P_1^k = \Big[M_1^{(q,p,p,p)}\Big]^k(\theta_0), \qquad Q_1^k = \Big[M_1^{(p,p,q,p)}\Big]^k(\theta_0),$$
$$P_2^k = \Big[M_2^{(p,p,p,q)}\Big]^k(\theta_0), \qquad Q_2^m = \Big[M_2^{(p,q,p,p)}\Big]^k(\theta_0)$$

and finally,

$$P_i = \bigcap_{k=1}^{\infty} P_i^k \qquad \text{and} \qquad Q_i = \bigcap_{k=1}^{\infty} Q_i^k \qquad \forall i \in I.$$

Let $\mathcal{P}$ be the profile of partitions $(\mathcal{P}_i)_{i\in I}$,

$$\mathcal{P}_i = \big\{P_i \setminus Q_i,\, Q_i \setminus P_i,\, P_i \cap Q_i,\, T_i^* \setminus (P_i \cup Q_i)\big\} \qquad \forall i \in I.$$

To verify that $\mathcal{P}$ is a frame, note that (29) implies

$$P_1^k \subseteq B_1^q(\theta_0 \times P_2^{k-1}) \subseteq P_1^{k-1} \qquad \forall k \geqslant 1$$

and hence

$$P_1 = \bigcap_{k=2}^{\infty} P_1^k \subseteq \bigcap_{k=2}^{\infty} B_1^q(\theta_0 \times P_2^{k-1}) = B_1^q(\theta_0 \times P_2) \subseteq \bigcap_{k=2}^{\infty} P_1^{k-1} = P_1.$$

Therefore, $P_1 = B_1^q(\theta_0 \times P_2)$. By analogous arguments,

$$Q_1 = B_1^p(\theta_0 \times Q_2), \qquad P_2 = B_2^p(\theta_0 \times Q_1), \qquad \text{and} \qquad Q_2 = B_2^p(\theta_0 \times P_1).$$

Thus, for all $i \in I$ and $t_i \in T_i^*$, the element of $\mathcal{P}_i$ containing t_i is determined by the values $\mu_i^*(\theta_0 \times P_{-i}|t_i)$ and $\mu_i^*(\theta_0 \times Q_{-i}|t_i)$, and hence, a fortiori, by the restriction of $\mu_i^*(\cdot|t_i)$ to $\Theta \times \mathcal{P}_{-i}$. This shows that $\mathcal{P}$ is a frame.

We now prove that t_1^n does not converge to t_1 uniform weakly on $\mathcal{P}$. Fix $0 < \delta < \min\{p, \Delta\}$. It is enough to show that for every positive integer n,

$$d_{1,\mathcal{P}}^{4n-3}(t_1, t_1^n) > \delta, \quad d_{2,\mathcal{P}}^{4n-2}(s_2, s_2^n) > \delta, \quad d_{1,\mathcal{P}}^{4n-1}(s_1, s_1^n) > \delta, \quad d_{2,\mathcal{P}}^{4n}(t_2, t_2^n) > \delta. \tag{30}$$

To prove this, we show that the first inequality holds for $n = 1$, that for each $n \geqslant 1$, if the first inequality holds for n, then all others also hold for n, and finally that if the last inequality holds for n, then the first holds for $n + 1$. In the proof we use the following facts, which are immediate from the definition of t_1, t_2, s_1, s_2 and the fact that $\mu_i^*(\theta_0|r_i) = 0$ and $P_i \cup Q_i \subseteq B_i^p(\theta_0)$ for each $i \in I$:

$$t_1 \in P_1, \quad s_1 \in Q_1, \quad s_2 \in Q_2, \quad t_2 \in P_2, \quad r_i \notin (P_i \cup Q_i)_{\mathcal{P}}^{\delta,1} \quad \forall i \in I. \tag{31}$$

Since $\mu_1^*(\theta_0|t_1) = 1$ and $\mu_1^*(\theta_0|t_1^1) = 0$, we have $d_{1,\mathcal{P}}^1(t_1, t_1^1) = 1 > \delta$, which proves the first inequality in (30) for $n = 1$. Now fix any $n \geqslant 1$ and assume the first inequality in (30) holds for n. Then, by (31), $\{t_1^n, r_1\} \cap (P_1)_{\mathcal{P}}^{\delta,4n-3} = \varnothing$. Since $\mu_2^*(\theta_0 \times \{t_1^n, r_1\}|s_1^n) = 1$ and, by (31), $\mu_2^*(\theta_0 \times P_1|s_2) = \mu_2^*(\theta_0, t_1|s_2) = p$, it follows that $d_{2,\mathcal{P}}^{4n-2}(s_2, s_2^n) \geqslant p > \delta$. Thus, the second inequality in (30) holds for n. The latter implies, by (31), $\{s_2^n, r_2\} \cap (P_2)_{\mathcal{P}}^{\delta,4n-2} = \varnothing$, and since $\mu_1^*(\theta_0 \times \{s_2^n, r_2\}|s_1^n) = 1$ and, by (31), $\mu_1^*(\theta_0 \times Q_2|s_1) = \mu_1^*(\theta_0, s_2|s_1) = p$, we also have $d_{1,\mathcal{P}}^{4n-1}(s_1, s_1^n) \geqslant p > \delta$, that is, the third inequality in (30) holds for n. This in turn implies, by (31), $\{s_1^n, r_1\} \cap (P_1)_{\mathcal{P}}^{\delta,4n-1} = \varnothing$, and since $\mu_2^*(\theta_0 \times \{s_1^n, r_1\}|t_2^n) = 1$ and, by (31), $\mu_2^*(\theta_0 \times Q_1|t_2) = \mu_1^*(\theta_0, s_1|t_2) = p$, we obtain $d_{2,\mathcal{P}}^{4n}(t_2, t_2^n) \geqslant p > \delta$. This proves that the fourth inequality in (30) holds for n, and hence, by (31), that $\{t_2^n, r_2\} \cap (P_2)_{\mathcal{P}}^{\delta,4n} = \varnothing$. It remains to show that the latter implies that the first inequality in (30) holds for $n + 1$. Indeed, since $\mu_1^*(\theta_0 \times \{t_2^n, r_2\}|t_1^{n+1}) = 1 - q + \Delta$ and, by (31), $\mu_1^*(\theta_0 \times P_2|t_1) = \mu_1^*(\theta_0, t_2|t_1) = q$, we have $d_{1,\mathcal{P}}^{4n+1}(t_1, t_1^{n+1}) \geqslant \Delta > \delta$, as claimed.

A.7 Proof of Corollary 2

An immediate implication of Ely and Pęski (2011, Theorem 1) is that every finite type is critical. This fact, together with the denseness of finite types in the strategic topology (Dekel, Fudenberg, and Morris, 2006), implies that the set of critical types is dense in the strategic topology.

Next, we show that the set of critical types is open in the strategic topology, or equivalently, that the set of regular types is closed. Suppose not. Then, there is a sequence of regular types t_i^n that converges to some critical type t_i. By Ely and Pęski (2011, Theorem 1), there is some $p > 0$ and some product-closed, proper subset $E \subset \Omega$ with $t_i \in C_i^{p\mathbf{1}}(E)$. Then, there is an integer $k \geqslant 1$ and $\delta \in (0, p)$ such that the d_i^k-closure of $E^{\delta,k}$ is a proper subset of Ω. Moreover, by Corollary 1, $t_i^n \to t_i$ in common beliefs, and hence $t_i^n \in C_i^{(p-\delta)\mathbf{1}}(E^{\delta,k})$ for all n large enough. It follows, again by Ely and Pęski (2011, Theorem 1), that t_i^n is a critical type for all n large enough, and this is a contradiction. The contradiction shows that the set of regular types is closed.

A.8 Proof of Theorem 3

Since the implications $(a) \Leftrightarrow (b) \Rightarrow (c) \Leftrightarrow (d) \Rightarrow (e)$ follow from Theorems 1, 2 and from Corollary 1, it is enough to prove the implication $(e) \Rightarrow (a)$. The latter is an immediate consequence of the following lemma.

Lemma 12. *Let (T_1, T_2) be a finite type space. There exist $\eta > 0$ and $k \geqslant 1$ such that for each $\delta \in (0, \eta)$, $\varepsilon \in [0, 1)$, $m \geqslant 0$, $i \in I$ and $(t_i, t_i') \in T_i \times T_i^*$ with $d_i^k(t_i, t_i') < \delta$, one has*

$$t_i' \in \left[B_i^{(1-\varepsilon,1-\varepsilon)}\right]^m \left(\Theta \times T_1^{\delta,k} \times T_2^{\delta,k}\right) \quad \Rightarrow \quad d_i^{k+m}(t_i, t_i') < \delta + \varepsilon \,. \tag{32}$$

Proof. Choose $k \geqslant 1$ and $\eta > 0$ so that, for each $i \in I$ and $t_i, t_i' \in T_i$, if $d_i^{k-1}(t_i, t_i') < 2\eta$ then $t_i' = t_i$. Since T_1 and T_2 are finite, such k and η exist. Fix $\delta \in (0, \eta)$. Thus, for each $n \geqslant 0$, $i \in I$, $t_i \in T_i$ and $E \subseteq \Theta \times T_{-i}$,

$$E^{\delta,k-1} \cap (\Theta \times T_{-i})^{\delta,k} \subseteq E^{\delta,k}. \tag{33}$$

The proof of (32) is by induction in m. Obviously, (32) holds for $m = 0$. Now assume that it holds for some $m \geqslant 0$ and fix $i \in I$, $t_i \in T_i$ and $t_i' \in T_i^*$ with $t_i' \in \left[B_i^{(1-\varepsilon,1-\varepsilon)}\right]^{m+1}(\Theta \times T_1^{\delta,k} \times T_2^{\delta,k})$ and $d_i^k(t_i, t_i') < \delta$. Since $\mu_i^*(\Theta \times T_{-i}|t_i) = 1$, in order to prove that $d_i^{k+m+1}(t_i, t_i') < \delta + \varepsilon$ it suffices to show that $\mu_i^*(E^{\delta+\varepsilon,k+m}|t_i') \geqslant \mu_i^*(E|t_i) - \delta - \varepsilon$ for each $E \subseteq \Theta \times T_{-i}$. Indeed,

$$\begin{aligned}
\mu_i^*\big(E^{\delta+\varepsilon,k+m}\big|t_i'\big) &\geq \mu_i^*\big(E^{\delta,k} \cap \left[B_{-i}^{(1-\varepsilon,1-\varepsilon)}\right]^m\big(\Theta \times T_1^{\delta,k} \times T_2^{\delta,k}\big)\big|t_i'\big) \\
&\geq \mu_i^*\big(E^{\delta,k-1} \cap (\Theta \times T_{-i})^{\delta,k} \cap \left[B_{-i}^{(1-\varepsilon,1-\varepsilon)}\right]^m\big(\Theta \times T_1^{\delta,k} \times T_2^{\delta,k}\big)\big|t_i'\big) \\
&\geq \mu_i^*\big(E^{\delta,k-1}\big|t_i'\big) - \varepsilon \geq \mu_i^*\big(E\big|t_i\big) - \delta - \varepsilon,
\end{aligned}$$

where the first inequality follows from the induction hypothesis, the second from (33), the third from $t_i' \in \left[B_i^{(1-\varepsilon,1-\varepsilon)}\right]^{m+1}(T^{\delta,k})$ and the fourth from $d_i^k(t_i, t_i') < \delta$. ■

A.9 Proofs of Theorems 4 and 5

The proofs of the theorems rely on the four lemmas below. Before presenting the results, we need some preliminaries. First, in the proofs we use the following notation: for each $\delta > 0$, $i \in I$ and measurable $E \subseteq \Theta \times T_i^*$, we let $E^{\delta,\infty}$ denote the uniform-weak δ-neighborhood of E, that is,

$$E^{\delta,\infty} = \bigcap_{k \geqslant 1} E^{\delta,k}.$$

Second, we say a finite type space $T = (T_i)_{i \in I}$ has *full support* if $\mu_i^*(E|t_i) > 0$ for every $i \in I$, $t_i \in T_i$ and $E \subseteq \Theta \times T_{-i}$. Third, generalizing our notations h_{d^P} and h_{d^S}, we define the metric d^{UW} on T^* by letting $d^{UW}(t, t') = \max_{i \in I} d_i^{UW}(t_i, t_i')$ for every $t = (t_i)_{i \in I}$ and $t' = (t_i)_{i \in I}$ in T^*, and given two type spaces $T = (T_i)_{i \in I}$ and $T' = (T_i')_{i \in I}$ in $\mathcal{T}$ we write $h_{d^{UW}}(T, T')$ for the Hausdorff distance between T and T' under d^{UW}, that is,

$$h_{d^{UW}}(T, T') = \max\left\{\sup_{t \in T} \inf_{t' \in T'} d^{UW}(t, t'),\ \sup_{t' \in T'} \inf_{t \in T} d^{UW}(t, t')\right\}.$$

Since for each $i \in I$ the uniform-weak topology on T_i^* is finer than the product topology on T_i^*, and hence product-closed sets in T_i^* are also uniform-weakly closed, $h_{d^{UW}}$ is also a metric on $\mathcal{T}$.

Lemma 13. *Let $T = (T_i)_{i \in I}$ be a finite type space with full support. Let $j \in I$ and $\bar{t}_j \in T_j$ and suppose there is a sequence of common prior types t_j^n such that $d_j^{UW}(\bar{t}_j, t_j^n) \to 0$. Then there is a sequence of finite common prior type spaces T^n such that $h_{d^{UW}}(T, T^n) \to 0$.*

Proof. Since T is a finite type space with full support, there exists $\eta > 0$ such that, for each $i \in I$, both of the following hold:

$$d_i^{UW}(t_i, t_i') < 2\eta \quad \Rightarrow \quad t_i' = t_i \qquad \forall t_i, t_i' \in T_i, \tag{34}$$
$$\eta < \mu_i^*(E|t_i) \qquad \forall t_i \in T_i,\ \forall \varnothing \neq E \subseteq \Theta \times T_{-i}. \tag{35}$$

Fix $\delta \in (0, \eta)$ and choose n such that $d_j^{UW}(\bar{t}_j, t_j^n) < \delta$. Let π^n be a common prior with $\pi^n(t_j^n) > 0$. Since π^n is a common prior, for each $i \in I$ we have

$$\pi^n\big(\{t_i\}^{\delta,\infty}\big) > 0 \ \Rightarrow\ \Big|\pi^n\big(E^{\delta,\infty}\big|\{t_i\}^{\delta,\infty}\big) - \mu_i^*\big(E|t_i\big)\Big| < \delta \quad \forall t_i \in T_i,\ \forall E \subseteq \Theta \times T_{-i}.\text{[49]} \tag{36}$$

Since $\pi^n\big(\{\bar{t}_j\}^{\delta,\infty}\big) \geqslant \pi^n(t_j^n) > 0$, it follows from (35) and (36) that, for each $i \in I$,

$$\pi^n\big(\{t_i\}^{\delta,\infty}\big) > 0 \ \text{ and } \ \Big|\pi^n\big(E^{\delta,\infty}\big|\{t_i\}^{\delta,\infty}\big) - \mu_i^*\big(E|t_i\big)\Big| < \delta \quad \forall t_i \in T_i,\ \forall E \subseteq \Theta \times T_{-i}. \tag{37}$$

Let $(T_i^n)_{i \in I}$ be the type space defined as follows. For each $i \in I$, $T_i^n = \{f_i^n(t_i) : t_i \in T_i\}$, where

$$\mu_i^*\big(\theta \times f_{-i}^n(t_{-i})\big| f_i^n(t_i)\big) = \frac{\pi^n\big(\theta \times \{t_i\}^{\delta,\infty} \times \{t_{-i}\}^{\delta,\infty}\big)}{\pi^n\big(\Theta \times \{t_i\}^{\delta,\infty} \times T_{-i}^{\delta,\infty}\big)} \qquad \forall \theta \in \Theta,\ \forall t_i \in T_i,\ \forall t_{-i} \in T_{-i},$$

which is well defined by (34). Clearly, $(T_i^n)_{i \in I}$ is a common prior type space, with the common prior $\bar{\pi}^n$ defined as follows:

$$\bar{\pi}^n\big(\theta, f_1^n(t_1), f_2^n(t_2)\big) = \frac{\pi^n\big(\theta \times \{t_i\}^{\delta,\infty} \times \{t_{-i}\}^{\delta,\infty}\big)}{\pi^n\big(\Theta \times T_i^{\delta,\infty} \times T_{-i}^{\delta,\infty}\big)} \qquad \forall (\theta, t_1, t_2) \in \Theta \times T_1 \times T_2.$$

Moreover, for each $i \in I$, $t_i \in T_i$ and $k \geqslant 1$ we have $d_i^k(t_i, f_i^n(t_i)) < \delta$. To prove the claim for $k = 1$, note that for each $E \subseteq \Theta$ we have

$$\mu_i^*\big(E\big| f_i^n(t_i)\big) = \frac{\pi^n\big(E \times \{t_i\}^{\delta,\infty} \times T_{-i}^{\delta,\infty}\big)}{\pi^n\big(\Theta \times \{t_i\}^{\delta,\infty} \times T_{-i}^{\delta,\infty}\big)} \geqslant \pi^n\big(E \times T_{-i}^{\delta,\infty}\big|\{t_i\}^{\delta,\infty}\big) \geqslant \mu_i^*(E|t_i) - \delta,$$

where the second inequality follows from (37). Proceeding by induction, suppose the claim is true for some $k \geqslant 1$. Then, for each $E \subseteq \Theta \times T_{-i}$,

$$\begin{aligned}\mu_i^*\big(E^{\delta,k}\big| f_i^n(t_i)\big) &\geqslant \mu_i^*\big(\big\{(\theta, f_{-i}^n(t_{-i})) : (\theta, t_{-i}) \in E\big\}\big| f_i^n(t_i)\big) \\ &= \frac{\pi^n\big(E^{\delta,\infty} \times \{t_i\}^{\delta,\infty}\big)}{\pi^n\big(\Theta \times \{t_i\}^{\delta,\infty} \times T_{-i}^{\delta,\infty}\big)} \geqslant \pi^n\big(E^{\delta,\infty}\big|\{t_i\}^{\delta,\infty}\big) \geqslant \mu_i^*(E|t_i) - \delta,\end{aligned}$$

where the first inequality follows from the induction hyothesis, and the third from (37). Thus, $d_i^{UW}(t_i, f_i^n(t_i)) < \delta$ for each $i \in I$ and $t_i \in T_i$, and hence $h_{d^{UW}}(T, T^n) < \delta$. ■

[49]By definition of conditional probability, if π is a common prior, then, for each $i \in I$ and measurable $E \subseteq T_i^*$ with $\pi(E) > 0$, we have $\pi(F|E) = (1/\pi(E)) \int_E \mu_i^*(F|t_i)\pi(dt_i)$ for each measurable $F \subseteq \Theta \times T_{-i}^*$.

Lemma 14. *For each $\varepsilon > 0$ and finite type space T there exists a finite non-common prior type space with full support T' such that $h_{d^{UW}}(T, T') < \varepsilon$.*

Proof. Fix $\varepsilon > 0$ and a finite type space $T = (T_i)_{i \in I}$. In order to construct the type space T', let $\delta \in \left(0, \frac{\varepsilon}{2}\right)$ and consider first the type space $T'' = (T_i'')_{i \in I}$ defined as follows. For each $i \in I$, $T_i'' = \{f_i(t_i) : t_i \in T_i\}$, where for each $t_i \in T_i$ and $E \subseteq \Theta \times T_{-i}$,

$$\mu_i^*\left(\left\{(\theta, f_{-i}(t_{-i})) : (\theta, t_{-i}) \in E\right\} \middle| f_i(t_i)\right) = (1 - \delta)\, \mu_i^*(E|t_i) + \frac{\delta|E|}{|\Theta \times T_{-i}|}.$$

Clearly, T'' has full support, and an immediate induction argument shows that $d_i^{UW}(t_i, f_i(t_i)) < \frac{\varepsilon}{2}$ for each $i \in I$ and $t_i \in T_i$. Thus, $h_{d^{UW}}(T, T'') < \frac{\varepsilon}{2}$. If T'' is a non-common prior type space, then setting $T' = T''$ completes the proof. Otherwise, assume that T'' admits a common prior π, fix $(\bar{t}_1, \bar{t}_2) \in T''$ and two distinct $\theta', \theta'' \in \Theta$, let $0 < \delta' < \min_{i \in I, t_i'' \in T_i'', E \subseteq \Theta \times T_{-i}''} \mu_i^*(E|t_i)$, and let $T' = (T_i')_{i \in I}$ be the type space defined as follows. For each $i \in I$, $T_i' = \{f_i'(t_i'') : t_i'' \in T_i''\}$, where

$$\mu_2^*(\theta', f_1'(\bar{t}_1)|f_2'(\bar{t}_2)) = \mu_2^*(\theta', \bar{t}_1|\bar{t}_2) + \delta', \qquad \mu_2^*(\theta'', f_1'(\bar{t}_1)|f_2'(\bar{t}_2)) = \mu_2^*(\theta'', \bar{t}_1|\bar{t}_2) - \delta'$$

and, for every $i \in I$ and $(\theta, t_i'', t_{-i}'') \in \Theta \times T_i'' \times T_{-i}''$ for which $\mu_i^*(\theta, f_{-i}'(t_{-i}'')|f_i'(t_i''))$ is not yet defined,

$$\mu_i^*(\theta, f_{-i}'(t_{-i}'')|f_i'(t_i'')) = \mu_i^*(\theta, t_{-i}''|t_i'').$$

Clearly, T' has full support, and an immediate induction argument shows that $h_{d^{UW}}(T'', T') < \frac{\varepsilon}{2}$. Thus, $h_{d^{UW}}(T, T') \leqslant h_{d^{UW}}(T, T'') + h_{d^{UW}}(T'', T') < \varepsilon$. Moreover, T' does not admit a common prior, because

$$\begin{aligned}\frac{\mu_1^*(\theta', f_2'(\bar{t}_2)|f_1'(\bar{t}_1))}{\mu_1^*(\theta'', f_2'(\bar{t}_2)|f_1'(\bar{t}_1))} = \frac{\mu_1^*(\theta', \bar{t}_2|\bar{t}_1)}{\mu_1^*(\theta'', \bar{t}_2|\bar{t}_1)} &= \frac{\pi(\theta', \bar{t}_1, \bar{t}_2)}{\pi(\theta'', \bar{t}_1, \bar{t}_2)} \\ &= \frac{\mu_2^*(\theta', \bar{t}_1|\bar{t}_2)}{\mu_2^*(\theta'', \bar{t}_1|\bar{t}_2)} < \frac{\mu_2^*(\theta', f_1'(\bar{t}_1)|f_2'(\bar{t}_2))}{\mu_2^*(\theta'', f_1'(\bar{t}_1)|f_2'(\bar{t}_2))},\end{aligned}$$

where the second and third equality follow from the fact that π is a common prior for T''. ■

Lemma 15. *For every finite non-common prior type space with full support $T = (T_i)_{i \in I}$ there exists $\varepsilon > 0$ such that $h_{d^{UW}}(T, T') > \varepsilon$ for every common prior type space T'. Furthermore, $d_i^{UW}(t_i, t_i') > \varepsilon$ for every $i \in I$, $t_i \in T_i$ and common prior type $t_i' \in T_i^*$.*

Proof. The second claim is an immediate consequence of the first claim and the proof of Lemma 13. To prove the first claim, first note that since T is finite, there exists $\eta > 0$ such that

$$d_i^{UW}(t_i, t_i') < 2\eta \quad \Rightarrow \quad t_i' = t_i \qquad \forall t_i, t_i' \in T_i. \tag{38}$$

Moreover, since T is finite and has no common prior, there exists $f : \Theta \times T \to \mathbb{R}$ such that

$$\sum_{(\theta,t_2')\in\Theta\times T_2} f(\theta,t_1,t_2')\mu_1^*(\theta,t_2'|t_1) > 0 > \sum_{(\theta,t_1')\in\Theta\times T_1} f(\theta,t_1',t_2)\mu_2^*(\theta,t_1'|t_2) \qquad \forall (t_1,t_2)\in T. \quad (39)$$

(See, for instance, Samet, 1998) Now suppose, contrary to our claim, that $h_{d^{UW}}(T,T^n) \to 0$ for a sequence of common prior type spaces T^n. Choose N so that $h_{d^{UW}}(T,T^n) < \eta$ for each $n \geqslant N$, and for each such n define $f^n : \Theta \times T^n \to \mathbb{R}$ by letting

$$f^n(\theta,t_1^n,t_2^n) = f(\theta,t_1,t_2) \qquad \forall (t_1,t_2)\in T,\ \forall (t_1^n,t_2^n)\in T^n \cap (\{t_1\}^{\eta,\infty}) \times \{t_2\}^{\eta,\infty}), \quad (40)$$

which is well defined by (38) and by the fact that $T_i^n \subseteq T_i^{\eta,\infty}$ for each $i \in I$. Then, it follows from (39) and (40) that, for each sufficiently large n,

$$\sum_{(\theta,t_2')\in\Theta\times T_2^n} f^n(\theta,t_1^n,t_2')\mu_1^*(\theta,t_2'|t_1^n) > 0 > \sum_{(\theta,t_1')\in\Theta\times T_1^n} f(\theta,t_1',t_2^n)\mu_2^*(\theta,t_1'|t_2^n) \qquad \forall (t_1^n,t_2^n)\in T^n.$$

But, letting π^n denote the common prior for the type space T^n, we then reach the contradiction that $\sum_{(\theta,t^n)\in\Theta\times T^n} \pi^n(\theta,t^n)$ is both positive and negative. ■

Lemma 16. *Let $T = (T_i)_{i\in I}$ be a finite type space, and let $T^n = (T_i^n)_{i\in I}$ be a sequence of type spaces in $\mathcal{T}$. Then, $h_{d^P}(T,T^n) \to 0$ if and only if $h_{d^{UW}}(T,T^n) \to 0$.*

Proof. Since uniform-weak convergence implies product convergence, it suffices to prove the "only if" part. First note that, since $(T_i)_{i\in I}$ is a finite type space, by Lemma 12 there exist $\eta > 0$ and $k \geqslant 1$ such that for every $\delta \in (0,\eta)$, $i \in I$ and $t_i \in T_i$,

$$\{t_i\}^{\delta,k} \cap C_i^{\mathbf{1}}\big(\Theta \times T_1^{\delta,k} \times T_2^{\delta,k}\big) \subseteq \{t_i\}^{\delta,\infty}. \quad (41)$$

Now fix any $\delta \in (0,\eta)$. Since $h_{d^P}(T,T^n) \to 0$, there exists $N \geqslant 1$ such that

$$\sup_{t_i\in T_i}\inf_{t_i^n\in T_i^n} d_i^k(t_i,t_i^n) < \delta \quad \text{and} \quad \sup_{t_i^n\in T_i^n}\inf_{t_i\in T_i} d_i^k(t_i,t_i^n) < \delta \qquad \forall i\in I,\ \forall n \geqslant N. \quad (42)$$

Then, for each $i \in I$ and $n \geqslant N$, we have $T_i^n \subseteq T_i^{\delta,k}$ and hence, using the fact that $T_i^n \subseteq C_i^{\mathbf{1}}(\Theta\times T_1^n \times T_2^n)$, also $T_i^n \subseteq C_i^{\mathbf{1}}(\Theta \times T_1^{\delta,k} \times T_2^{\delta,k})$. It follows from (41) that for each $i \in I$, $n \geqslant N$, $t_i \in T_i$ and $t_i^n \in T_i^n$, if $d_i^k(t_i,t_i^n) < \delta$ then $d_i^{UW}(t_i,t_i^n) < \delta$. Thus, by (42), $h_{d^{UW}}(T,T^n) < \delta$. ■

We can now prove Theorems 4 and 5. Dekel, Fudenberg, and Morris (2006, Theorem 3) show that finite types are dense under the strategic topology. Since the uniform weak topology is finer than the strategic topology, it follows from Lemma 14 that the set of types that belong to some finite non-common prior type space with full support are also dense under the strategic topology. By Lemma 15 and Theorem 3, each such type belongs to a strategically open set that does not contain any common prior type. Thus, Theorem 4 follows.

Mertens and Zamir (1985, Theorem 3.1) show that the set of finite type spaces is dense in $\mathcal{T}$ under the product topology on type spaces. Since the uniform weak topology is finer than the product topology, it follows from Lemma 14 that the set of finite non-common prior type spaces with full support is also dense in $\mathcal{T}$ under the product topology on type spaces. By Lemma 15 and Lemma 16, each finite non-common prior type space with full support also belongs to a subset of $\mathcal{T}$ that is open under the product topology on type spaces and does not contain any common prior type space. Thus, Theorem 5 follows.

References

Aumann, Robert J. (1976), "Agreeing to disagree." *Annals of Statistics*, 4, 1236–1239.

Aumann, Robert J. (1987), "Correlated equilibrium as an expression of bayesian rationality." *Econometrica*, 55, 1–18.

Aumann, Robert J. (1998), "Common priors: A reply to Gul." *Econometrica*, 66, 929–938.

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

Bernheim, B. Douglas (1984), "Rationalizable strategic behavior." *Econometrica*, 52, 1007–1028.

Billingsley, P. (1999), *Convergence of Probability Measures*. John Wiley & Sons, 2nd edition.

Blackwell, David (1951), "Comparison of experiments." In "Proceedings of the second Berkeley symposium on mathematical statistics and probability," volume 1, pp. 93–102.

Blackwell, David (1953), "Equivalent comparisons of experiments." *The annals of mathematical statistics*, 24, 265–272.

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

Brandenburger, Adam and Eddie Dekel (1993), "Hierarchies of beliefs and common knowledge." *Journal of Economic Theory*, 59, 189–198.

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

Chen, Yi-Chun, Alfredo Di Tillio, Eduardo Faingold, and Siyang Xiong (2016), "Supplemental appendix for 'Characterizing the strategic impact of misspecified beliefs'."

Chen, Yi-Chun and Siyang Xiong (2013), "Genericity and robustness of full surplus extraction." *Econometrica*, 81, 825–847.

Crémer, Jacques and Richard P. McLean (1988), "Full extraction of the surplus in bayesian and dominant strategy auctions." *Econometrica*, 56, 1247–1257.

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

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

Ely, Jeffrey C. and Marcin Pęski (2011), "Critical types." *Review of Economic Studies*, 78, 907–937.

Gul, Faruk (1998), "A comment on Aumann's bayesian view." *Econometrica*, 66.

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

Heifetz, Aviad and Zvika Neeman (2006), "On the generic (im)possibility of full surplus extraction in mechanism design." *Econometrica*, 74, 213–234.

Kajii, Atsushi and Stephen Morris (1997), "The robustness of equilibria to incomplete information." *Econometrica*, 65, 1283–1309.

Kajii, Atsushi and Stephen Morris (1998), "Payoff continuity in incomplete information games." *Journal of Economic Theory*, 82, 267–276.

Kohlberg, Elon and Jean-Francois Mertens (1986), "On the strategic stability of equilibria." *Econometrica*, 54, 1003–1037.

Lehmann, Eric L. (1988), "Comparing location experiments." *Annals of Statistics*, 16, 521–533.

Lipman, Barton (2003), "Finite order implications of common priors." *Econometrica*, 71, 1255–1267.

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

Milgrom, Paul and Nancy Stokey (1982), "Information, trade and common knowledge." *Journal of Economic Theory*, 26, 17–27.

Monderer, Dov and Dov Samet (1989), "Approximating common knowledge with common beliefs." *Games and Economic Behavior*, 1, 170–190.

Monderer, Dov and Dov Samet (1996), "Proximity of information in games with incomplete information." *Mathematics of Operations Research*, 21, 707–725.

Morris, Stephen (1994), "Trade with heterogeneous prior beliefs and asymmetric information." *Econometrica*, 62, 1327–1347.

Morris, Stephen (1995), "The common prior assumption in economic theory." *Economics and Philosophy*, 11, 227–253.

Morris, Stephen, Hyun Song Shin, and Muhamet Yildiz (2015), "Common belief foundations of global games." *Working Paper, Princeton University*.

Myerson, R.B. (1981), "Optimal auction design." *Mathematics of Operations Research*, 6, 58–73.

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

Quah, John K-H and Bruno Strulovici (2009), "Comparative statics, informativeness, and the interval dominance order." *Econometrica*, 77, 1949–1992.

Rubinstein, Ariel (1982), "Perfect equilibrium in a bargaining model." *Econometrica*, 50, 97–109.

Rubinstein, Ariel (1989), "The electronic mail game: A game with almost common knowledge." *American Economic Review*, 79, 389–391.

Samet, Dov (1998), "Common priors and separation of convex sets." *Games and Economic Behavior*, 24, 172–174.

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

Yildiz, Muhamet (2003), "Bargaining without a common prior–an immediate agreement theorem." *Econometrica*, 71, 793–811.