Misspecified Information in Dynamic Environments*

Evan Piermont

Royal Holloway, University of London

Peio Zuazo-Garin

HSE University, International College of Economics and Finance

We study dynamic games in which the *set* of all possible payoff-relevant parameters is *not* commonly known—thus covering usual economic scenarios that entail discrepancies about the forms in which asymmetric information can take place, or misjudgments about the information that own choices may reveal to others. Within a novel framework that allows for this flexibility, we: (1) characterize the robustness of different dynamic variants of rationalizability according to their reliance on observed behavior for belief-updating: while those that may neglect unexpected observed moves are robust (*weak* and *backward* rationalizability), that which exploits them is not (*strong* rationalizability); (2) identify sufficient conditions on private information for the robustness of strong rationalizability, and establish the latter's *genericity* by showing that these conditions hold everywhere in the universal type space, except in a meager set of "knife-edge" cases; (3) unveil an impossibility theorem that confronts the economic analyst with the following dilemma: either excluding payoff-states in ad hoc fashion, or giving up on sharpening predictions via dynamic criteria; (4) show that the predictions that can be uniquely selected by weak rationalizability when small noise about beliefs and information à la Weinstein and Yildiz (2007) is introduced coincide, exactly, with those of strong rationalizability.

Keywords: Dynamic Games, Asymmetric Information, Rationality, Robustness, Misspecifications, Forward and Backward Induction, Common Knowledge, Unique Selections

JEL Classification: C72, D82, D83

1 Introduction

In the theory of games with incomplete information, a central tool in Economics, the rationale behind strategic behavior can crucially hinge on "knife-edge" common knowledge assumptions,[1] so that *arbitrarily small* misspecifications of the latter can suffice to render predictions invalid—not infrequently, to the extent of overturning the economic conclusions

*November 13, 2022 (latest version). Contact: evan.piermont@rhul.ac.uk and p.zuazogarin@hse.ru

[1] This is deftly exhibited by Rubinstein's (1989) *Email game*: Consider a two-player static game with complete information, where: $E_0 \equiv$ "There are two strict Nash equilibria $a = (a_1, a_2)$ and $b = (b_1, b_2)$, and the former Pareto dominates the latter" holds. Notice that, if $E_1 \equiv$ "Player 1 believes with high enough probability that there is payoff-uncertainty and action b_2 is strictly dominant for Payer 2" held, then Player 1's unique possible best-reply would be b_1 (the *bad* equilibrium action); if $E_2 \equiv$ "Player 2 is certain of E_0 but also believes with high enough probability in E_1" held, then b_2 would be Player 2's unique possible best-reply (to b_1, what she would consider to be the unique possible best-reply of Player 1); if $E_3 \equiv$ "Player

of the model.[2] While this sensitivity calls into questions the solidity of general theoretical insights, it also provides an avenue towards equilibrium selection, allowing the selection of equilibria on the basis of their robustness to relaxations of common knowledge assumptions, for instance, *global games* techniques.

The question of which predictions remain pertinent when common knowledge assumptions are relaxed emerges naturally. There is an extensive literature in game theory and mechanism design addressing the impact of an player's misspecified probabilistic beliefs about individual preferences, modeled, with very limited exception, as her beliefs at the beginning of the strategic interaction.[3] In contrast, misspecifications of a players' persistent information, manifest in their *information-types*, have remained virtually unexplored.[4] By *persistent* information, we mean those restrictions on both the values of the payoff-relevant parameters of the environment, and the information and beliefs of others, that a player can deem as possible at *some* stage of a dynamic interaction, even after observing initially surprising events. Due to the subtlety of modeling such restrictions, it is unsurprising that specifications of information-types abound with stringent persistent informational assumptions that, often, pass unnoticed.

For example, in auction settings it is usual to model bidders' preferences as determined by a privately known parameter that is uncertain to the rest of participants. However, no bidder expects another bidder to be uncertain about her own preference and, moreover, the *set* of possible pieces of private information that each bidder can hold is commonly known. Similar observations apply to more general settings that allow for interdependent preferences: the *set* of possible profiles of relevant parameters (either privately known or residual) is commonly known, and with it, so are several other subtler informational aspects. For instance, whether asymmetric information can take place or not, or the fact that bidders share a common understanding of how belief-update would take place in a sequential scenario: if bid b_i could only be rationalized on the grounds of a specific subset of private valuation of bidder i, V_i', then a bidder j that observed b_i would infer that i's private valuation is contained in V_i', and update her beliefs accordingly—moreover, this would also be known in advance by i. Assumptions of this kind are not universally valid for the representation of real-life economic settings, and this naturally motivates an interest in better understanding the trade-off between enhanced tractability and aggravated

1 is certain of E_0 and of Player 2 being certain of E_0, but also believes with high enough probability in E_2" held, then b_1 would be Player 1's unique possible best-reply… That is, no matter how much common belief in the scenario in E_0 was approximated (i.e., which E_k held), the good equilibrium actions would never be played.

[2]The observation is eloquently synthesized by Chen, Tillio, Faingold and Xiong (2017, pp. 1424–1425), who recall some well-known examples of this phenomenon: the (in)dependence of private values and the possibility of full surplus extraction (Myerson, 1981, and Crémer and McLean, 1988), whether the probabilistic assessment of economic fundamentals is common or not and the possibility of commonly known mutually beneficial trade (Milgrom and Stokey, 1982, and Morris, 1994), and the certainty of acting as a proposer in bargaining settings and the immediacy of agreement (Rubinstein, 1982, and Yildiz, 2003).

[3]E.g., Dekel and Fudenberg (1990), Lipman (2003), Weinstein and Yildiz (2007), Dekel, Fudenberg and Morris (2006), Chen, Di Tillio, Faingold and Xiong (2010), Ely and Peski (2011), Penta (2012), Chen (2012), Chen, Tillio, Faingold and Xiong (2017), Heifetz and Kets (2018) or Germano, Weinstein and Zuazo-Garin (2020) for game theory, and Bergemann and Morris (2005, 2009, 2016), Oury and Tercieux (2012), Penta (2015) or Chen, Mueller-Frank and Pai (2020) for mechanism and information design.

[4]Exceptions include Penta (2012), on which we expand below, or Evsyukova, Innocenti and Lomys (2021).

missrepresentation that different specifications of information entail.

Prompted by the above, this paper approaches the problem of robustness from a new angle by accounting for *almost all* relaxations of informational assumptions. We introduce a universal-type space enabling for this flexibility and, within this framework, study the robustness of different dynamic variants of rationalizability at *interim* level—i.e., the upper-hemicontinuity of different correspondences mapping players' types to subsets of strategies. We find that disparities in the robustness of different solution concepts can be traced back to their reliance on observed behavior for updating beliefs after unexpected moves: while *weak* and *backward* rationalizability (the variants which may neglect unexpected moves) are robust, *strong* rationalizability (that which exploits them) is not.[5] A closer look reveals that strong rationalizability is *generically* robust—that is, upper-hemicontinuous in and open and dense subset of the universal type-space—and furthermore, that the conditions guaranteeing its robustness refer to absences of ties in payoffs and are, in consequence, economically relevant and easy to implement. Next, we show that if each players' information about others' information is always broad enough to permit making sense of every possible observed behavior (that is, is *rich*), then the predictions of weak, strong and backward rationalizability all coincide. We interpret this as an 'impossibility theorem' posing a dilemma between either restricting information in ad hoc fashion, or giving up on sharpening predictions via dynamic criteria. Finally, we address the validity of equilibrium selections based on robustness. We present a new dynamic version of Weinstein and Yildiz's (2007) 'structure theorem' showing that strongly rationalizable predictions characterize the set of weakly rationalizable predictions that can be uniquely selected by perturbing information and beliefs—in particular, it follows that strongly backward rationalizable predictions backward rationalizable ones. We thus conclude that selection arguments based on perturbations allow for nontrivial refinements of weak rationalizability.

Getting into further detail, our work builds on Penta's (2012) pioneering investigation of the interplay between information and higher-order uncertainty in dynamic games, and its impact on the robustness of game-theoretic predictions. Our departure from his work is threefold. First, we do *not* require players *not to have* information about other players' information. Second, we study robustness to *perturbations*, not *arbitrary changes*, of informational assumptions. Third, we do permit (but not require) information to be *wrong*.[6] To

[5] *Weak* rationalizability (c.f. Ben Porath, 1997, and Penta, 2012) captures the behavioral implication of sequential rationality and common *initial* belief thereof, so that it places no restrictions in beliefs at noninitial histories beyond update via conditional probability. *Backwards* rationalizability (c.f. Catonini and Penta, 2022) is a 'backward induction' version of rationalizability, so that erratic observed moves are rendered as uninformative 'trembles' and keeps common belief in *future* sequential rationality (Perea, 2014). *Strong* rationalizability (or *extensive-form rationalizability* in previous literature, e.g., Pearce, 1984) is a forward induction version of rationalizability that conjectures about future behavior by making sense of observed choices—via Battigalli's (1997) 'best rationalization principle' (also Battigalli and Siniscalchi, 2003).

[6] Penta (2012) focuses on payoff-states is of the form $\Theta = \Theta_0 \times \prod_{j \in I} \Theta_j$ so that, at each state $\theta \in \Theta$, each player i is privately informed of θ_i and thus considers every state in $\{\theta_i\} \times \Theta_0 \times \prod_{j \neq i} \Theta_j$ as possible (possibly, with probability 0). Thus, the information that player i has about other players' types is always the same, and independent of her information. While an extension to nonproduct sets of payoff-states is hinted at (p. 653), this is beyond the focus of that paper. Besides, the latter considers a notion of robustness to misspecifications of the set of states, *model invariance*, that requires predictions to be invariant to arbitrary changes of this set. Instead, we focus on perturbations (possibly at the higher-order) of the set of payoff-states—what we view as methodologically aligned with the approach to robustness to higher-order beliefs.

allow for these, our starting point is not an exogenously fixed *payoff-information structure* (roughly, a set of states, possibly also describing information-types) that implicitly entails numerous informational assumptions,[7] but instead, a novel type-space that encompasses every universal type-space built upon every possible payoff-information structure. This is achieved by, loosely, 'bundling together' payoff-information structures and type-spaces by introducing two minimal tweaks: (1) a set of payoff-states consisting in the set of *all possible* profiles of utility-functions from the set of outcomes, a primitive of the model, into the interval $[0, 1]$ (not an exogenously fixed set whose parameters later serve as arguments for a utility-function), and (2) considering types whose informational component consists in a subset of payoff-states *and* other players' types (not simply payoff-states).[8] This approach has two conceptual advantages. First, the persistent higher-order beliefs that a type holds via its informational component are fully encapsulated in the type, without further reference to any set of states, or payoff-information structure. Hence, the effects of relaxing informational assumptions can be studied in isolation, by simply focusing on perturbations of individual types' information. Second, the interim solution concepts that we study become *type-space invariant*; that is, they depend solely on the hierarchy of information and beliefs encoded in each type, not on the specific type-space that is employed for its representation. The analysis in our universal type-space is thus without loss of generality.

Within this framework, our examples in Section 2 document the negative—yet hardly surprising—observation that the predictions of strong rationalizability are *not* robust. On the positive side, our first main result (Theorem 1) shows that two standard solution concepts are robust in the whole universal type-space, namely weak and backward rationalizability. While the second part of the theorem is—to the best our knowledge—fully novel, the robustness of weak rationalizability reinforces previous insights by Dekel, Fudenberg and Morris (2006, Theorem 1) and Penta (2012, Proposition 1), who show, respectively, that weak rationalizability is robust in settings with no private information,[9] and to perturbation that do not affect types' information about others' types information. In principle, the disparate robustness properties of weak and backward rationalizability on the one hand, and strong rationalizability on the other, point towards a potentially more problematic phenomenon: that despite the conceptual appeal of the forward induction reasoning behind strong rationalizability, the conditions for it to be applied successfully may crucially rely on knife-edge assumptions that are tremendously demanding, and hard to meet in practice. Our next main result (Theorem 2) shows that this apparent intuition is misleading: strong rationalizability is robust in an open and dense subset of types in the universal type-space; thus, it is not the validity of forward induction what is knife-edge, but rather, the conditions

[7]See the previous footnote for the case of a commonly known set of payoff-state with product structure. Allowing for more general sets $\Theta \subseteq \Theta_0 \times \prod_{j \in I} \Theta_j$ entail less stringent common knowledge assumptions, but by virtue of Θ being commonly known, is not free of them.

[8]Technically, the novelty is minimal, and the construction of the universal type-space consists in a trivial combination of Mariotti, Meier and Piccione (2005) and Brandenburger and Dekel (1993). Conceptually, Bergemann and Morris (2016) argue that the description of a Bayesian game should separate the basic game (including a set of payoff-states) from the type-space. We propose to separate utility-functions and sets of states (including details about private information) from the basic game, and placing them with beliefs.

[9]And therefore, robust to misspecifications that do not affect information.

for its robustness to fail.

The intuition behind the contrasting nature displayed by the robustness of weak, backward and strong rationalizability encapsulates one of the main insights of the paper. To get a grasp of why strong rationalizability is not robust, consider a two-player dynamic game with perfect information, where: (A) player 1 (the first-mover) has two actions in the beginning of the game, a_1 and b_1, and the minimum utility that she can obtain with a_1 coincides with the maximum utility that she can obtain with b_1, and (B) after b_1, player 2 has two actions available, a_2 and b_2, and b_2 always yields a strictly lower utility than a_2, except if player 1 chooses a strictly dominated action after (b_1, b_2), in which case b_2 always yields a strictly higher utility than a_2. In this scenario, observing b_1 does not suffice for player 2 to conclude that player 1 is not rational: 1's beliefs could be such that a_1 and b_1 lead to the worst and best possible scenarios following them, respectively, so that according to (A), player 1 would be indifferent between both actions. Since b_1 does not falsify the rationality of player 1 in 2's eyes, following the best rationalization principle (that serves as the foundation of strong rationalizability), 2 should believe that 1 will be rational in her next move. This eliminates action b_2 for player 2: if player 1 is to be rational in her next move, according to (B), a_2 must necessarily be better for 2. Now, suppose that, instead of (A), player 2's information is consistent with: (A^n) the utilities that player 1 obtains after a_1 are the same as in (A) plus $1/n$ (for $n \in \mathbb{N}$). This time, no matter how large n is, player 2 cannot rationalize b_1 anymore—a rational player 1 would have necessarily chosen a_1. It is then legit (though not necessary) for player 2 to believe that player 1 may incur on a new mistake in the future, and this, in turn, may justify that 2 chose b_2. Thus, we find a violation of upper-hemicontinuity: b_2 is justifiable for a player 2 with information consistent with (A^n) (for every $n \in \mathbb{N}$) but not with information consistent with (A). But this failure of upper-hemicontinuity is simply a manifestation of a failure of *lower-hemicontinuity*—not of behavior, but of the set of histories in which restrictions on beliefs are placed: b_1 is rationalizable for a player 2 with information consistent with (A) but not with information consistent with (A^n) (for every $n \in \mathbb{N}$). As we argue next, this observation has general validity, and makes the intuition behind Theorems 1 and 2, trivial.

The first theorem is a conclusion of the fact that both weak and backward rationalizability place exactly the same restrictions at every history.[10] Since the set of histories in which these restrictions are placed is *constant* on the type's information, the lower-hemicontinuity problem sketched above cannot arise. In consequence, neither can failures of upper-hemicontinuity of behavior. For the intuition behind Theorem 2, notice the crucial role that the *tie* specified in (A) plays in the failure of robustness of strong rationalizability. If player 2's information was broad enough as to include states in which this tie can be broken in either direction, then the failure of lower-hemicontinuity of the histories in which 2 can rationalize choice b_1 would not arise. Thus, the intuition behind the genericity in Theorem 2 can be found in the fact that any *arbitrarily small* 'inflation' of a type's information suffices for this information to be broad enough in the aforementioned sense. These

[10] Each their own: weak rationalizability, certain assumptions in the beginning of the game, and *none* afterwards; backward rationalizability, that certain assumptions hold in the *continuation* play.

ideas provide the first main conclusion of the paper, and can be summarized as:

Main finding 1. *Weakly and backward rationalizable predictions are robust to misspecifications of informational assumptions; strongly rationalizable ones are not. This a is manifestation of failures of lower-hemicontinuity in the histories in which forward induction reasoning places constraints over beliefs. However, strongly rationalizable predictions are* generically *robust.*

We next restrict the informational setting under consideration, to allow for a better comparison with Chen (2012) and Penta (2012)'s findings. The extensions to dynamic settings of Weinstein and Yildiz's (2007) seminal structure theorem in these papers can be cast in the following terms: no strict refinement of weak rationalizability is robust in environments where: (C1) there are no information asymmetries about other players' information and the residual payoff-relevant component (i.e., the payoff-information structure is based on a *product* set of payoff-states), and (C2) each player's information about others' information is always broad enough to permit making sense of every possible behavior that she observes (the set of payoff-states is *rich*). This should already raise some questions. On the one hand, we have seen above that backward rationalizability, a *refinement* of weak rationalizability, is robust. On the other, as we prove in Proposition 2, it is easy to see that strong rationalizability—another refinement of weak rationalizability—is robust to the perturbations that respect (C1).[11] Corollary 1 reconciles these seemingly contradictory facts in the only possible way: under richness assumptions à la (C1) and (C2), the predictions of weak, backward and strong rationalizability all coincide. Back to the example above, it is tempting to argue that the failure of upper-hemicontinuity of strong rationalizability sketched there is an artifact of an excessively demanding constraint (that player 2 is unable to envision a payoff-state that rationalizes b_1), and from there, to conclude that a good modeling practice should look to avoid such pathological phenomena by prescribing that types' informational components to be rich enough. Corollary 1 shows that such a modeling approach results on the refinements of weak rationalizability based on dynamic criteria losing their bite.[12] Hence the interpretation of Corollary 1 as an 'impossibility' result: it is unavoidable to either restrict the information that enables for rationalization of observed choices in ad hoc fashion, or to give up on refinements based on dynamic criteria. That is:

Main finding 2. *If there are no asymmetries of information about others' information and the residual payoff-relevant component, and each player's information about others' information is always broad enough to permit making sense of every possible behavior that she observes, then the dynamic refinement criteria presupposed by backward and forward induction reasoning have no bite, and fully coincide with those obtained in mainly ex ante terms (that is, solely alluding to the interim normal-form of the game).*

Finally, we study the validity of the equilibrium selection criteria based on robustness to misspecifications of higher-order information and beliefs. Loosely speaking, the idea behind

[11] If information about others' information is constant, no lower-hemicontinuity issues appear.

[12] A weaker version of this result can be found in Battigalli and Siniscalchi (2007, Proposition 5).

these techniques (which include *global games*) is that, in settings with multiple equilibria, *small noise* at the higher-orders can suffice to create a contagion argument that result on the emergence of endogenous coordination in some equilibrium.[13] Weinstein and Yildiz's (2007) structure theorem raises a critique to this paradigm: in static games, and for any profile of types, every weakly rationalizable action profile can be *uniquely* selected by some perturbation of the type profile. In consequence, selecting an equilibrium on the basis of some perturbation is, essentially, as ad hoc as selecting said equilibrium without further argument (as all equilibria admit an analogous justifying perturbation).[14] The works by Chen (2012) and Penta (2012) extend this insight to dynamic games, but reveal significant caveats that add a critical twist on the message. Chen (2012) requires players not to hold any private information, and Penta (2012) relaxes this by simply requiring players not to have information about other players' information. The fact that these results rely on strong informational assumptions suggests that the presence of information, or persistent beliefs, may mitigate the severity of the Weinstein and Yildiz critique in dynamic settings.

Theorem 3 shows that there is some truth to this intuition. There, we show that, for *almost* every profile of finite types t, and any outcome z that is strongly rationalizable for t, there exists some perturbation of t along which weak rationalizability *uniquely* selects outcome z. Some minimal informational assumption on the type profile is required though (hence the 'almost'); namely, that there exists some payoff-state that is consistent with the information that each type has. Note that this 'consistency' requirement holds by construction in virtually every model employed in economic theory (following from the informational component of types being assumed *correct*) and, furthermore, that no richness assumptions are imposed.[15] Moreover, we document via a counterexample that an analogous selection is not always possible for backward (and thus weak) rationalizable predictions. Our theorem not only covers the results by Weinstein and Yildiz (2007), Chen (2012) and Penta (2012) as special cases, but (together with the counterexample) also clarifies the extent to which the Weinstein and Yildiz critique applies to dynamic games. On the one hand, only some predictions of weak rationalizability (including those of strong rationalizability) are signified, in the sense of admitting unique selections. On the other, the critique remains pertinent if restricted to equilibrium outcomes that are consistent with strong rationalizability. In particular, together with part 2 of Theorem 1, our structure theorem automatically implies that backward rationalizable predictions contain those by strong rationalizability for consistent type profiles. In summary:

Main finding 3. *Strongly rationalizable outcomes characterize the predictions of weak rationalizability that can be uniquely selected via a small noise in information and beliefs—regardless of informational assumptions, including richness ones. Hence, Weinstein and Yildiz's (2007) critique extends to dynamic games if the focus is on strongly rationalizable predictions, but not otherwise.*

The rest of the paper is structured as follows. Section 2 presents the approach to

[13]Once again, the logic follows that sketched in the explanation of Rubinstein's (1989) *Email game*.
[14]Morris, Shin and Yildiz (2016) eloquently argue against taking this view too seriously.
[15]On the benchmark type profile—richness assumptions do play a crucial role in the perturbation.

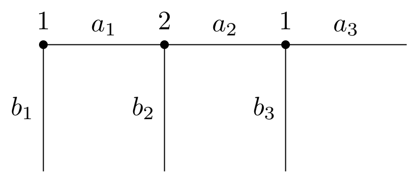
uncertainty and the main results in a simplified way, with a focus on examples. Section 3 introduces the model of payoff-uncertainty and the game-theoretic preliminaries that our analysis relies on. Section 4 presents the main findings of the paper. Section 5 ends with some literature review and additional discussion. All proofs are relegated to the appendix.

2 Informal example

In this section we provide a (rather) nontechnical example aimed at providing the main intuitions behind the more formal presentation later found in Section 3. We begin with a description of the game-theoretical settings covered by the paper. Next, we jump to a succinct overview of the two ingredients of our model of payoff-uncertainty: the use of profiles of utility-functions as states, and the incorporation in types of an informational component that puts *persistent* restrictions on players' higher-order beliefs about the payoff-state. Finally, we present an example illustrating the differences in the robustness to informational assumptions that different variants of dynamic reasoning display.

2.1 Extensive-forms and payoff-states

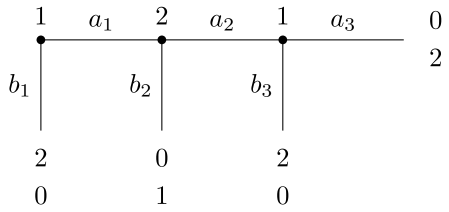
We focus on situations of sequential choice consisting of finitely many stages. At each stage, some players (possibly one), already informed about all the choices in previous stages, choose actions simultaneously. Previous sequence of choices determines which players are active at each stage, and which are the actions available to each of these players. The following *extensive-form* provides an illustration of such an scenario:



Here, players $i \in I = \{1, 2\}$ choose alternatively. At the *initial* stage, player 1 chooses between a_1 and b_2. If she chooses b_1 the game is over; otherwise, the game advances to its second stage and player 2 is called into action. At this second stage, player 2 (who knows that 1 chose a_1 at the first stage) chooses between a_2 and b_2. If she chooses b_2, the is game over; otherwise, the game advances to its final stage, where player 1 (who knows that 2 chose a_2 at the first stage) chooses between a_3 and b_3 and terminates the game. Thus, the set of possible terminal histories of choices, or *outcomes*, is $Z = \{b_1, (a_1, b_2), (a_1, a_2, a_3), (a_1, a_2, b_3)\}$. Obviously, situations where each stage involves more than one player, with more than two actions available to each of them, are allowed for.

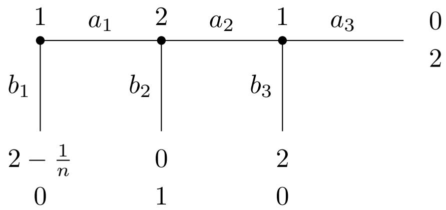
Players have von Neumann-Morgenstern preferences over outcomes but face uncertainty about each other's preferences (possibly, even about their own). This form of *payoff-uncertainty* is formalized by representing the possible profile of utility-functions over outcomes that players may have as *payoff-states*. By standard expected utility theory, it is w.l.o.g. to assume that players' utilities take values in $[0, 1]$. Accordingly, we let

$\Theta^* = \prod_{i \in I} [0, 3]^Z$ denote the set of all possible payoff-states given players I and outcomes Z.[16] For the extensive-form above we may have, for example, payoff state θ,

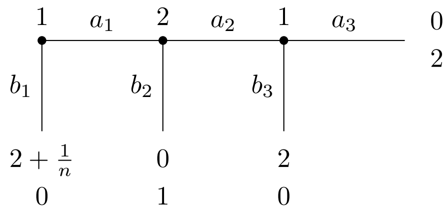


Payoff-state θ^0

or, for each $n \in \mathbb{N}$, payoff-states θ^{-n} or θ^{+n},



Payoff-state θ^{-n}



Payoff-state θ^{+n}

so that, for example, the utilities that outcome b_1 yields to players 1 and 2 are $2 - 1/n$ and 0, respectively, at payoff-state θ^{-n}, and $2 + 1/n$ and 0, respectively, at payoff-state θ^{+n}.

2.2 Information and (other) beliefs

In the context above, each player's assessment of payoff-uncertainty is represented by a *type* consisting of two different components: *information* and *beliefs*. These elements can be understood in hierarchical terms, as usual:

(1) Player i's *first-order model* is a pair $M_i^1 = (\Delta_i^1, \tau_i^1)$ where Δ_i^1 is set of payoff-states (i's first-order information) and τ_i^1 is a probabilistic belief about payoff-states (i's first-order belief) with probabilities concentrated in Δ_i^1. The interpretation is the following. τ_i^1 represents the beliefs that player i holds at the beginning of the game, which are subject to evolve in response to the choices of others' that she observes throughout the interaction. The informational component Δ_i^1, on the contrary, is assumed to remain constant along the whole interaction, and collects *the only* states that player i can consider as possible as the game unfolds—so that no $\theta \notin \Delta_i^1$ can ever get positive probability for i. Thus, the information and belief components place *persistent* and *initial* restrictions, respectively: the latter fully describes the initial likelihood that player i assigns to each state, and the former establishes limits on the support of the updated beliefs.

[16] That is, a payoff-state $\theta \in \Theta^*$ is a profile $(\theta_i)_{i \in I}$ whose ith component, $\theta_i : Z \to [0, 3]$, describes the utility-function that represents player i's preferences over Z at state θ. The admittedly arbitrary choice of 3 as maximum bound is to simplify the exposition that follows; in Section 3 this is normalized to 1.

(2) Player i's *second-order model* is a pair $M_i^2 = (\Delta_i^2, \tau_i^2)$, where Δ_i^2 and τ_i^2 represent player i's *second-order* information and beliefs, respectively, as follows. Δ_i^2 is a collection of pairs (θ, M_{-i}^1) where θ is a payoff-state and M_{-i}^1 is a description of other players' first-order models. τ_i^2 is a probabilistic belief about payoff-states *and* other players' first-order models with probabilities concentrated in Δ_i^2. Similarly as before, τ_i^2 represents the beliefs that player i holds at the beginning of the game, which are subject to evolve in response to the choices of others' that she observes throughout the interaction. The informational component Δ_i^2 is again assumed to remain constant along the whole interaction, and collects *the only* pairs of states and others' first-order models that i can consider as possible as the game unfolds. Again, information and beliefs place *persistent* and *initial* restrictions on beliefs, respectively.

...

A natural iterative procedure gives rise to an infinite hierarchy of finite-order models $t_i = (M_i^1, M_i^2, \dots, M_i^k, \dots)$, which we refer to as *type* for player i.[17] This formulation allows for representing all kind of usual (and less usual) dynamic games, depending on the specific type-profile that we endows players with. For example, within the context of the extensive-form in the previous paragraph we can have:

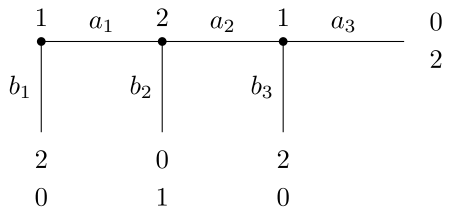
(i) *Payoff-uncertainty with symmetric information and a common prior.* Consider set of payoff-states $\Theta = \{\theta^{-n}, \theta^{+n}\}$ and probability measure q, where $q[\theta^{-n}] = q[\theta^{+n}] = 1/2$. Then, pick profile of types (t_1, t_2) where the first-order model of each t_i is $M_i^1 = (\Theta, q)$ and its higher-order models represent persistent and initial common belief in (M_1^1, M_2^1).[18] In this scenario, throughout the whole interaction, players will only assign positive probability to states in Θ and, while these probabilistic assessments may evolve as the game unfolds, each player will *initially* have a uniform assessment of the likelihood of the payoff-states in Θ.

(ii) *Complete information.* Suppose that each player i is endowed with type t_i that represents both persistent and initial common belief in payoff-state θ^0.[19] In such a case, no matter how the game unfolds, players will persistently believe that the following game depicts the situation:

[17] The usual higher-order *coherency* requirement between lower and higher-order models is assumed. Namely, for all $k \geq 1$, between Δ_i^k and a projection of Δ_i^{k+1}, and between τ_i^k and a marginal of τ_i^{k+1}.

[18] That is, type t_i's second-order model is $M_i^2 = (\Theta \times \{M_{-i}^1\}, \tau_i^2)$, where the marginal on Θ^* of τ_i^2 is q, and $\tau_i^2[\Theta \times \{M_{-i}^1\}] = 1$. This represents the facts that the only first-order model that player i persistently envisions as possible for the other player is M_{-i}^1, and that the latter is the only first-order model of the other that i can initially assigns positive probability to. The higher-order models that capture these ideas of persistent and initial common belief can be easily retrieved in iterative fashion.

[19] Similarly as before, player i's first-order model is $M_i^1 = (\{\theta^0\}, p)$ with $p[\theta^0] = 1$; her second-order model is $M_i^2 = (\{(\theta^0, M_{-i}^1)\}, p \times q)$, where $q[M_{-i}^1] = 1$; and so on.



(iii) *Incorrect asymmetric information.* The framework allows for representing situations in which players' persistent beliefs can greatly differ, to the extent that, as explained in the paragraph below, certain nonstandard phenomena manifest. For example, suppose that the following two hold: (A) for player 2, the dynamic game corresponds to the complete-information game parametrized by payoff-state θ^{-n}, and (B) the latter is known by player 1, for whom the true payoff-state is θ^{+n}. This situation can be formalized by: (A) endowing player 2 with type t_2^n with first-order information $\Delta_2^1 = \{\theta^{+n}\}$ and first-order beliefs τ_2^1 that put probability 1 on θ^{+n}, and whose higher-order models represent persistent common belief of said scenario, and (B) endowing player 1 with type t_1^n with first-order information $\Delta_1^1 = \{\theta^{-n}\}$ and first-order beliefs τ_1^1 that put probability 1 on θ^{-n}, and whose higher-order models represent persistent and initial higher-order belief in the scenario represented by t_2^n.[20]

It is important to highlight the differences between the vignettes described in (i) and (iii). In the former, both players may eventually envision θ^{-n} and θ^{+n} as possible; in the latter, player 2 will never consider θ^{+n} to be the true state, no matter what behavior by 1 she observes throughout the game (similarly, player 1 will never take θ^{-n} as the true state, but will be aware of the fact that, in 2's mind, θ^{-n} is the only state that matters).

2.3 Misspecificified information and strategic behavior

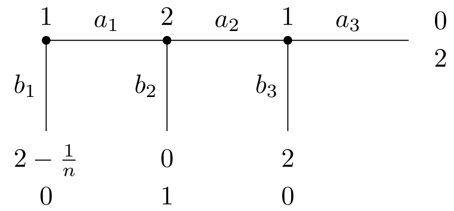
Consider again the game with complete information and utility-functions represented by payoff-state θ^0, already discussed in Example (ii) above. The prediction in this game seems self-evident: player 1 should choose b_1 in her first turn, regardless of whether she reasons in *backward* or *forward* inductive way:

- If players reason in backward inductive way (according to what we call, below, 'backward rationalizability'), then the logic is clear: in her second turn, where the decision does not involve any strategic uncertainty, player 1 would choose b_3; anticipating this, player 2 would choose b_2 in her turn and thus, anticipating this, player should decide to terminate the game at the first stage. Hence b_1.
- If players reason in forward inductive way (according to what we call, below, 'strong rationalizability'), trying to draw inferences about future behavior based on observations of previous behavior, the same conclusion is reached. First, a rational player 1

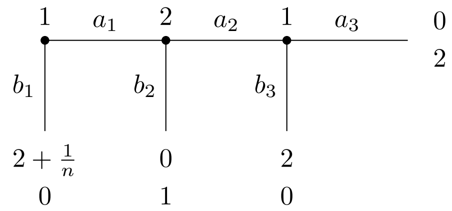
[20] As said, the first-order information that players 1 and 2 have is, respectively, $\Delta_1^1 = \{\theta^{-n}\}$ and $\Delta_2^1 = \{\theta^{+n}\}$ (the first-order beliefs are thus obvious); the second-order information that players 1 and 2 have, is respectively, $\Delta_1^2 = \{\theta^{-n}, \{\theta^{+n}\}\}$ and $\Delta_2^2 = \{\theta^{+n}, \{\theta^{+n}\}\}$; the third-order information that players 1 and 2 have is, respectively, $\Delta_1^3 = \{\theta^{-n}, \{\theta^{+n}\}, \{\theta^{+n}, \{\theta^{+n}\}\}\}$ and $\Delta_2^2 = \{\theta^{+n}, \{\theta^{+n}\}, \{\theta^{+n}, \{\theta^{+n}\}\}\}$; and so on.

would never choose a_3 in her second turn. Second, if player 2 had observed a_1, while she could be surprised that 1 took the risk of advancing instead of guaranteeing utility 2, she need not abandon the belief that 1 is rational: player 1 may have advanced, rationally, by believing that 2 does not expect her to be rational in her second turn, and thus believing that 2 will advance as well. In consequence, by keeping the belief that 1 is rational and will thus choose b_3 in her second turn, player 2 would choose b_2. Anticipating this, the first choice of player 1 is clear.

These conclusions are in contrast with what we could expect to happen in the scenario depicted in Example (*iii*), where players held different information (or persistent beliefs) about the payoff-state—for 2 the setting is as in complete-information game parametrized by θ^{+n}, and for player 1, who is aware of the previous fact, the true payoff-state is θ^{-n}:



Payoff-state θ^{-n}



Payoff-state θ^{+n}

In this case:

- Things do not seem to vary much if players reason in backward inductive way. At her second turn, player 1 would choose b_3 regardless of her beliefs about the payoff-state: in both cases it yields a higher utility than a_3. Anticipating this, Player 2, whose strategic considerations are fully focused on the game depicted in the right side, would choose b_2. Anticipating this, player 1, who considers the game in the left to represent the true situation, but is aware of the fact that in 2's mind it is the game in the right the one that counts, would assume that aiming for outcome (a_1, a_2, b_3) (which yields a higher utility than b_1) is unrealistic, and thus choose b_1.

- Now, if players reason in forward inductive way, new predictions arise. To see it, let us focus first on player 2. If called into action, player 2 *knows* that player 1 has chosen a_1. Now, given that, for her, the only relevant state is θ^{+n} (she totally omits θ^{-n} from her analysis of the situation), this choice of 1 is impossible to rationalize: by choosing a_1 player 1 has given up on utility $2 + 1/n$ knowing the maximum utility that she can obtain by doing so is 2. There are two possible ways in which player 2 can react to this: either (A) she believes that choosing a_1 must have been a mistake that is not indicative of future mistakes (as when reasoning in backward inductive way), or (B) she takes a_1 as indisputable evidence that 1 is likely to incur on erratic choices in her second turn. Thus, depending on how 2 interprets 1's choice, she will incline towards b_2 or a_2. The former in case of (A) for her, the latter in case of (B).

  Given the above, whether player 1 choose between a_1 or (2) will depend on how she expects 2 to interpret a_1. If player 1 expects a_1 to be interpreted as a mistake, she will opt for b_2 while if she expects it to be interpreted a sign of persistent erratic behavior, she will opt for a_1.[21] In consequence, we obtain three possible predictions: b_1, (a_1, b_2) and (a_1, a_2, b_3).

Notice that these conclusions are true for *every* $n \in \mathbb{N}$; that is, no matter how closely this second scenario approaches the one with complete information—notice that as n grows, the utility-functions specified by both θ^{-n} and θ^{+n} become arbitrarily close to those specified by θ^0 and hence arbitrarily close to complete information.

Thus, the example reveals a few phenomena—some of them nonstandard. First, that backward inductive predictions are robust to the particular misspecification of information described in the example: the perturbation of the benchmark scenario (complete information with utilities given by θ^0) does not give rise to new predictions (we always have only b_1); below we explore whether this observation generalizes or not. Second, that forward inductive prediction are *not* robust: any arbitrary small perturbation (arbitrarily large n) suffices for predictions omitted in the benchmark model to appear; below we study the extent to which this lack of robustness is critical or not. Finally, the examples suggests that the classic insight regarding the use of forward induction as a refinement argument for backward inductive predictions may be very fragile:[22] arbitrarily small perturbations suffice for forward inductive predictions not to refine backward inductive ones.

3 Framework

We now introduce the formal tools required to give some the necessary rigor to the ideas sketched above. In Section 3.1 we present our model of payoff-uncertainty, which serves as a primitive for the kind of games that we focus on: dynamic Bayesian games, formalized in Section 3.2. Once the framework is clear, in Section 3.3 we present the adaptation to the present setting of the different models of behavior, or solution concepts, whose robustness we later analyze.

3.1 Model of payoff-uncertainty

As mentioned above, our model of payoff-uncertainty is mostly standard, except for two minor novelties regarding the particular kind of payoff-states that we focus on, and the formalization of the possible information that players may hold—the latter does not only concern payoff-states but also other players' types. We first formally describe our type-spaces and briefly recall how to construct a universal type-space that allows for relaxations of common knowledge assumptions regarding the set of payoff-states and the typology of information. Next, we show how our modeling allows for said typology to be a feature of each type, and we end with some discussion.

[21] The strategy of *playing dumb* is a commonplace in popular culture; as Suetonius (121) writes about Emperor Claudius: "He did not even keep quiet about his own stupidity, but in certain brief speeches he declared that *he had purposely feigned it* under [Caligula], because otherwise he could not have escaped alive and attained his present station. But he convinced no one, and within a short time a book was published, the title of which was 'The Elevation of Fools' and its thesis, that no one feigned folly." The italics are ours.

[22] Or, more precisely, that strong rationalizability refines those backward rationalizability *outcomewise* (see Reny, 1992, Battigalli, 1996, Chen and Micali, 2013, or Perea, 2018a,c)

Type-spaces. The members of a finite set of *players* I have preferences over the elements of a finite set of *outcomes* Z. The preference of each player i can be represented by a (von Neumann-Morgenstern) *utility-function* $\theta_i : Z \to \mathbb{R}$, and players may face uncertainty about the profile of utility-functions $\theta \in \prod_{j \in I} \mathbb{R}^Z$, or *payoff-state*, that describes each other's utility-functions. The information and beliefs that players may hold about the payoff-state is represented by a *type-space*:[23]

Definition 1 (Type-spaces). *Let I and Z be finite sets of players and outcomes, respectively. Then, a* type-space *is a list $\mathscr{T} = (\Theta, (T_j, \Delta_j, \tau_j)_{j \in I})$, where $\Theta \subseteq \prod_{j \in I} \mathbb{R}^Z$ is a compact Polish set of payoff-states and, for each player i:*

1. *T_i is a compact Polish set of types.*
2. *$\Delta_i : T_i \rightrightarrows \Theta \times T_{-i}$, where $T_{-i} := \prod_{j \neq i} T_j$, is a continuous* possibility-correspondence *with nonemtpy and compact values.*
3. *$\tau_i : T_i \to \Delta(\Theta \times T_{-i})$ is a continuous* belief-map *with $\tau_i(t_i)\left[\Delta_i(t_i)\right] = 1$ for every t_i.*

That is, each type t_i describes a *model* $(\Delta_i(t_i), \tau_i(t_i))$ where $\Delta_i(t_i)$ is set of payoff-states and other players' types (t_i's *information*) and $\tau_i(t_i)$ is a probabilistic belief about payoff-states (t_i's *beliefs*) with probabilities concentrated in $\Delta_i(t_i)$. The interpretation is the following. $\tau_i(t_i)$ represents the probabilistic beliefs that type t_i initially holds, which are subject to evolve in response to how she may observe other players to choose. The informational component $\Delta_i(t_i)$, on the contrary, is assumed to always remain constant, and collects *the only* payoff-states and other players' types that t_i can consider as possible, regardless of how she observes others to choose—so that no $(\theta, t_{-i}) \notin \Delta_i(t_i)$ can ever get positive probability for i. Thus, the information and belief components place *persistent* and *initial* restrictions, respectively: the latter fully describes the initial likelihood that player i assigns to each payoff-state and others' types, and the former establishes limits on the support of updated beliefs.

Universal type-space. Given a type-space $\mathscr{T}$, each type t_i encodes a unique *hierarchy* of *models* $\mu_i(t_i) := (\mu_i^k(t_i))_{k \in \mathbb{N}}$. Type t_i's *first-order model*, describing information and beliefs about the payoff-state is $\mu_i^1(t_i) := (\Delta_i^1(t_i), \tau_i^1(t_i)) := (\mathrm{Proj}_\Theta \Delta_i(t_i), \mathrm{marg}_\Theta \tau_i(t_i))$. Type t_i's *second-order model*, describing joint information and beliefs about the payoff-state and others' first-order models, is $\mu_i^2(t_i) := (\Delta_i^2(t_i), \tau_i^2(t_i))$, where $\Delta_i^2(t_i) := \{(\theta, \mu_{-i}^1(t_{-i})) | \, (\theta, t_{-i}) \in \Delta_i(t_i)\}$ and, for every measurable $E \subseteq \Theta \times \prod_{j \neq i} \Delta(\Theta)$,

$$\tau_i^2(t_i)[E] := \tau_i(t_i) \left[\left\{ (\theta, t_{-i}) \in \Theta \times T_{-i} \,\middle|\, (\theta, \mu_{-i}^1(t_{-i})) \in E \right\} \right].$$

A standard recursive procedure give rise to the complete hierarchy $\mu_i(t_i)$. A straightforward combination of Mariotti, Meier and Piccione (2005) and Brandenburger and Dekel (1993)

[23] Throughout, finite sets and reals are endowed with the discrete and usual topology, respectively. Products of topological spaces are endowed with the product topology, and the space of compact subsets of a topological space is endowed with the Hausdorff metric. For a topological space X the set of probability measures on its Borel σ-algebra is denoted by $\Delta(X)$ and equipped with the topology of weak convergence.

establishes the existence of a *universal* type-space $\mathscr{T}^* = (\Theta^*, (T_j^*, \Delta_j^*, \tau_j^*)_{j\in I})$ where $\Theta^* := \prod_{j\in I}[0,1]^Z$, each T_i^* consists in the collection of all possible hierarchies of models for player i, and the usual *universality* properties à la Mertens and Zamir (1985) hold. First, $\mathscr{T}^*$ encapsulates all the possible information and beliefs about the payoff-state and others' hierarchies of models that a player may hold in the form of a specific hierarchy of models. Second, the set of types of player i of every type-space can be envisioned as a subset of T_i^* in a way that information and beliefs remain qualitatively invariant.[24]

Typologies of information. Within a type-space $\mathscr{T}$, each type t_i is interpreted as only considering as possible states $\theta \in \mathrm{Proj}_\Theta \Delta_i(t_i)$ and systematically excluding the rest. On top of this, it is a frequent in economic modeling to assume that players commonly believe that each type considers as possible *all* the types of others' whose information is *consistent* with hers—and only those.[25] To capture this, we say that a type-space $\mathscr{T}$ is *information-based* if, for every player i and every type t_i:

$$\Delta_i(t_i) = \left\{ (\theta, t_{-i}) \in \Theta \times T_{-i} \;\middle|\; \theta \in \bigcap_{j\in I} \mathrm{Proj}_\Theta \Delta_j(t_j) \right\}.$$

Each type's private information is often interpreted as *objectively* correct, so that a profile of types t is *consistent* if $\bigcap_{j\in I} \mathrm{Proj}_\Theta \Delta_j(t_j) \neq \emptyset$. The following cases have received particular attention in economics:[26]

- *Basic information.* Settings in which the information about the payoff-states is completely uniformative about others players' types, or more formally, where $\Delta_i(t_i) = \mathrm{Proj}_\Theta \Delta_i(t_i) \times T_{-i}$ for every player i and every type t_i.
- *No information.* Settings in which it is common knowledge that players hold the same information about the payoff-state and have no additional information about others players' types. That is, type-spaces where $\Delta_i(t_i) = \Theta \times T_{-i}$ for every player i and every type t_i.
- *Private values.* Setting in which players always *know* their own utility function but may be uncertain about others'; that is, type-spaces where, for every player i and type t_i, where elements in $(\mathrm{Proj}_\Theta \Delta_i(t_i))_i$ are positive transformations of each other.

[24]Section A.2 provides a more rigorous construction, but let us give some further detail also here. First, T_i^* is homeomorphic to the space of all models (K_i, P_i) where P_i is a compact subset of $\Theta^* \times T_{-i}^*$ and P_i is a probability measure on $\Theta^* \times T_{-i}^*$ with support contained in K_i. Thus, every possible model can be models can be encoded as a hierarchy of (finite-order) models. Second, for every type-space $\mathscr{T} = (\Theta, (T_j, \Delta_j, \tau_j)_{j\in I})$ there exist embeddings: (1) $\phi_0^{\mathscr{T}} : \Theta \to \Theta^*$ such that $(\phi_0(\theta))_i$ is a positive transformation of θ_i for every player i and every payoff-state θ, and (2) for every player i, $\phi_i^{\mathscr{T}} : T_i \to T_i^*$ such that, for every type t_i, we have $\Delta_i^*(\phi_i^{\mathscr{T}}(t_i)) = \{(\phi_0^{\mathscr{T}}(\theta), (\phi_j^{\mathscr{T}}(t_j))_{j\neq i}) | (\theta, t_{-i}) \in \Delta_i(t_i)\}$, and $\tau_i^*(\phi_i^{\mathscr{T}}(t_i))[E] = \tau_i(t_i)[\{(\theta, t_{-i}) \in \Theta_i \times T_{-i} | (\phi_0^{\mathscr{T}}(\theta), (\phi_j^{\mathscr{T}}(t_j))_{j\neq i}) \in E\}]$ for every measurable $E \subseteq \Theta \times T_{-i}^*$. Thus, every type-space can be regarded as a subspace of $\mathscr{T}^*$.

[25]E.g., in the literature in robust mechanism design—see Bergemann and Morris (2009, 2011), Penta (2012), Ollár and Penta (2017) or Müller (2016, 2020).

[26]The papers cited in the previous footnote all deal with the first case. For examples with trivial information, see Weinstein and Yildiz (2007) and Chen (2012) in the literature in robustness to higher-order beliefs, or Oury and Tercieux (2012) in mechanism design. For private values in mechanism design, see Heifetz and Neeman (2006) or Chen and Xiong (2013).

The modeling above allows for determining which is the particular typology of information that a canonical type $t_i \in T_i^*$ has, without the need to fully specify the particular type-space in which t_i is encoded. More precisely, we can think of which the *most stringent* typology of information that a type-space in which the canonical type can be encoded satisfies, so that a canonical type $t_i \in T_i^*$ is *information-based* if there exists some information-based type-space $\mathscr{T} = (\Theta, (T_j, \Delta_j, \tau_j)_{j \in I})$ such that $\phi_i^{\mathscr{T}}(t_i) = t_i^*$ for some type $t_i \in T_i$. Accordingly, t_i has *basic information* if there is some information-based type-space with basic information that encodes t_i, and has *no information* if, among these type-spaces, some has no information.

Discussion. Two features of the type-spaces introduced here are nonstandard. First, we envision payoff-states as profiles of utility-functions, not as exogenous parameters that, together with outcomes, are plugged into some general utility-function. Second, our possibility-correspondences map types to subsets of $\Theta \times T_{-i}$, not simply Θ. The reasons for both modeling choices are intertwined, but the main advantage lies on allowing for a canonical type-space that captures types with arbitrary typology of information and/or consistent with arbitrary common belief assumptions (or lack thereof) about said typology. On the one hand, working directly with profiles of utility-functions allows for dispensing with the common knowledge implicit in the mathematical structure of an exogenously set of parameters that usually serve as payoff-states. On the other hand, by leveraging on the persistent restrictions that our richer possibility-correspondence allows for, we can capture types that exhibit any of the aforementioned common knowledge assumptions. In consequence, the ability to capture all forms of beliefs about the typology of information emerges as particularly useful—if not strictly necessary—for the focus of the paper: the study of misspecifications of informational assumptions and its interplay with the impact of higher-order beliefs on strategic behavior on dynamic settings.[27]

Another important advantage, discussed in Section 3.3 below, is that the persistent restrictions on others' types permitted by our possibility-correspondences result on the *type-space invariance* of the variants of rationalizability that we study—that is, the predictions of our models of strategic behavior will only depend on the hierarchy associated to each type, not on the specific type-space that encodes the hierarchy. In consequences, the modeling allows for the formalization of strategic behavior being *robust* to the specific type-space chosen by the analyst to encode the hierarchy that captures exogenous restrictions on information and beliefs. In consequence, our focus on canonical types is fully justified.

3.2 Dynamic Bayesian games

A *dynamic Bayesian game* is a pair $(\mathscr{E}, \mathscr{T})$ where $\mathscr{E}$ is an *extensive-form* describing the players' possible sequences of choices and their information about previous moves, and $\mathscr{T}$ is a type-space whose set of payoff-states describes the possible profiles of players' utility-

[27] Section A provides details on the connection of our formalism with various standard tools in the literature. We show that *payoff-information structures* (e.g., Penta, 2012), Δ-*restrictions* with types (e.g., Battigalli and Siniscalchi, 2003), and settings with *context misalignment* (c.f., Guarino and Ziegler, 2022) all fall under the umbrella of the types here.

functions over over terminal sequences of choices. More specifically, we have:

- $\mathscr{E} = (I, (A_j)_{j\in I}, H, Z)$, where I is s finite set of *players*, and for each player i, A_i is a finite set of *actions*. H and Z are finite sets of *histories*, finite sequences of possibly simultaneous choices representing the possible ways in which the game may unfold. We say that history h' *follows* h if it obtains from the latter via the concatenation of finitely many possibly simultaneous choices, thus inducing a natural order on $H \cup Z$, assumed to be an oriented tree with root h^0—the *initial* history—and set of *terminal* nodes, or *outcomes*, Z. Histories in H are referred to as *partial*, and by H_i we denote the set of partial histories in which player i is *active*, i.e., those in which the set of actions *available* to i at h, $A_i(h)$, is nonempty.[28]

  Within $\mathscr{E}$, player i's set of *strategies* is defined as $S_i := \prod_{h\in H_i} A_i(h)$ and, as usual, we denote the set of i's opponents strategies by $S_{-i} := \prod_{j\neq i} S_j$. The set of strategy *profiles* is $S := \prod_{j\in I} S_j$ and, beginning at a partial history h, each profile s induces a unique *conditional* outcome $z(s|h)$. For each partial history h, we let $S_i(h)$ denote the set of player i's strategies that reach h, and write $S_{-i}(h) = \prod_{j\neq i} S_j(h)$. For each strategy s_i, $H_i(s_i)$ denotes the set of i's histories that can be reached by s_i.[29]

- $\mathscr{T} = (\Theta, (T_j, \Delta_j, \tau_j)_{j\in I})$, where, as discussed in the previous section, $\Theta \subseteq \prod_{j\in I} \mathbb{R}^Z$ is a set of payoff-states and, for each player i, $\theta_i : Z \to \mathbb{R}$ describes the *utility-function* over outcomes that corresponds to i at each state θ.

The focus of our analysis is on the robustness of different models of individual behavior at interim level, that is, in *interim solution concepts* $\mathcal{I}_i : T_i \rightrightarrows S_i$ that describe the dependence of the strategies that a player may choose with her assessment of the payoff-uncertainty. The notion of robustness that we study is conservative: small misspecifications of a player's type should not give rise to new predictions. Formally, an interim solution concept $\mathcal{I}_i : T_i \rightrightarrows S_i$ is *robust* if it is an upper-hemicontinuous correspondence.[30]

3.3 Solution concepts

We study three variants of rationalizability for dynamic games: *weak*, *strong* and *backward* rationalizability. All these interim solution concepts characterize the behavioral implications of rational players who engage on different versions of strategic reasoning. We first recall the usual notion of sequential rationality, and then present the restrictions on conjectures that allow for formalizing the variants of strategic reasoning that we focus on.

(Sequential) Rationality. In a dynamic Bayesian game each player i is uncertain about the payoff-state and about other players' strategies and types, and we represent the

[28]Formally, a history consists in a finite sequence with components in $\{h^0\} \cup A$, where $A := \bigcup_{J\subseteq I} A_J$ and $A_J := \prod_{i\in J} A_i$ for each nonempty $J \subseteq I$. h follows h' if there is some $(a^n)_{n\leq N}$ with components in A such that $h' = (h, a^1, a^2, \dots, a^N)$. We focus on *multistage-games with observable actions* (e.g., Fudenberg and Tirole, 1991) but the analysis extends to extensive-forms with *ambiguous* orderings of information sets (e.g., Topolyan, 2020).

[29]That is, $S_i(h) = \{s_i \in S_i | h \preceq z(s_{-i}; s_i|h^0)$ for some $s_{-i} \in S_{-i}\}$ and $S_{-i}(h) = \prod_{j\neq i} S_j(h)$ on the one hand, and $H_i(s_i) := \{h \in H_i | s_i \in S_i(h)\}$ on the other.

[30]I.e., for any sequence $(t_i^n)_{n\in\mathbb{N}}$ with limit t_i, and any $s_i \in \bigcap_{n\in\mathbb{N}} \mathcal{I}_i(t_i^n)$, it must be the case that $s_i \in \mathcal{I}_i(t_i)$.

beliefs that i holds at the different stages of the game via a *conjecture* $\mu_i : H_i \cup \{h^0\} \to \Delta(S_{-i} \times \Theta \times T_{-i})$ satisfying that: (C1) at each $h \in H_i \cup \{h^0\}$ player i *knows* that h has been reached and, (C2) whenever possible, i updates her beliefs via conditional probability:

(C1) $\mu_i(h)[S_{-i}(h) \times \Theta \times T_{-i})] = 1$.

(C2) For any $h' \in H_i$ such that $S_{-i}(h') \subseteq S_{-i}(h)$ and $\mu_i(h)[S_{-i}(h') \times \Theta \times T_{-i}] > 0$, and every measurable $E \subseteq S_{-i} \times \Theta \times T_{-i}$,

$$\mu_i(h')[E] = \frac{\mu_i(h')[E \cap (S_{-i}(h) \times \Theta \times T_{-i})]}{\mu_i(h)[S_{-i}(h') \times \Theta \times T_{-i}]}.$$

A conjecture μ_i induces a *conditional expected utility* for each strategy s_i at every history $h \in H_i \cup \{h^0\}$. We can thus recall the usual notion of (sequential) rationality, according to which s_i is a (*sequential*) *best-reply* for μ_i if it maximizes i's conditional expected utility at every history that it reaches. Hence, player i's set of best-replies to μ_i is:

$$r_i(\mu_i) := \left\{ s_i \in S_i \;\middle|\; s_i \in \bigcap_{h \in H_i(s_i)} \arg\max_{s'_i \in S_i} \int_{S_{-i} \times \Theta} \theta_i(z((s_{-i}, s'_i)|h)) \mathrm{d}(\mathrm{marg}_{S_{-i} \times \Theta} \mu_i(h)) \right\}.$$

Our analysis focuses on players already endowed with a type; thus, the conjectures that are consistent with the restrictions that a type t_i imposes over the belief $\Theta \times T_{-i}$ are of special interest. We say that conjecture μ_i is *consistent* with t_i if the probability of every $\mu_i(h)$ is concentrated on $S_{-i}(h) \times \Delta_i(t_i)$ and, at the initial history, $\mu_i(h^0)$ marginalizes to $\tau_i(t_i)$. Thus the aforementioned interpretation of $\Delta_i(t_i)$ as restrictions on beliefs that are *never* abandoned throughout the game, and $\tau_i(t_i)$, as initial beliefs that may evolve. Formally, μ_i is consistent with t_i if:

(C3) $\mu_i(h)[S_{-i}(h) \times \Delta_i(t_i)] = 1$ for every $h \in H_i \cup \{h^0\}$.

(C4) $\mathrm{marg}_{\Theta \times T_{-i}} \mu_i(h^0) = \tau_i(t_i)$.

We denote the set of player i's conjectures consistent with type t_i by $\mathrm{C}_i^{\mathscr{T}}(t_i)$.

VARIANTS OF STRATEGIC REASONING. The main interim solution concepts we analyze throughout the paper are based on rationality and assumptions on how players reason strategically—the main differences arising from the specific beliefs and higher-order beliefs about mutual rationality that players hold at each history of the game. It is thus convenient to recall first the constraints on beliefs that capture the solution concept below:

DEFINITION 2 (Initial, strong and future belief). *Let $(\mathscr{E}, \mathscr{T})$ be a dynamic Bayesian game and let $(\mathcal{I}_j)_{j \in I}$ be a profile of interim solution concepts. Then, for any player i and any type t_i we say that conjecture $\mu_i \in \mathrm{C}_i^{\mathscr{T}}(t_i)$ displays:*

1. Initial belief *in $\mathcal{I}_{-i}$ if, for history $h = h^0$,*

$$\mu_i(h)[M \times \Theta] = 1 \text{ for some measurable } M \subseteq \mathrm{Graph}(\mathcal{I}_{-i}).$$

2. Strong belief *in* $\mathcal{I}_{-i}$ *if, for every* $h \in H_i \cup \{h^0\}$ *s.t.* $S_{-i}(h) \times \Delta_i(t_i) \cap \mathrm{Graph}(\mathcal{I}_{-i}) \times \Theta \neq \emptyset$,

$$\mu_i(h)[M \times \Theta] = 1 \text{ for some measurable } M \subseteq \mathrm{Graph}(\mathcal{I}_{-i}).$$

3. Future belief *in* $\mathcal{I}_{-i}$ *if, for every* $h \in H_i \cup \{h^0\}$,

$$\mu_i(h)[M \times \Theta] = 1 \text{ for some measurable } M \subseteq \mathrm{Graph}(\mathcal{I}^h_{-i}),$$

where, for each player $j \neq i$ *and each type* $t_j \in T_j$,

$$\mathcal{I}^h_j(t_j) := \bigcup_{s_j \in \mathcal{I}_j(t_j)} \{s'_j \in S_j \mid s'_j(h') = s_j(h') \text{ for all } h' \in H_j \text{ s.t. } S_j(h') \subseteq S_j(h)\}.$$

That is, a conjecture μ_i (consistent with type t_i) that initially believes in $\mathcal{I}_{-i}$ represents a player i who, at the beginning of the game believes that other players' behavior is consistent with the predictions of $\mathcal{I}_{-i}$ but, upon reaching some *unexpected* history (i.e., one such that $\mu_i(h^0)[S_{-i}(h) \times \Theta \times T_{-i}] = 0$), may hold any arbitrary belief about others' play. Strong belief constrains belief updates at unexpected histories by requiring that, if reaching an unexpected h is consistent with $\mathcal{I}_{-i}(t_{-i})$ for some t_{-i} in $\Delta_i(t_i)$, then belief in $\mathcal{I}_{-i}$ is kept. Finally, future belief captures the usual *continuation consistency* requirement of backward induction procedures: even if the history reached is inconsistent with $\mathcal{I}_{-i}$, belief in *future* play following $\mathcal{I}_{-i}$ is always held. Applying each of these version of beliefs iteratively, we obtain three different dynamic variants of rationalizability:

Definition 3 (Weak rationalizability, c.f. Penta, 2012). *Let* $(\mathscr{E}, \mathscr{T})$ *be dynamic Bayesian game. The* weakly rationalizable *strategies of player* i *are given by* $\mathcal{W}^{\mathscr{T}}_i : T_i \rightrightarrows S_i$ *where, for each type* t_i *we have* $\mathcal{W}^{\mathscr{T}}_i(t_i) := \bigcap_{k \geq 0} \mathcal{W}^{\mathscr{T}}_{i,k}(t_i)$ *with* $\mathcal{W}^{\mathscr{T}}_{i,0}(t_i) := S_i$*, and, for every* $k \geq 0$,

$$\mathcal{W}^{\mathscr{T}}_{i,k+1}(t_i) := \left\{ s_i \in \mathcal{W}^{\mathscr{T}}_{i,k}(t_i) \;\middle|\; \begin{array}{ll} \multicolumn{2}{l}{\text{There exists some } \mu_i \in \mathrm{C}^{\mathscr{T}}_i(t_i) \text{ such that:}} \\ (1) & s_i \in r_i(\mu_i), \\ (2) & \mu_i \text{ displays initial belief in } \mathcal{W}^{\mathscr{T}}_{-i,k} \end{array} \right\}.$$

Definition 4 (Strong rationalizability, cf. Pearce, 1984; Battigalli, 1997). *Let* $(\mathscr{E}, \mathscr{T})$ *be dynamic Bayesian game. The* strongly rationalizable *strategies of player* i *are given by* $\mathcal{S}^{\mathscr{T}}_i : T_i \rightrightarrows S_i$ *where, for each type* t_i *we have* $\mathcal{S}^{\mathscr{T}}_i(t_i) := \bigcap_{k \geq 0} \mathcal{S}^{\mathscr{T}}_{i,k}(t_i)$ *with* $\mathcal{S}^{\mathscr{T}}_{i,0}(t_i) := S_i$*, and, for every* $k \geq 0$,

$$\mathcal{S}^{\mathscr{T}}_{i,k+1}(t_i) := \left\{ s_i \in \mathcal{S}^{\mathscr{T}}_{i,k}(t_i) \;\middle|\; \begin{array}{ll} \multicolumn{2}{l}{\text{There exists some } \mu_i \in \mathrm{C}^{\mathscr{T}}_i(t_i) \text{ such that:}} \\ (1) & s_i \in r_i(\mu_i), \\ (2) & \mu_i \text{ displays strong belief in } \mathcal{S}^{\mathscr{T}}_{-i,k} \end{array} \right\}.$$

DEFINITION 5 (Backward rationalizability, c.f. Catonini and Penta, 2022). *Let* $(\mathscr{E},\mathscr{T})$ *be dynamic Bayesian game. The* backward rationalizable *strategies of player* i *are given by* $\mathcal{B}_i^{\mathscr{T}} : T_i \rightrightarrows S_i$ *where, for each type* t_i *we have* $\mathcal{B}_i^{\mathscr{T}}(t_i) := \bigcap_{k\geq 0} \mathcal{B}_{i,k}^{\mathscr{T}}(t_i)$ *with* $\mathcal{B}_{i,0}^{\mathscr{T}}(t_i) := S_i$*, and, for every* $k \geq 0$*,*

$$\mathcal{B}_{i,k+1}^{\mathscr{T}}(t_i) := \left\{ s_i \in \mathcal{B}_{i,k}^{\mathscr{T}}(t_i) \;\middle|\; \begin{array}{ll} \multicolumn{2}{l}{\text{There exists some } \mu_i \in \mathrm{C}_i^{\mathscr{T}}(t_i) \text{ such that:}} \\ (1) & s_i \in r_i(\mu_i), \\ (2) & \mu_i \text{ displays future belief in } \mathcal{B}_{-i,k}^{\mathscr{T}} \end{array} \right\}.$$

This way, weak rationalizability captures the behavioral consequences of sequential rationality and common *initial* belief thereof, and is a straightforward adaptation to our setting of Penta's (2012) *interim sequential rationalizability*.[31] Strong rationalizability (or *extensive-form rationalizability*, as in Pearce, 1984, and Battigalli, 1997) refines weak rationalizability via a forward induction criterion: At each history h, player i is required to believe that others play following the $\mathcal{S}_{-i,k}^{\mathscr{T}}$ with the *highest* k that reaching h is consistent with. That is, players rationalize observed behavior to the highest possible degree.[32] Finally, backward rationalizability captures the idea of 'continuation consistency' (or, loosely speaking, subgame-perfection) usual in backward induction procedures.[33]

Notably, under the notion of type that we employ, all these solution concepts exhibit the robustness property known astype-space invariant, so that the strategies corresponding to each type only depend on the type's hierarchy of models (not on the specific type-space). This feature is specific to our notion of type, where each $\Delta_i(t_i)$ only considering types T_{-i} throughout the whole game and thus, embedding a type in a larger type-space does not alter the way in which beliefs can be updated.

PROPOSITION 1. *Let* $(\mathscr{E},\mathscr{T})$ *be a dynamic Bayesian game. Then for every player* i *and every type* t_i *the following three hold:*

1. $\mathcal{W}_i^{\mathscr{T}}(t_i) = \mathcal{W}_i^{\mathscr{T}^*}(\phi_i^{\mathscr{T}}(t_i))$.
2. $\mathcal{S}_i^{\mathscr{T}}(t_i) = \mathcal{S}_i^{\mathscr{T}^*}(\phi_i^{\mathscr{T}}(t_i))$.
3. $\mathcal{B}_i^{\mathscr{T}}(t_i) = \mathcal{B}_i^{\mathscr{T}^*}(\phi_i^{\mathscr{T}}(t_i))$.

Based on this, for the rest of the paper we drop superscript $\mathscr{T}^*$ when applying the solution concepts to dynamic Bayesian games where the type-space is universal.

---

[31]Which, in turn, is the *sequential* rationality counterpart to Dekel, Fudenberg and Morris's (2007) interim correlated rationalizability (ICR), employed by Chen (2012), and founded on *ex ante*, not sequential, rationality. See Ben Porath (1997) for a study of sequential rationality and common initial belief thereof in settings with complete information.

[32]I.e., Battigalli's (1996) *best rationalization principle* applies. See Battigalli and Siniscalchi (2002) for an analysis on the foundations of the solution concept, based on 'rationality and common *strong* belief therein'.

[33]The version of backward rationalizability here is a straightforward adaptation of Catonini and Penta's (2022) original notion. The epistemic foundations of the solution concept are studied by Perea (2014) and Battigalli and De Vito (2018), and founded on 'common belief in future rationality' in the former, and 'common full belief in optimal planning and in belief in continuation consistency' in the latter.

4 Results

We can now formally present the main results of the paper. In Section 4.1 we first identify and then compare the robustness properties of weak and backward rationalizability on the one hand, and strong rationalizability on the other. This exercise is revealing about the the extent to which the interplay between players' information and their ability to employ observed behavior to conjecture about future is critically sensitivity to modeling details. Next, in Section 4.2 we show that, if players are allowed to hold private information that is 'too broad', then usual game-theoretic refinement criteria based on dynamic considerations (e.g., subgame perfection) become innocuous and the resulting analysis, essentially static. Finally, Section 4.3 establishes that selection arguments based on robustness to perturbations of informational assumptions allow for discarding some weakly and backward rationalizable outcomes, but not for discriminating among strongly rationalizable outcomes.

4.1 Comparative robustness

Our first main result establishes that both weak and backward rationalizability are universally robust to misspecifications of informational assumptions—or, more formally, that $\mathcal{W}_i$ and $\mathcal{B}_i$ are upper-hemicontinuous correspondences in the whole universal type-space. The first part of the following theorem reinforces the message about the robustness of weak rationalizability in previous literature, namely, the extension to dynamic settings of Dekel, Fudenberg and Morris's (2006) result due to Penta (2012). The former pertains static settings—where informational assumptions are fully captures by beliefs and therefore play no role—, and the latter focuses on types-spaces where players hold no information about other players' types. In comparison, our result shows that the robustness of weak rationalizability holds even if players' information about other players' information varies across types, and even if these dependence is perturbed. The second part of the theorem further supports this positive message by showing that also the predictions of backward rationalizability satisfy these remarkable demanding robustness criteria:

Theorem 1. *Let $(\mathscr{E}, \mathscr{T}^*)$ be a dynamic Bayesian game. Then, for any player i the following two hold:*

1. *$\mathcal{W}_i : T_i^* \rightrightarrows S_i$ is upper-hemicontinuous.*

2. *$\mathcal{B}_i : T_i^* \rightrightarrows S_i$ is upper-hemicontinuous.*

These results are in contrast with the failure of robustness of strong rationalizability documented in the closing example of Section 2.3. Notably, the mechanics of that violation of upper-hemicontinuity greatly clarify the main intuition of why analogous problems do not arise in the case of weak and backward rationalizability. As discussed above, the constraints over beliefs placed by the forward inductive logic behind strong rationalizability are information-dependent: different restrictions apply at each history depending on whether the information that a player holds allows for making sense of observed behavior or not. Moreover, as a player may be able to rationalize observed behavior given some particular piece of private information of hers, but not as soon as this information is perturbed, it is easy to come up with examples (as the one in Section 2.3), the aforementioned dependence on information of is not lower-hemicontinuous. It was precisely this feature what triggered problems in the upper-hemicontinuity of strong rationalizability: the less demanding the constraints over the beliefs about others' rationality are, the wider the range of own behavior that can be sustained. It is then easy to see why weak and backward rationalizability are immune to this problem. Both solutions concepts place constant constraint over beliefs—in the case of weak rationalizability, *only* at the initial history and in the case of backward rationalizability, at *every* history. Thus, whether a constraint holds at a history or not is, in both cases, information-independent, and thus, the lower-hemicontinuity problem cannot arise. Neither can, in consequence, problems with upper-hemicontinuity.

It is then natural to inquire about the extent to which the the critical sensitivity to informational assumptions of strong rationalizability is general. Our next results unveils that the problem is, fortunately, remarkably marginal: while not universally robust, strong rationalizability is *generically* robust—it is an upper-hemicontinuous correspondence in an open and dense subset of the universal type-space. It turns out that the robustness of strong rationalizability goes beyond: the latters' predictions are universally robust when restricted to *information-based* types and perturbations (the usual ones in economic theory):[34]

Theorem 2. *Let $(\mathscr{E}, \mathscr{T}^*)$ be a dynamic Bayesian game. Then, for any player i the following three hold:*

1. *$\mathcal{S}_i : T_i^* \rightrightarrows S_i$ is generically upper-hemicontinuous.*
2. *$\mathcal{S}_{i|T_i^0} : T_i^0 \rightrightarrows S_i$ is upper-hemicontinuous.*

While the proof of the result requires to go through several technical steps, the main intuition behind is rather simple, and related to the crucial role that both the inconsistency in information and tie in Player 1's utilities player in the problematic example of Section 2.3. Notice that it is the absolute, persistent certainty of Player 2 on this tie what allows for perturbing her information so that she becomes convinced that choice a_1 of Player 1 is impossible to make sense of. Formally, this is operationalized by endowing Player 2 with a type t_2 such that, for every t_1 consistent with information $\Delta_2(t_2)$, $\Delta_1(t_1) = \{\theta^0\}$—this θ^0 being the payoff-state that includes the tie. Suppose instead that Player 2 was endowed with a type t_2^n with information $\Delta_2(t_2)$ that was consistent with types of Player 1 whose information about the payoff-state contained some (arbitrarily small) neighborhood of θ^0. In such case, *every* perturbation of t_2^n would have information consistent with types of Player 1 whose information was consistent with payoff-states where the tie in θ^0 may be broken in favor of a_1. In consequence, such types t_2^n would be immune to the lower-hemicontinuity problem with the set of histories in which strong rationalizability places constraints on beliefs. This way, part (1) of Theorem 2 captures the observation that, if strong rationalizability exhibits some upper-hemicontinuity problem at some type, it suffices with 'inflating'

[34]The set of player i's information-based types, T_i^0, collects all the canonical types of player i that are consistent with persistent common certainty of correct information about the payoff-state (see Section 3.1).

this type's information in arbitrarily small fashion. Part (2) guarantees that this procedure also works for types that are consistent with common certainty of correct information about payoff-states, which are, essentially, the only kind of types employed in economics. Now, while the second result may provide certain alleviation by suggesting that robustness problems only arise for types that are commonly neglected in practice, we consider that such an interpretation is overly optimistic. Whereas information-based types are convenient for for economic modeling (at least in order to avoid trivial explanation of phenomena based on severe inconsistencies among types), a genuine concern for robustness should account for severe misspecifications of the informational assumptions presupposed by the analyst—among them, of course, those related to common persistent belief in information about the payoff-state being correct. By excluding this possibility, the perturbations allowed for in part (2) of Theorem 2 pose, in our view, an excessively lenient robustness criterion.

4.2 Impossibility theorem

The robustness problem of strong rationalizability that the example in Section 2.3 illustrates raises an immediate objection about whether, by relying on very stringent informational assumptions, such examples have a questionable relevance. These informational assumptions required that Player 2 was unable to envision a payoff-state that would explain Player 1's behavior, and was instead comfortable with the conclusion that Player 1 may not be rational. It is thus natural to wonder whether sound modeling practice should require players' information about other players' information to be broad enough as to always deems as possible states that make sense of all possible observed behavior. Next, we explore the consequences of such a view.

To this end, we first introduce two formal definitions that adapt Weinstein and Yildiz (2007), Chen (2012) and Penta's (2012) *richness assumption* to our setting. Notably, we can leverage on the informational component in our definition of types to recast said assumptions as a property of individual types, not games:

Definition 6 (Conditional dominance). *Let $(\mathscr{E}, \mathscr{T})$ be a dynamic Bayesian game. Then, for any player i, any type t_i a strategy s_i is* conditionally dominant *for t_i if, for every strategy s_i' such that $s_i'(h') \neq s_i(h')$ for some history $h' \in H_i(s_i)$, there exists some history $h \in H_i$ such that, for every $s_{-i} \in S_{-i}(h)$ and every $\theta \in \mathrm{Proj}_\Theta \Delta_i(t_i)$,*

$$\theta_i(z((s_i, s_{-i})|h)) > \theta_i(z((s_i', s_{-i})|h)).$$

Definition 7 (Richness). *Let $(\mathscr{E}, \mathscr{T})$ be a dynamic Bayesian game. Then, we say that type t_i is* rich *if, for every $s_{-i} \in S_{-i}$, there exists some $t_{-i}^{s_{-i}} \in \mathrm{Proj}_{T_{-i}^*} \Delta_i(t_i)$ such that the following two hold:*

(1) *s_j is conditionally dominant for $t_j^{s_{-i}}$ for every $j \neq i$.*

(2) $(\mathrm{Proj}_\Theta(\Delta_i(t_i) \cap \Theta \times \{t_{-i}^{s_{-i}}\}))_i = [0,1]^Z$.

In words, a strategy is *conditionally dominant* for a type if, given the type's information has about the payoff-state, every behaviorally different strategy (that is, prescribing a different choice at some history that both strategies reach) yields unambiguously lower utilities after some history. Based on this, we say that a type is *rich* if, for every possible behavior of others', it envisions as possible some profile of types of theirs such that: (1) said behavior is conditionally dominant for the types in this profile, and (2) every possible utility-function of the player is consistent with this profile. While the second condition is of mainly technical nature, the first condition captures the essence of a type's information being 'broad' enough as to make sense of every possible behavior of others': no matter which history h player i reaches throughout the game, she can always envision some type t_{-i} of others for whom reaching h would have been reasonable.

Now, results by Penta (2012) and Chen (2012) show that in dynamic Bayesian games consisting of type-spaces where every type of every player is rich and *basic*,[35], then no strict refinement of weak rationalizability is robust. However, we saw in the previous paragraph that backward rationalizability is robust. Moreover, it is easy to see that for type-spaces consisting of rich and basic types strong rationalizability is robust too (if perturbations do not affect information about others' information, then the set of histories in which constraints on beliefs are placed cannot exhibit the lower-hemicontinuity problem explained above), a fact summarized in the following result:

Proposition 2. *Let $(\mathscr{E}, \mathscr{T})$ be a dynamic Bayesian game where every type t_i of every player i basic. Then, for every player i the correspondence $\mathcal{S}_i^{\mathscr{T}} : T_i \rightrightarrows S_i$ is upper-hemicontinuous.*

Obviously, backward and strong rationalizability *both refine* weak rationalizability. The conclusion is then immediate—under the informational assumptions in Chen (2012) and Penta (2012) all these solution concepts coincide:

Corollary 1. *Let $(\mathscr{E}, \mathscr{T}^*)$ be a dynamic Bayesian game. Then for rich and basic types, weak, strong, and backward rationalizability coincide; i.e., for any player i and any rich and basic type t_i:*

$$\mathcal{W}_i(t_i) = \mathcal{S}_i(t_i) = \mathcal{B}_i(t_i).$$

We thus reach meaningful conclusions about the limits of game-theoretic modeling in dynamic environments. Allowing for sufficiently rich information may preclude robustness problems, but only at the expense of trivializing the bite that accounting for the dynamic nature of the settings enables for.[36] Alternatively, this compelling extra bite is revealed to crucially rely on restricting, on somewhat ad hoc fashion, players' ability to rationalize observed behavior.

[35] A type of player i is basic if it is consistent with common persistent belief in the event that each player's types have the same information about other players' types (see Section 3.1).

[36] A similar insight can be found in Battigalli and Siniscalchi (2007) who, in our language, show that by *embedding* a dynamic game with complete information into one with payoff-uncertainty by expanding types' information about the (originally unique) payoff-state to allow for richness, the predictions of weak and strong rationalizability coincide. Garcia-Galocha, Jaromír Kovářík and Zuazo-Garin (2022) also replicate this result in the context of higher-order uncertainty about sequential rationality.

4.3 Selections via perturbations

4.3.1 Structure theorem

Our final main result sheds light on the refining power of equilibrium selection criteria based on robustness to misspecifications of players' higher-order information and beliefs. The main idea behind these techniques is based on creating contagion arguments that favor endogenous equilibrium coordination. As a sketchy illustration, consider a complete-information game with two symmetric players where (a, a) is one (of possibly many) Nash equilibrium, and introduce *diminishing* payoff-uncertainty as follows: there is a type t^1 of players for whom a is strictly dominant and thus, the only rationalizable choice for t^1; there is a type t^2 of players whose first-order beliefs are as in the complete-information benchmark case and who assigns probability 1 type t^1 of the other player, so that a is the only rationalizable choice for t^2; ... ; there is a type t^{k+1} of players whose kth-order beliefs are as in the complete-information benchmark case and who assigns probability 1 type t^k of the other player, so that a is the only rationalizable choice for t^{k+1}; ... . It seems then that perturbing arbitrarily high-order beliefs (even if leaving lower-order beliefs intact) can suffice for triggering a contagion that makes a uniquely rationalizable along the whole perturbation—regardless of how much the complete information benchmark case is approximated. This intuition has been operationalized with remarkable success by *global games*. These represent situations with payoff-uncertainty where, for type-profiles consistent with complete information the Nash equilibria are multiple and, by perturbing these types' higher-order beliefs via small asymmetries of information, *only one* equilibrium is selected. Weinstein and Yildiz (2007) raise a criticism to this approach by showing that, in static games with payoff-uncertainty, *every* outcome that is consistent with rationalizability in the interim normal-form of the game can be uniquely selected by an appropriate perturbation. In consequence, no rationalizable outcome—let alone equilibrium outcome—should deserve special preponderance on the basis of being possible to select by *some* perturbation of higher-order beliefs. Chen (2012) and Penta (2012) provide certain extensions of this insight to dynamic games. However, since their results only hold under rather stringent informational assumptions (that there are no information asymmetries about other players' information and the residual payoff-relevant component), the severity of Weinstein and Yildiz's (2007) critique to the refinement program in dynamic games remains unclear.

Before discussing the findings that address the problem above, let us be more precise and formally define first when do the predictions of a solution concepts admit to be selected via a perturbation:

Definition 8 (Unique selections via weak rationalizability). *Let $(\mathscr{E}, \mathscr{T}^*)$ be a dynamic Bayesian game, let t be a profile of types, and let $\mathcal{I} = (\mathcal{I}_i)_{i\in I}$ be a profile of interim solution concepts. Then, we say that the predictions of $\mathcal{I}$ for type-profile t* admit unique selections via weak rationalizability *if, for every $s \in \mathcal{I}(t)$ there exists a sequence of type-profiles $(t^n)_{n\in\mathbb{N}}$ converging to t such that, for every $n \in \mathbb{N}$ and every $s^n \in \mathcal{W}(t^n)$,*

$$z(s|h^0) = z(s^n|h^0)$$

In words, the predictions of a solution concept for a given specification of types can be *uniquely selected* by a perturbation when, for each z among these outcomes, it is possible to find a sequence that approximates the benchmark type-profile in a way that, for each component of the sequence, the *unique* prediction of weak rationalizability is z. Given this, and as explained below, the theorem and remark (based on the counterexample in the next paragraph) below provide answers to the scope of the Weinstein and Yildiz critique in dynamic games:

Theorem 3. *Let $(\mathscr{E}, \mathscr{T}^*)$ be a dynamic Bayesian game and let t be a consistent profile of finite, information-based types. Then, the predictions of strong rationalizability for t admit unique selections via weak rationalizability.*

Remark 1. *There exists some dynamic Bayesian game $(\mathscr{E}, \mathscr{T}^*)$ and a consistent profile of finite, information-based types t whose backward rationalizable predictions do not admit unique selections. In particular, neither do those of weak rationalizability.*

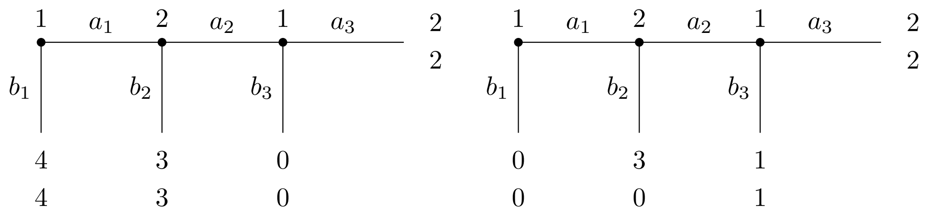
That is, in dynamic games, every strongly rationalizable outcome can be uniquely selected by some perturbation, (almost) regardless of the informational assumptions set by the analyst and even the absence of richness assumptions over types' information. To guarantee this flexibility of information, we leverage on the fact that perturbing types' informational components allows for approximating, for instance, settings with persistent common belief in no information using types that may exhibit some belief, possibly at the higher-orders, in the presence of private information. Furthermore, as we show via an example in the next paragraph, not all backward rationalizable outcomes can be uniquely selected this way. As a result, on the one hand, we show that informational assumptions do not constitute some barrier for the Weinstein and Yildiz critique: unique selections are possible (almost) in the absence of specific informational assumptions. On the other, this possibility excludes some backward rationalizable outcomes—and thus, also some weakly rationalizable ones. These observations provide two main insights into the open question about the Weinstein and Yildiz critique of refinements through perturbations. First, in settings where strongly rationalizable predictions refine weakly rationalizable ones and are consistent with some equilibrium outcome, selection criteria based on robustness to perturbations do have some bite, and this is regardless of informational assumptions. Second, these selection criteria remain subject to the Weinstein and Yildiz critique when applied to discriminating between equilibrium predictions consistent with strong rationalizability.

Some additional clarifications about Theorem 3 are in order. First, while our theorem does not require types to be rich, richness does play a crucial role in the perturbation, which is based on a standard contagion argument à la Email game. Now, the latter does not crucially compromise the generality of the former. If the profile of types to be perturbed was required to satisfy richness assumptions, then, by virtue of Corollary 1, the result would only be applicable to settings where dynamic refinements add no bite. Our result, on the contrary, is of broad scope: it applies to predictions within every dynamic game with finite extensive-form and, loosely speaking, irrespective of which type profile is chosen. Second, precisely because we do not require richness assumptions, our selection argument applies to *outcomes*, not profiles of strategies. This is due to the fact that, in the absence of richness, it is impossible to 'propagate' initial beliefs to *every* history of the game—reaching some histories in a strongly rationalizable way may be impossible and hence a zero probability event. Third, as mentioned above, our result allows for unique selections of *strongly* rationalizable outcomes, and we document via a counterexample that analogous selections are not always possible for backwards rationalizable (and hence weakly rationalizable) predictions. Finally, together with the upper-hemicontinuity of backward rationalizability, Theorem 1 provides a new proof of a well known result:[37]

Corollary 2 (c.f. Battigalli, 1996, Chen and Micali, 2013, Perea, 2018a). *Let $(\mathscr{E}, \mathscr{T}^*)$ be a dynamic Bayesian game and let t be a consistent profile of finite, information-based types. Then, the strongly rationalizable predictions given t are contained in the backward rationalizable predictions given t.*

4.3.2 Counterexample for backward rationalizability

The following example illustrates the claim above about not every backward rationalizable prediction admitting a perturbation of information that uniquely selects it. Consider the following dynamic game with utility-functions parametrized by payoff-states θ_1 and θ_2:



Payoff-state θ_1 Payoff-state θ_2

Suppose in addition that Player 1 always knows the payoff-state, that Player 2 never knows it, and that these two are common knowledge. The set of backward rationalizable strategies of Player 2 for arbitrary type t_2 is:

(A) $\{a_2\}$, if t_2 assigns positive probability to θ_2. At state θ_2 strategy $(a_1; a_3)$ is strictly dominant for player 1 (who, remember, knows that the state is θ_1). Thus, if player 2 observes that player 1 advances in her first round, she updates her beliefs by excluding the possibility of θ_1 being the true state. In consequence, a_2 becomes her only rational choice.

(B) $\{a_2, b_2\}$, if t_2 assigns null probability to θ_2. In this case it is unexpected for player 2 to observe player 1 advance, b_1 would have been strictly dominant if the state had been θ_1. Thus, player 2 needs to perform an update of her beliefs from scratch, what allows for the following two possibilities: Was choosing a_1 a mistake and, effectively,

[37]It is worth noting that the first proof relating strongly and backward rationalizable outcomes, due to Battigalli (1997), relied on *strategic stability* à la Kohlberg and Mertens (1986), a notions based on robustness to perturbations. While the perturbations that we consider are of different nature, and with an unclear connection with those employed for strategic stability, the analogy does not seem coincidental.

> the state is θ_1? Or where player 2's beliefs wrong and the state is θ_2, what justifies player 1's decision to advance? Both updating alternatives (and the corresponding mixed beliefs) are admissible according to backward rationalizable reasoning, and the corresponding only rational choices are b_2 and a_2, respectively (and one or either of them, if the updated is a mixed belief).

Pick now model $M := ((\Delta_1, \tau_1), (\Delta_2, \tau_2))$, where (i) $\mathrm{Proj}_{\Theta^*}(\Delta_i) = \{\theta_1, \theta_2\}$ and (ii) τ_1 and τ_2 represent initial common belief in states θ_2 and θ_1, respectively. We claim now that no perturbation of M can lead to b_2 being uniquely selected for player 2 and thus, to outcome (a_1, b_2) (which is backward rationalizable for M) being uniquely selected. To see it notice first that if a subset of states is close enough that of Δ_i then it can be written as the disjoint union of two sets of states, Θ^1 and Θ^2, such that it is strictly dominant for player 1 to choose b_1 at every $\theta'_1 \in \Theta^1$ and to choose $(a_1; a_3)$ at every $\theta'_2 \in \Theta^2$. Accordingly, for any perturbation of (Δ_2, t_2), the same argument as in (A) and (B) leads to the following conclusions: (A') if in the perturbed model Θ^2 gets positive probability then a_2 is player 2's unique backward rationalizable strategy, and (B') if in the perturbed model Θ^2 gets zero probability then both a_2 and b_2 are backward rationalizable for player 2. As a result, there is no perturbation of M in which a_2 is *not* backward rationalizable for player 2.

5 Related literature

Our approach bears some similarity with the notion of *unawareness* (see Fagin and Halpern, 1988; Modica and Rustichini, 1994), and in particular with the state-space or semantic approach to modeling awareness (see Dekel, Lipman and Rustichini, 1998; Heifetz, Meier and Schipper, 2006, 2008; Piermont, 2019), in which agent's are endowed with a coarse understanding of the true space of uncertainty. In particular, these approaches often model uncertainty via a partial order of increasingly expressive state-spaces: agent's residing in more expressive state-spaces can only reason about events in lower state-spaces. Agents who are *introspectively* unaware, however, might reason that there exist contingencies they are unaware of, without, of course, knowing exactly what such contingencies entail (see Halpern and Rêgo, 2009; Halpern and Piermont, 2019). The informational components of our type-spaces can capture both naive and introspective unawareness by changing restrictions on the relation between $\Delta_i(t_i)$ and the $\Delta_j(t_j)$ that each t_j consistent with $\Delta_i(t_i)$ holds: when the projection on Θ of $\Delta_i(t_i)$ is required to contain the one of $\Delta_j(t_j)$ then the agent i is naively unaware, as he does not consider it possible that j considers a contingency he does not. Note, most game-theoretic analyses of unawareness—Heifetz, Meier and Schipper (2014), Perea (2018b) and Guarino (2020), for instance—have focused on unawareness w.r.t. actions or strategies, not payoff states.

Relaxation of common knowledge of type-spaces has also been studied by Ziegler (2019) and Guarino and Ziegler (2022), and alternative misspecifications of models, by Esponda and Pouzo (2016) and, in the context of macroeconomic models, Hansen and Sargent (2001) and Cho and Kassa (2017). The study of disagreements about payoff-states and informational components follows the literature on how small changes in beliefs and information at high orders affects strategic behavior. Rubinstein's (1989) *email game* documents that behavior under common knowledge or under *almost* common knowledge can vary drastically. Later, Weinstein and Yildiz (2007) show that these discontinuities of behavior are not an isolated phenomenon, but rather, a pervasive feature of games with incomplete information. Penta (2012) and Chen (2012) extend this observation to dynamic games. Ely and Peski (2011) and Ruiz G. (2018) characterize the types in which these discontinuities arise in static and dynamic settings, respectively. Penta and Zuazo-Garin (2022) study the strategic impact of the discontinuities corresponding to higher-order uncertainty about the observability of choices, not preferences. Within the literature of robust mechanism design, Oury and Tercieux (2012) and Chen, Mueller-Frank and Pai (2020) study which social choice functions are implementable is a way robust to small misspecifications of higher-order beliefs. To this respect, the new notion of continuity in this paper and our results on the continuity of different solution concepts suggest novel questions and techniques for the study of implementability in dynamic mechanism design.

The counterintuitive nature behind the the stark discontinuities of behavior on information has sprung a literature arguing that these phenomenon is an artifact of very specific formalization or unrealistic assumptions on behavior. An approach consists in varying the notions of 'similarity' of belief hierarchies or 'approximation', as studied by Dekel, Fudenberg and Morris (2006), Chen, Di Tillio, Faingold and Xiong (2010; 2017) or Morris, Shin and Yildiz (2016). Another approach consists in showing that these discontinuities vanish with the introduction of bounded rationality and, specifically, under arbitrarily small departures from the benchmark of rationality and common belief thereof, as in Strzalecki (2014), Heifetz and Kets (2018), Germano, Weinstein and Zuazo-Garin (2020), Murayama (2020) or Jimenez-Gomez (2019). On the contrary, the literature on global games, starting from Carlsson and van Damme (1993), has embraced the discontinuities as an intrinsic feature of strategic behavior and leveraged on them to explain diverse economic phenomena such as currency crises (Morris and Shin, 1998), bank runs (Angeletos, Hellwig and Pavan, 2006; 2007), conflict (Baliga and Sjöström, 2012), overvaluation in financial markets (Han and Kyle, 2017) or disclosure policies for stress tests (Inostroza and Pavan, 2018).

Acknowledgments and disclaimers

Very preliminary versions of this project were occasionally presented under the titles "Heterogeneously Perceived Incentives in Dynamic Environments: Rationalization, Robustness and Unique Selections" and "Rationalization and robustness in dynamic games with incomplete information." Thanks are due to Pierpaolo Battigalli, Emiliano Catonini, Yi-Chun Chen, Pierfrancesco Guarino, Jaromír Kovářík, Avi Lichtig, Gabriel Ziegler and, especially, Antonio Penta, for insightful comments and valuable feedback, and to conference and seminar audiences at Bar-Ilan University, ESMES 2022, GAMES 2020, HSE University-International College of Economics and Finance, NUS, SAET 2020, TARK 2015, and Università Bocconi. Zuazo-Garin acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, from the Department of Education, Language Policy and Culture of the Basque Government (grants ECO2012-31346 and POS-2016-2-

0003 and IT568-13, respectively), from the ERC Programme (ERC Grant 579424) and from the Russian Academic Excellence Project '5-100'.

References

Angeletos, George-Marios, Christian Hellwig and Alessandro Pavan (2006). "Signaling in a global game: Coordination and policy traps". *Journal of Political Economy*, **114**, 452–484.

Angeletos, George-Marios, Christian Hellwig and Alessandro Pavan (2007). "Dynamic global games of regime change: Learning, multiplicity and timing of attacks". *Econometrica*, **75**, 711–756.

Baliga, Sandeep and Tjomas Sjöström (2012). "The strategy of manipulating conflict". *American Economic Review*, **102**, 2897–2922.

Battigalli, Pierpaolo (1996). "Strategic rationality ordorder and the best rationalization principle". *Games and Economic Behavior*, **13**, 178–200.

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

Battigalli, Pierpaolo and Nicodemo De Vito (2018). "Beliefs, plans, and perceived intentions in dynamic games". IGIER Working Paper Series #629.

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

Battigalli, Pierpaolo and Marciano Siniscalchi (2003). "Rationalization and incomplete information". *The B.E. Journal of Theoretical Economics*, **3**, 1–46.

Battigalli, Pierpaolo and Marciano Siniscalchi (2007). "Interactive Epistemology in Games with Payoff Uncertainty". *Research in Economics*, **61**, 165–184.

Ben Porath, Elchanan (1997). "Rationality, Nash equilibrium and backward induction in perfect information games". *The Review of Economic Studies*, **64**(1), 22–46.

Bergemann, Dirk and Stephen Morris (2005). "Robust mechanism design". *Econometrica*, **73**, 1771–1813.

Bergemann, Dirk and Stephen Morris (2009). "Robust Implementation in Direct Mechanisms". *The Review of Economic Studies*, **76**(4), 1175–1204.

Bergemann, Dirk and Stephen Morris (2011). "Correlated equilibrium in games with incomplete information". Cowles Foundation Discussion Paper No. 1822.

Bergemann, Dirk and Stephen Morris (2016). "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games". *Theoretical Economics*, **11**, 487–522. Mimeo.

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

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

Catonini, Emiliano and Antonio Penta (2022). "Backward Induction Reasoning Beyond Backward Induction". Mimeo. TSE Working Papers.

Chen, Jing and Silvio Micali (2013). "The order independence of iterated dominance in extensive games". *Theoretical Economics*, **8**, 25–63.

Chen, Yi-Chun (2012). "A structure theorem for rationalizability in the normal form of dynamic games". *Games and Economic Behavior*, **75**, 587–597.

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

Chen, Yi-Chun, Manuel Mueller-Frank and Mallesh M. Pai (2020). "Continuous implementation with direct revelation mechanisms". Mimeo.

Chen, Yi-Chun, Alfredo Di Tillio, Eduardo Faingold and Siyang Xiong (2017). "Characterizing the Strategic Impact of Misspecified beliefs". *The Review of Economic Studies*, **84**, 1424–1471.

Chen, Yi-Chun and Siyang Xiong (2013). "Genericty and Robustness of Full Surplus Extraction". *Econometrica*, **81**, 825–847.

Cho, In-Koo and Kenneth Kassa (2017). "Model averaging and persistent disagreement". *Federal Reserve Bank of St. Louis Review,*, **99**, 279–294.

Crémer, Jacques and Richard P. McLean (1988). "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions". *Econometrica*, **56**, 1247 1257.

Dekel, Eddie and Drew Fudenberg (1990). "Rational behavior with payoff uncertainty". *Journal of Economic Theory*, **52**, 243–267.

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.

Dekel, Eddie, Barton L. Lipman and Aldo Rustichini (1998). "Standard state-space models preclude unawareness". *Econometrica*, **66**, 159–173.

Ditor, Seymour Z. and Larry Q. Eifler (1972). 'Some Open Mapping Theorems for Measures. *Transactions of the American Mathematical Society*, **164**, 287–293.

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

Esponda, Ignacio and Demian Pouzo (2016). "Berk-Nash equilibrium: A framework for modeling agents with misspecified models". *Econometrica*, **84**, 1093–1130.

Evsyukova, Yulia, Federico Innocenti and Niccolò Lomys (2021). "Information Design with Optimal Frame Choice". Mimeo.

Fagin, Ronald and Joseph Y Halpern (1988). Belief, awareness, and limited reasoning. *Artificial intelligence*, **34**(1), 39–76.

Fudenberg, Drew and Jean Tirole (1991). *Game Theory*. MIT University Press. Cambridge.

Garcia-Galocha, Aleix, Jaromír Kovářík and Peio Zuazo-Garin (2022). Centipede Games: A Fragile Paradox. Mimeo.

Germano, Fabrizio, Jonathan Weinstein and Peio Zuazo-Garin (2020). "Uncertain rationality, depth of reasoning and robustness in games with incomplete information". *Theoretical Economics*, **15**, 89–122.

Guarino, Pierfrancesco (2020). "An epistemic analysis of dynamic games with unawareness". *Games and Economic Behavior*, **120**, 257–288.

Guarino, Pierfrancesco and Gabriel Ziegler (2022). "Strategic Behavior under Context Misalignment". Mimeo.

Halpern, Joseph Y and Evan Piermont (2019). Partial awareness. In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 2851–2858.

Halpern, Joseph Y. and Leandro C. Rêgo (2009). "Reasoning about knowledge of unawareness". *Games and Economic Behavior*, **67**, 503–525.

Han, Jungsuk and Albert S. Kyle (2017). "Speculative equilibrium with differences in higher-order beliefs. *Management Science*, **69**, 4317–4332.

Hansen, Lars Peter and Thomas J. Sargent (2001). "Robust control and model uncertainty". *American Economic Review*, **91**, 60–66.

Heifetz, Aviad and Willemien Kets (2018). "Robust multiplicity with a grain of naiveté. *Theoretical Economics*, **13**, 415–465.

Heifetz, Aviad, Martin Meier and Burkhard Schipper (2006). "Interactive unawareness". *Journal of Economic Theory*, **130**, 78–94.

Heifetz, Aviad, Martin Meier and Burkhard Schipper (2008). "A canonical model for interactive unawareness". *Games and Economic Behavior*, **62**, 304–324.

Heifetz, Aviad, Martin Meier and Burkhard Schipper (2014). "Dynamic unawareness and rationalizable behavior". *Games and Economic Behavior*, **81**, 50–68.

Heifetz, Aviad and Zvika Neeman (2006). "On the Generic (Im)Possibility of Full Surplus Extraction in Mechanism Design". *Econometrica*, **74**, 213–233.

Inostroza, Nicolas and Alessandro Pavan (2018). ""Persuasion in global games with application to stress testing". Mimeo.

Jimenez-Gomez, David (2019). "False consensus in games: Embedding level-k models into games of incomplete information". Mimeo.

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

Lipman, Barton (2003). "Finite Order Implications of Common Priors". *Econometrica*, **71**, 1255 1267.

Mariotti, Thomas, Martin Meier and Michele Piccione (2005). "Hierarchies of Beliefs for Compact Possibility Models". *Journal of Mathematical Economics*, **41**, 303–324.

Mertens, Jean-François 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.

Modica, Salvatore and Aldo Rustichini (1994). Awareness and partitional information structures. *Theory and decision*, **37**(1), 107–124.

Morris, Stepehen, Hyung Song Shin and Muhamet Yildiz (2016). "Common belief foundations of global games". *Journal of Economic Theory*, **163**, 826–848.

Morris, Stephen (1994). "Trade with Heterogeneous Prior Beliefs and Asymmetric Information". *Econometrica*, **62**, 1327 1347.

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

Müller, Christoph (2016). "Robust Virtual Implementation Under Common Strong Belief in Rationality". *Journal of Economic Theory*, **162**, 407–450.

Müller, Christoph (2020). "Robust Implementation in Weakly Perfect Bayesian Equilibrium". *Journal of Economic Theory*, **189**, 105–138.

Murayama, Kota (2020). "Robust predictions under finite depth of reasoning". *The Japanese Economic Review*, **71**, 59–84.

Myerson, Roger B. (1981). "Optimal Auction Design . *Mathematics of Operations Research*, **6**, 58 73.

Ollár, Mariann and Antonio Penta (2017). "Full Implementation and Belief Restrictions". *American Economic Review*, **107**, 2243–2277.

Oury, Marion and Olivier Tercieux (2012). "Continuous implementation". *Econometrica*, **80**, 1605–1637.

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

Penta, Antonio (2012). "Higher order uncertainty and information: Static and dynamic games". *Econometrica*, **80**, 631–660.

Penta, Antonio (2015). "Robust dynamic implementation". *Journal of Economic Theory*, **160**, 280–316.

Penta, Antonio and Peio Zuazo-Garin (2022). "Rationalizability, observability and common knolwedge. *Review of Economic Studies*, **89**, 948–975.

Perea, Andrés (2014). "Belief in the opponent's future rationality". *Games and Economic Behavior*, **81**, 235–254.

Perea, Andrés (2018a). "Why forward induction leads to the backward induction outcome: A new proof for Battigalli's theorem". *Games and Economic Behavior*, **110**, 120–138.

Perea, Andrés (2018b). "Common belief in rationality in games with unawareness". Mimeo (previous version, Epicenter Working Paper No. 13).

Perea, Andrés (2018c). "Order independence in dynamic games". Mimeo (previous version, Epicenter Working Paper No. 8).

Piermont, Evan (2019). "Algebraic semantics for propositional awareness logics". *arXiv preprint arXiv:1910.07275*.

Reny, Philipp J. (1992). "Backward induction, normal-form perfection and explicable equilibria". *Econometrica*, **60**, 627–649.

Rubinstein, Ariel (1982). "Perfect Equilibrium in a Bargaining Model". *Econometrica*, **50**, 97 109.

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

Ruiz G., David (2018). "Critical types in dynamic games". Mimeo.

Strzalecki, Tomasz (2014). "Depth of reasoning and higher order beliefs". *Journal of Economic Behavior and Organization*, **108**, 108–122. Mimeo.

Suetonius, Gaius Tranquillus (121). *The Twelve Caesars*. Pengüin Classics.

Topolyan, Irina (2020). "On Common Belief in Future Rationality in Games with Ambiguous Orderings ofInformation Sets". *Dynamic Games and Applications*, **10**, 183–201.

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.

Ziegler, Gabriel (2019). "Adversarial bilateral information design". Mimeo.

A Type-spaces

A.1 Preliminaries

Definition 9 (Type-space). *Let I be a finite set. Then, a* type-space *is a list* $\mathscr{T} = (\Theta_i, T_i, \Delta_i, \tau_i)_{i\in I}$ *where, for each $i \in I$, we have:*

1. *A compact Polish space Θ_i.*
2. *A compact Polish set T_i.*
3. *A continuous correspondence with nonempty and compact values $\Delta_i : T_i \rightrightarrows \Theta_i \times T_{-i}$, where $T_{-i} := \prod_{j\neq i} T_j$.*
4. *A continuous map $\tau_i : T_i \to \Delta(\Theta_i \times T_{-i})$ where $\tau_i(t_i)\,[\Delta_i(t_i)] = 1$ for every $t_i \in T_i$.*

Definition 10 (Completeness). *A* type-space $\mathscr{T} = (\Theta_i, T_i, \Delta_i, \tau_i)_{i\in I}$ *is* complete *if, for every $i \in I$, the following map is surjective:*

$$\begin{array}{rcl} T_i & \longrightarrow & \{(K_i, \beta_i) \in 2^{\Theta_i\times T_{-i}} \times \Delta(\Theta_i \times T_{-i}) |\, K_i \text{ is nonempty and compact, and } \beta_i[K_i] = 1\} \\ t_i & \mapsto & (\Delta_i(t_i), \tau_i(t_i)) \end{array}$$

Definition 11 (Terminality). *Let I be a finite set of players. Then,* type-space $\mathscr{T} = (\Theta_i, T_i, \Delta_i, \tau_i)_{i\in I}$ *is* terminal *if, for every type-space $\mathscr{T} = (\tilde\Theta_i, \tilde T_i, \tilde\Delta_i, \tilde\tau_i)_{i\in I}$ where $\tilde\Theta_i \subseteq \Theta_i$ for every player i, there exists, for every player i, a continuous map $\phi_i : \tilde T_i \to T_i$ such that, for every type $\tilde t_i \in \tilde T_i$ the following two hold:*

(1) $\Delta_i(\phi_i(\tilde t_i)) = \{(\tilde\theta_i, (\phi_j(\tilde t_j))_{j\neq i}) | (\tilde\theta_i, \tilde t_{-i}) \in \tilde\Delta_i(\tilde t_i)\}$

(2) $\tau_i(\phi_i(\tilde t_i))[E] = \tilde\tau_i(\tilde t_i)\{(\tilde\theta_i, (\tilde t_j)_{j\neq i}) \in \tilde\Theta_i \times \tilde T_{-i} | (\tilde\theta_i, (\phi_j(\tilde t_j))_{j\neq i}) \in E\}$ *for every measurable* $E \subseteq \Theta_i \times T_{-i}$.

A.2 Universal type-space

A.2.1 The space of all hierarchies of models

Fix a set of player I and, for each player i, a compact Polish set of basic uncertainty Θ_i. We will now construct a type-space:

$$\mathscr{T}^* = (\Theta_i, \Delta_i^*, \tau_i^*)_{i\in I}$$

that plays the role of a universal type-space in our analysis:

Set of types. Player i's set *hierarchies of models* is:

$$T_i^* := \left\{ (\mu_i^k)_{k\in\mathbb{N}} \,\middle|\, (\mu_i^1, \dots, \mu_i^k) \in M_i^k \text{ for every } k \in \mathbb{N} \right\},$$

where each M_i^k is player i's set of *kth-order models* and is defined inductively via the process that we detail next (Lemma 1 guarantees that the following steps are sound):

- First-order models. Set first $M_{-i}^0 := \Theta_i$ for each $i \in I$ and then, define the latter's set of *first-order models* as:

$$M_i^1 := \left\{ (\Delta_i^1, \tau_i^1) \in 2^{M_{-i}^0} \times \Delta(M_{-i}^0) \,\middle|\, \begin{array}{ll} (1) & \Delta_i^1 \neq \emptyset \text{ is compact} \\ (2) & \tau_i^1[\Delta_i^1] = 1 \end{array} \right\}.$$

Set $M^1_{-i} := \prod_{j \neq i} M^1_j$ and, for every $\mu^1_i = (\Delta^1_i, \tau^1_i) \in M^1_i$, define:

$$\Delta^1_i(\mu^1_i) := \Delta^1_i \ \text{ and } \ \tau^1_i(\mu^1_i) := \tau^1_i.$$

- *Higher-order models.* For each $k \geq 1$ and for each $i \in I$ define the latter's set of $(k+1)$*th-order models* as:

$$M^{k+1}_i := \left\{ (\mu^k_i, (\Delta^{k+1}_i, \tau^{k+1}_i)) \in M^k_i \times 2^{M^k_{-i}} \times \Delta(M^k_{-i}) \;\middle|\; \begin{array}{ll} (1) & \Delta^{k+1}_i \neq \emptyset \text{ is compact}, \\ (2) & \tau^{k+1}_i[\sigma^{k+1}_i] = 1, \\ (3) & \mathrm{Proj}_{M^k_i} \Delta^{k+1}_i = \Delta^k_i(\mu^k_i), \\ (4) & \mathrm{marg}_{M^k_i} \tau^{k+1}_i = \tau^k_i(\mu^k_i) \end{array} \right\}.$$

and set $M^{k+1}_{-i} := \prod_{j \neq i} M^{k+1}_j$ and, for every $\mu^{k+1}_i = (\Delta^{k+1}_i, \tau^{k+1}_i) \in M^{k+1}_i$, define:

$$\Delta^{k+1}_i(\mu^{k+1}_i) := \Delta^{k+1}_i \ \text{ and } \ \tau^{k+1}_i(\mu^{k+1}_i) := \tau^{k+1}_i.$$

Given T^*_i, we denote $T^*_{-i} := \prod_{j \neq i} T^*_j$ and define, for each hierarchy of of models $t_i = (\mu^k_i)_{k \in \mathbb{N}}$:

$$\Delta_i(t_i) := (\Delta^k_i(t_i))_{k \in \mathbb{N}} := (\Delta^k_i(\mu^k_i))_{k \in \mathbb{N}} \quad \text{and} \quad \tau_i(t_i) := (\tau^k_i(t_i))_{k \in \mathbb{N}} := (\tau^k_i(\mu^k_i))_{k \in \mathbb{N}}.$$

Lemma 1. *T^*_i is well-defined, and compact Polish.*

Proof. Let us introduce some auxiliary notation first. For each player i set $\hat{O}^1_i := \mathcal{K}(\Theta) \times \Delta(\Theta)$ and $\bar{M}^1_i := \hat{O}^1_i$, and then, define inductively:

$$\hat{O}^{k+1}_i := \mathcal{K}(\bar{M}^k_{-i}) \times \Delta(\bar{M}^k_{-i}) \quad \text{and} \quad \bar{M}^{k+1}_i := \bar{M}^k_i \times \hat{O}^{k+1}_i,$$

where $\bar{M}^k_{-i} = \prod_{j \neq i} \bar{M}^k_j$ for every $k \geq 1$. For every player i, $\hat{O}^k_i$ is compact Polish for every $k \in \mathbb{N}$. Set now, for each player i, $\hat{M}^0_i = \prod_{k \in \mathbb{N}} \hat{O}^k_i$, and, for every $k \in \mathbb{N}$,

$$\hat{M}^k_i := M^k_i \times \prod_{\ell > k} \hat{O}^\ell_i.$$

Notice that $\hat{M}^0_i$ is compact Polish and Hausdorff.

Then, we proceed by induction on $k \geq 0$. It is trivially true that, for $k = 0$, $\hat{M}^k_i$ is compact Polish for every player i, so pick $k \geq 0$ and let us check that if $\hat{M}^k_i$ is compact Polish for every i then so is $\hat{M}^{k+1}_i$. For every i, that $\hat{M}^k_i$ is compact Polish implies that M^k_i is compact Polish and thus, it follows that both $\mathcal{K}(M^k_{-i}) \times \Delta(M^k_{-i})$ and hence $M^k_i \times \mathcal{K}(M^k_{-i}) \times \Delta(M^k_{-i})$ are compact Polish as well. Since M^{k+1}_i is a closed subset of the latter, we conclude that it is compact Polish and, in consequence, that so is $\hat{M}^{k+1}_i$. Now, notice that we have both (a) $T^*_i = \bigcap_{k \in \mathbb{N}} \hat{M}^k_i$ and (b) $\hat{M}^k_i \subseteq \hat{M}^0_i$ for every $k \in \mathbb{N}$. Then, T^*_i is a countable intersection of compact and Polish subspaces of a Hausdorff space. Thus, T^*_i is compact Polish. ∎

Possibility-correspondence. Player i possibility-correspondence is $\Delta^*_i : T^*_i \rightrightarrows \Theta_i \times T^*_{-i}$, where:

$$\Delta^*_i(t_i) := \left\{ (\theta_i, t_{-i}) \in \Theta_i \times T^*_{-i} \;\middle|\; \theta_i \in \Delta^1_i(t_i) \text{ and } (\Delta^k_{-i}(t_{-i}), \tau^k_{-i}(t_{-i})) \in \Delta^{k+1}_i(t_i) \text{ for every } k \in \mathbb{N} \right\}.$$

Lemma 2. *Δ^*_i is surjective and continuous with nonempty and compact values.*

Proof. For each $k \in \mathbb{N}$ define correspondence $\hat{\Delta}_i^k : T_i^* \rightrightarrows \Theta_i \times \prod_{j \neq i} \hat{M}_j^0$ by setting, for each $t_i \in T_i^*$,

$$\hat{\Delta}_i^k(t_i) := \Delta_i^k(t_i) \times \prod_{j \neq i} \prod_{\ell > k} \hat{O}_j^\ell.$$

Obviously, $\hat{\Delta}_i^k$ is nonempty-valued, compact-valued and continuous for every $k \in \mathbb{N}$, and it holds that:

$$\Delta_i^*(t_i) = \bigcap_{k \in \mathbb{N}} \hat{\Delta}_i^k(t_i).$$

Now, since $\Theta_i \times T_{-i}^*$ is compact Polish, we know that (1) the compactness of every $\hat{\Delta}_i^k(t_i)$ implies the compactness of $\Delta_i^*(t_i)$, and (2) the fact that collection $\{\hat{\Delta}_i^k(t_i) | \, k \in \mathbb{N}\}$ satisfies the finite intersection property implies that $\Delta_i^*(t_i)$ is nonempty. That Δ_i^* is upper-hemicontinuous follows from the fact that every $\hat{\Delta}_i^k$ is continuous. To see that Δ_i^* is lower-hemicontinuous fix $t_i \in T_i^*$ and open set $V_i \subseteq \Theta_i \times T_{-i}^*$ where $V_i \cap \Delta_i(t_i) \neq \emptyset$. Since V_i is open in $\Theta_i \times T_{-i}^*$ we know that there exist:

- Some $V_i' \subseteq \Theta_i \times \hat{M}_{-i}^0$ where $V_i = V_i' \cap \Theta_i \times T_{-i}^*$.
- Some $k \in \mathbb{N}$ and some open $W_i \subseteq \Theta_i \times \prod_{j \neq i} \prod_{\ell=1}^k \hat{O}_j^\ell$ such that $V_i' := W_i \times \prod_{j \neq i} \prod_{\ell > k} \hat{O}_j^\ell$.

Now, since $\hat{\Delta}_i^k$ is lower-hemicontinuous we know that there exists some open $U_i \subseteq T_i^*$ such that $W_i \cap \hat{\Delta}_i^k(t_i') \neq \emptyset$ for every $t_i' \in T_i^*$. It follows that, for every $t_i' \in U_i$, $V_i \cap \Delta_i^*(t_i') \neq \emptyset$ and hence, that Δ_i^* is lower-hemicontinuous. Finally, to see that Δ_i^* is surjective pick arbitrary nonempty and compact $K_i \subseteq \Theta_i \times T_{-i}^*$ and set $\Delta_i = (\Delta_i^k)_{k \in \mathbb{N}}$ were $\Delta_i^k := \mathrm{proj}_{\hat{O}_i^k} K_i$ for every $k \in \mathbb{N}$. Obviously, there exists some $t_i \in T_i^*$ such that $\Delta_i^k(t_i) = \Delta_i^k$ for every $k \in \mathbb{N}$ and hence, such that $\Delta_i^* = K_i$. ■

Belief-map. Define first the following set:

$$\hat{T}_i^* := \left\{ t_i = (\Delta_i^k, \tau_i^k)_{k \in \mathbb{N}} \in \hat{M}_i^0 \;\middle|\; \begin{array}{ll} (1) & \tau_i^k[\Delta_i^k] = 1 \\ (2) & \text{For every } k \in \mathbb{N},\ \mathrm{Proj}_{\bar{M}_{-i}^k} \Delta_i^{k+1} = \Delta_i^k \\ (3) & \text{For every } k \in \mathbb{N},\ \mathrm{marg}_{\bar{M}_{-i}^k} \tau_i^{k+1} = \tau_i^k \end{array} \right\}.$$

Then extend map τ_i above to $\hat{T}_i^*$ by defining $\hat{\tau}_i : \hat{T}_i^* \to \prod_{k \in \mathbb{N}} \Delta(\bar{M}_{-i}^k)$ as follows:

$$(\Delta_i^k, \tau_i^k)_{k \in \mathbb{N}} \mapsto (\tau_i^k)_{k \in \mathbb{N}}.$$

The Kolmogorov Extension Theorem guarantees the existence of a unique map $g_i : \hat{\tau}_i(\hat{T}_i^*) \to \Delta(\Theta_i \times \hat{M}_{-i}^0)$ such that $\mathrm{marg}_{\bar{M}_{-i}^k} g_i((\tau_i^k)_{k \in \mathbb{N}}) = \tau_i^k$—a property that, in turn, implies the continuity of g_i. Given all the above, define map $\tau_i^* : T_i^* \to \Delta(\Theta_i \times T_{-i}^*)$ as follows:

$$\tau_i^* := g_{i|\tau_i(T_i^*)} \circ \tau_i.$$

Lemma 3. *τ_i^* is surjective, continuous and satisfies the following two for every $t_i \in T_i^*$:*

(1) $\tau_i^*(t_i)[\Delta_i^*(t_i)] = 1$

(2) $\mathrm{marg}_{M_{-i}^k} \tau_i^*(t_i) = \tau_i^{k+1}(t_i)$ *for every $k \geq 0$*

Proof. To see that τ_i^* pick $\beta_i \in \Delta(\Theta_i \times T_{-i}^*)$ and define $\beta_i^\infty := (\beta_i^k)_{k \in \mathbb{N}}$ where $\beta_i^k := \mathrm{marg}_{M_{-i}^k} \beta_i$ for every $k \in \mathbb{N}$. Obviously, $\beta_i^\infty \in \tau_i(T_i^*)$ and, obviously as well, $\tau_i^*(t_i) = \beta_i$ for every $t_i \in \tau_i^{-1}(\beta_i^\infty)$. The continuity of τ_i^* follows from the obvious facts that both g_i and τ_i are continuous. That (2) holds is immediate. To see that (1) holds fix $t_i = (\Delta_i^k, \tau_i^k)_{k \in \mathbb{N}} \in T_i^*$ and notice that, for every $k \in \mathbb{N}$ we have that $\tau_i^k[\Delta_i^k] = 1$, and thus, also that:

$$\mathrm{marg}_{M_{-i}^k} \tau_i^*(t_i)[\mathrm{Proj}_{M_{-i}^k} \Delta_i^*(t_i)] = \tau_i^*(t_i) \left[\hat{\Delta}_i^k(t_i)\right] = 1,$$

what clearly implies that $\tau_i^*(t_i)[\Delta_i^*(t_i)] = 1$. ■

A.2.2 Universality of the space of all hierarchies of models

PROPOSITION 3. *Let I be a finite set and, for each $i \in I$, let Θ_i be a compact Polish set. Then, the type-space $\mathscr{T}^* = (\Theta_i, T_i^*, \Delta_i^*, \tau_i^*)_{i \in I}$ as defined above is complete and terminal.*

Proof. To verify completeness pick arbitrary pair (K_i, β_i) where P_i is a nonempty and compact subset of $\Theta_i \times T_{-i}^*$ and $\beta_i \in \Delta(\Theta_i \times T_{-i}^*)$ satisfies that $\beta_i[K_i] = 1$. Then, define, for every $k \in \mathbb{N}$ the following two objects:

- $K_i^k := \mathrm{Proj}_{M_{-i}^k} K_i$.
- $\beta_i^k := \mathrm{marg}_{M_{-i}^k} \beta_i$.

Then, the completeness of $\mathscr{T}^*$ follows from the immediate fact that $t_i := (K_i^k, \beta_i^k)_{k \in \mathbb{N}}$ is an element of T_i^* such that $\Delta_i^*(t_i) = K_i$ and $\tau_i^*(t_i) = \beta_i$. To verify terminality fix type-space $\mathscr{T} = (\Theta_i', T_i, \Delta_i, \tau_i)_{i \in I}$ where $\Theta_i' \subseteq \Theta_i$ for every player i. The, for every player i and every $t_i \in T_i$ define first:

- $K_i^1(t_i) := \mathrm{Proj}_{\Theta_i'} \Delta_i(t_i)$
- $\beta_i^1(t_i)[E] := \mathrm{marg}_{\Theta_i'} \tau_i(t_i)[E \cap \Theta_i']$ for every measurable $E \subseteq \Theta_i$.

and set $X_{-i}^1 := \prod_{j \neq i}(K_j^1 \times \beta_j^1)(T_j)$. Then, define recursively:

- $K_i^{k+1}(t_i) := \mathrm{Proj}_{X_{-i}^k} \Delta_i(t_i)$
- $\beta_i^{k+1}(t_i)[E] := \mathrm{marg}_{X_{-i}^k} \tau_i(t_i)[E \cap X_{-i}^k]$ for every measurable $E \subseteq\ M_{-i}^k$.

and set $X_{-i}^{k+1} := \prod_{j \neq i}(K_j^{k+1} \times \beta_j^{k+1})(T_j)$. It should be noted that maps $\phi_i^{k+1,\mathscr{T}} : T_i \to X_{-i}^k$ given by $t_i \mapsto (K_i^k(t_i), \beta_i^k(t_i))$ are continuous and open, by virtue of projection and marginalization being continuous and open.[38] Finally, for each player i, define map $\phi_i^{\mathscr{T}} : T_i \to T_i^*$ by setting:

$$\phi_i^{\mathscr{T}}(t_i) := (K_i^k(t_i), \beta_i^k(t_i))_{k \in \mathbb{N}}$$

for every $t_i \in T_i$. Obviously, $\phi_i^{\mathscr{T}}$ is continuous, and satisfies properties (1) and (2) in 11. Thus, we conclude that $\mathscr{T}^*$ is terminal. ■

A.3 TYPE-SPACE INVARIANCE

A.3.1 Auxiliary definitions

DEFINITION 12 (Interim solution concept). *Let $\mathscr{E}$ be an extensive-form and let i be a player. Then, an* interim solution concept $\mathcal{I}_i$ *for player i is a mapping that associates each pair $(\mathscr{T}, t_i)$, where $\mathscr{T}$ is a type-space and t_i is a type of player i, with a subset of i's strategies $\mathcal{I}_i^{\mathscr{T}}(t_i) \subseteq S_i$.*

[38] Continuity is obvious, and so is the openness of projections. For openness of marginalization see Ditor and Eifler (1972).

DEFINITION 13 (Robustness). *Let $\mathscr{E}$ be an extensive-form and let i be a player. Then, an interim solution concept $\mathcal{I}_i$ for player i is* robust *if, for every type-space $\mathscr{T}$ the correspondence $\mathcal{I}_i^{\mathscr{T}} : T_i \rightrightarrows S_i$ is upper-hemicontinuous.*

DEFINITION 14 (Type-space invariance). *Let $\mathscr{E}$ be an extensive-form and let i be a player. Then, an interim solution concept $\mathcal{I}_i$ for player i is* type-space invariant *if, for every type-space $\mathscr{T}$ and every type t_i,*

$$\mathcal{I}_i^{\mathscr{T}}(t_i) = \mathcal{I}_i^{\mathscr{T}^*}(\phi_i^{\mathscr{T}}(t_i)).$$

A.3.2 Auxiliary lemmas

LEMMA 4. *Let $(\mathscr{E}, \mathcal{T}^*)$ be a dynamic Bayesian game, let i be a player and let $(\mathcal{I})_{j\neq i}$ be a partial profile of type-space invariant interim solution concepts. Then, for any type-space $\mathscr{T}$, any type $t_i \in T_i$ and any conjecture $\mu_i \in \mathrm{C}_i^{\mathscr{T}}(t_i)$ there exists some conjecture $\mu_i^* \in \mathrm{C}_i^{\mathscr{T}^*}(\phi_i^{\mathscr{T}}(t_i))$ such that $r_i(\mu_i^* = r_i(\mu_i)$ and the following three hold:*

1. *If μ_i displays initial belief in $\mathcal{I}_{-i}^{\mathscr{T}}$, then μ_i^* displays initial belief in $\mathcal{I}_{-i}^{\mathscr{T}^*}$.*
2. *If μ_i displays strong belief in $\mathcal{I}_{-i}^{\mathscr{T}}$, then μ_i^* displays strong belief in $\mathcal{I}_{-i}^{\mathscr{T}^*}$.*
3. *If μ_i displays future belief in $\mathcal{I}_{-i}^{\mathscr{T}}$, then μ_i^* displays future belief in $\mathcal{I}_{-i}^{\mathscr{T}^*}$.*

Proof. Fix conjecture $\mu_i \in \mathrm{C}_i^{\mathscr{T}}(t_i)$. Then, we define μ_i^* in four steps:

- First, define map $\psi_i^{\mathscr{T}} : S_{-i} \times \Theta \to S_{-i} \times \Theta^*$ as follows:

$$(s_{-i}, \theta, t_{-i}) \mapsto (s_{-i}, \theta, (\phi_j^{\mathscr{T}}(t_{-i}))_{j\neq i}).$$

  Obviously, $\psi_i^{\mathscr{T}}$ is continuous.

- Now, let $\mathrm{H}_i(\mu_i)$ collect the set of histories in $H_i \cup \{h^0\}$ where μ_i updates beliefs from scratch:

$$\mathrm{H}_i(\mu_i) := \left\{ h \in H_i \cup \{h^0\} \;\middle|\; \begin{array}{c} \text{For every } h' \in H_i \cup \{h^0\} \text{ s.t. } S_{-i}(h') \subseteq S_{-i}(h): \\ \mathrm{marg}_{S_{-i}} \mu_i(h')[S_{-i}(h)] = 0 \end{array} \right\}$$

- Then, for every $h \in \mathrm{H}_i(\mu_i)$ and every event $E_{-i} \subseteq S_{-i} \times \Theta^* \times T_{-i}^*$ set:

$$\mu_i^*(h)[E_{-i}] = \mu_i(h)\left[(\psi_i^{\mathscr{T}})^{-1}(E_{-i})\right].$$

  The continuity of each $\psi_i^{\mathscr{T}}$ guarantees that the above is well-defined.

- For every $h \notin \mathrm{H}_i(\mu_i)$ define $\mu_i^*(h)$ via conditional probability.

Obviously, $\mu_i^* := (\mu_i^*(h))_{h\in H_i\cup\{h^0\}}$ is a well-defined element of $\mathrm{C}_i^{\mathscr{T}^*}(\phi_i^{\mathscr{T}}(t_i))$ and furthermore, the marginals on $S_{-i} \times \Theta^*$ of $\mu_i^*(h)$ and $\mu_i(h)$ coincide for every $h \in H_i \cup \{h^0\}$. It follows from the latter that $r_i(\mu_i^*) = r_i(\mu_i^*)$. Now, notice that the following facts hold:

(F0) For every pair of measurable events $E_{-i} \subseteq S_{-i} \times \Theta^* \times T_{-i}$ and $F_{-i} \subseteq S_{-i} \times \Theta^* \times T_{-i}^*$ such that $(\psi_i^{\mathscr{T}})^{-1}(E_{-i}) \subseteq F_{-i}$ and every history $h \in H_i \cap \{h^0\}$, if there exists some measurable $M \subseteq E_{-i}$ such that $\mu_i(h)[M] = 1$ then there exists some measurable $M^* \subseteq \psi_i^{-1}(F_{-i})$ such that $\mu_i^*(h)[M^*] = 1$.

(F1) Since every $\mathcal{I}_j$ is type-space invariant, then, among the following two sets, the first one is included in the second:

$$\Theta^* \times \mathrm{Graph}(\mathcal{I}_{-i}^{\mathscr{T}}) = \Theta^* \times \left\{ (s_{-i}, t_{-i}) \in S_{-i} \times T_{-i} \,\middle|\, s_{-i} \in \mathcal{I}_{-i}^{\mathscr{T}}(t_i) \right\},$$

$$(\psi_i^{\mathscr{T}})^{-1}\left(\Theta^* \times \mathrm{Graph}(\mathcal{I}_{-i}^{\mathscr{T}^*})\right) = \Theta^* \times \left\{(s_{-i}, t_{-i}) \in S_{-i} \times T_{-i} \middle| s_{-i} \in \prod_{j \neq i} \mathcal{I}_j^{\mathscr{T}^*}(\phi_j^{\mathscr{T}}(t_j))\right\}.$$

(F2) Since every $\mathcal{I}_j$ is type-space invariant, then, for every $h \in H_i \cup \{h^0\}$

$$S_{-i}(h) \times \Delta_i(t_i) \cap \Theta^* \times \mathrm{Graph}(\mathcal{I}_{-i}^{\mathscr{T}}) \neq \emptyset \Longleftrightarrow S_{-i}(h) \times \Delta_i^*(\phi_i^{\mathscr{T}}(t_i)) \cap \Theta^* \times \mathrm{Graph}(\mathcal{I}_{-i}^{\mathscr{T}^*}) \neq \emptyset.$$

(F3) Since every $\mathcal{I}_j$ is type-space invariant, then, among the following two sets, for every $h \in H_i \cup \{h^0\}$ the first one is included in the second:

$$\left[\Theta^* \times \mathrm{Graph}(\mathcal{I}_{-i}^{\mathscr{T}})\right]_h = \Theta^* \times \left\{(s_{-i}, t_{-i}) \in S_{-i} \times T_{-i} \middle| [s_{-i}]_h \cap \mathcal{I}_{-i}^{\mathscr{T}}(t_{-i}) \neq \emptyset\right\},$$

$$(\psi_i^{\mathscr{T}})^{-1}\left(\left[\Theta^* \times \mathrm{Graph}(\mathcal{I}_{-i}^{\mathscr{T}^*})\right]_h\right) = \Theta^* \times \left\{(s_{-i}, t_{-i}) \in S_{-i} \times T_{-i} \middle| [s_{-i}]_h \cap \prod_{j \neq i} \mathcal{I}_j^{\mathscr{T}^*}(\phi_j^{\mathscr{T}}(t_j)) \neq \emptyset\right\}.$$

Then, claim 1 in the lemma follows directly from (F0) and (F1), claim 2 follows from (F0), (F1) and (F2), and claim 3 follows from (F0), (F1) and (F3). ■

Lemma 5. *Let $(\mathscr{E}, \mathcal{T}^*)$ be a dynamic Bayesian game, let i be a player and let $(\mathcal{I})_{j \neq i}$ be a partial profile of robust, type-space invariant interim solution concepts. Then, for any type-space $\mathscr{T}$, any type $t_i \in T_i$ and any conjecture $\mu_i^* \in \mathrm{C}_i^{\mathscr{T}^*}(\phi_i^{\mathscr{T}}(t_i))$ there exists some conjecture $\mu_i \in \mathrm{C}_i^{\mathscr{T}}(t_i)$ such that $r_i(\mu_i) = r_i(\mu_i^*)$ and the following three hold:*

1. *If μ_i^* displays initial belief in $\mathcal{I}_{-i}^{\mathscr{T}^*}$ then μ_i displays initial belief in $\mathcal{I}_{-i}^{\mathscr{T}}$.*
2. *If μ_i^* displays strong belief in $\mathcal{I}_{-i}^{\mathscr{T}^*}$ then μ_i displays strong belief in $\mathcal{I}_{-i}^{\mathscr{T}}$.*
3. *If μ_i^* displays future belief in $\mathcal{I}_{-i}^{\mathscr{T}^*}$ then μ_i displays future belief in $\mathcal{I}_{-i}^{\mathscr{T}}$.*

Proof. Fix conjecture $\mu_i^* \in \mathrm{C}_i^{\mathscr{T}^*}(\varphi_i^{\mathscr{T}}(t_i))$. Then, we define μ_i in four steps:

- First, define correspondence $\varphi_i^{\mathscr{T}} : S_{-i} \times \Theta \times T_{-i}^* \rightrightarrows S_{-i} \times \Theta^* \times T_{-i}$ as follows:

  $$(s_{-i}, \theta, t_{-i}) \mapsto (\psi_i^{\mathscr{T}})^{-1}(s_{-i}, \theta, (t_{-i}).$$

  Notice that $\varphi_i^{\mathscr{T}}$ is weakly measurable. To see it, notice that for every open $V \subseteq S_{-i} \times \Theta^* \times T_{-i}$ we have that:

  $$\left\{(s_{-i}, \theta, t_{-i}) \in S_{-i} \times \Theta^* \times T_{-i}^* | \varphi_i^{\mathscr{T}}(s_{-i}, \theta, t_{-i}) \cap V \neq \emptyset\right\} = \psi_i^{\mathscr{T}}(V),$$

  and that:

  $$\psi_i^{\mathscr{T}}(V) = \bigcap_{k \in \mathbb{N}} \left\{(s_{-i}, \theta, t_{-i}) \in S_{-i} \times \Theta^* \times T_{-i}^* \middle| ((\Delta_j^k(t_j), \tau_j^k(t_j))_{j \neq i} \in \Theta \times \phi_{-i}^{k,\mathscr{T}}(V)\right\}.$$

  The fact that every $\phi_i^{k,\mathscr{T}}$ is an open guarantees then that $\psi_i^{\mathcal{T}}(V)$ is measurable. The weak measurability of $\varphi_i^{\mathscr{T}}$ follows then. We know then by the Kuratowski and Ryll-Nardzewski Selection Theorem that $\varphi_i^{\mathscr{T}}$ admits a measurable selector $\sigma_i^{\mathscr{T}} : S_{-i} \times \Theta \times T_{-i}^* \rightrightarrows S_{-i} \times \Theta^* \times T_{-i}$.[39]

[39] That is, σ_i is measurable and such that $\sigma_i^{\mathscr{T}}(s_{-i}, t_{-i}, t_{-i}^*) \in \varphi_i^{\mathscr{T}}(s_{-i}, \theta, t_{-i})$ for every $(s_{-i}, \theta, t_{-i}) \in S_{-i} \times \Theta^* \times T_{-i}^*$.

- Now, let $\mathrm{H}_i(\mu_i)$ collect the set of histories in $H_i \cup \{h^0\}$ where μ_i^* updates beliefs from scratch:

$$\mathrm{H}_i(\mu_i^*) := \left\{ h \in H_i \cup \{h^0\} \;\middle|\; \begin{array}{c} \text{For every } h' \in H_i \cup \{h^0\} \text{ s.t. } S_{-i}(h') \subseteq S_{-i}(h): \\ \mathrm{marg}_{S_{-i}} \mu_i^*(h')[S_{-i}(h)] = 0 \end{array} \right\}$$

- Then, for every $h \in \mathrm{H}_i(\mu_i^*)$ define $\mu_i(h) \in \Delta(S_{-i} \times \Theta^* \times T_{-i})$ by setting, for every measurable $E \subseteq S_{-i} \times \Theta^* \times T_{-i}$:
$$\mu_i(h)[E] := \mu_i^*(h)[(\sigma_i^{\mathscr{T}})^{-1}(E)].$$

- For every $h \notin \mathrm{H}_i(\mu_i^*)$ define $\mu_i(h)$ via conditional probability.

Obviously, $\mu_i := (\mu_i(h))_{h \in H_i \cup \{h^0\}}$ is a well-defined element of $\mathrm{C}_i^{\mathscr{T}}(t_i)$ and furthermore, the marginals on $S_{-i} \times \Theta^*$ of $\mu_i(h)$ and $\mu_i^*(h)$ coincide for every $h \in H_i \cup \{h^0\}$. It follows from the latter that $r_i(\mu_i^*) = r_i(\mu_i^*)$. Now, notice that the following facts hold:

(F0) For every pair of measurable events $E_{-i} \subseteq S_{-i} \times \Theta^* \times T^*_{-i}$ and $F_{-i} \subseteq S_{-i} \times \Theta^* \times T_{-i}$ such that $(\sigma_i^{\mathscr{T}})^{-1}(E_{-i}) \subseteq F_{-i}$ and every history $h \in H_i \cap \{h^0\}$, if there exists some measurable $M^* \subseteq E_{-i}$ such that $\mu_i^*(h)[M] = 1$ then there exists some measurable $M \subseteq \sigma_i^{-1}(F_{-i})$ such that $\mu_i(h)[M] = 1$.

(F1) Since every $\mathcal{I}_j$ is type-space invariant, then, among the following two sets, the first one is included in the second:

$$\Theta^* \times \mathrm{Graph}(\mathcal{I}_{-i}^{\mathscr{T}^*}) = \Theta^* \times \left\{ (s_{-i}, t_{-i}) \in S_{-i} \times T^*_{-i} \;\middle|\; s_{-i} \in \mathcal{I}_{-i}^{\mathscr{T}^*}(t_{-i}) \right\},$$

$$(\sigma_i^{\mathscr{T}})^{-1}\left(\Theta^* \times \mathrm{Graph}(\mathcal{I}_{-i}^{\mathscr{T}})\right) = \Theta^* \times \left\{ (s_{-i}, t_{-i}) \in S_{-i} \times T^*_{-i} \;\middle|\; s_{-i} \in \prod_{j \neq i} \mathcal{I}_j^{\mathscr{T}}(\sigma_{T_j}^{\mathscr{T}}(t_j)) \right\},$$

where $(s_{-i}, \theta, (\sigma_{T_j}^{\mathscr{T}}(t_j))_{j \neq i}) = \sigma_i^{\mathscr{T}}(s_{-i}, \theta, t_{-i})$ for every $(s_{-i}, \theta, t_{-i}) \in S_{-i} \times \Theta^* \times T^*_{-i}$.

(F2) Since every $\mathcal{I}_j$ is type-space invariant, then, for every $h \in H_i \cup \{h^0\}$

$$S_{-i}(h) \times \Delta_i(t_i) \cap \Theta^* \times \mathrm{Graph}(\mathcal{I}_{-i}^{\mathscr{T}}) \neq \emptyset \Longleftrightarrow S_{-i}(h) \times \Delta_i^*(\phi_i^{\mathscr{T}}(t_i)) \cap \Theta^* \times \mathrm{Graph}(\mathcal{I}_{-i}^{\mathscr{T}^*}) \neq \emptyset.$$

(F3) Since every $\mathcal{I}_j$ is type-space invariant, then, among the the following two sets, for every $h \in H_i \cup \{h^0\}$ the first one is included in the second:

$$\left[\Theta^* \times \mathrm{Graph}(\mathcal{I}_{-i}^{\mathscr{T}^*})\right]_h = \Theta^* \times \left\{ (s_{-i}, t_{-i}) \in S_{-i} \times T^*_{-i} \;\middle|\; [s_{-i}]_h \cap \mathcal{I}_{-i}^{\mathscr{T}^*}(t_{-i}) \neq \emptyset \right\},$$

$$(\sigma_i^{\mathscr{T}})^{-1}\left(\left[\Theta^* \times \mathrm{Graph}(\mathcal{I}_{-i}^{\mathscr{T}})\right]_h\right) = \Theta^* \times \left\{ (s_{-i}, t_{-i}) \in S_{-i} \times T^*_{-i} \;\middle|\; [s_{-i}]_h \cap \prod_{j \neq i} \mathcal{I}_j^{\mathscr{T}^*}(\sigma_{T_j}^{\mathscr{T}}(t_j)) \neq \emptyset \right\}.$$

Then, claim 1 in the lemma follows directly from (F0) and (F1), claim 2 follows from (F0), (F1) and (F2), and claim 3 follows from (F0), (F1) and (F3). ■

A.3.3 Proof of Proposition 1

PROPOSITION 1. *Let $(\mathscr{E}, \mathscr{T})$ be a dynamic Bayesian game. Then for every player i and every type t_i the following three hold:*

1. $\mathcal{W}_i^{\mathscr{T}}(t_i) = \mathcal{W}_i^{\mathscr{T}^*}(\phi_i^{\mathscr{T}}(t_i))$.

2. $\mathcal{S}_i^{\mathscr{T}}(t_i) = \mathcal{S}_i^{\mathscr{T}^*}(\phi_i^{\mathscr{T}}(t_i))$.

3. $\mathcal{B}_i^{\mathscr{T}}(t_i) = \mathcal{B}_i^{\mathscr{T}^*}(\phi_i^{\mathscr{T}}(t_i))$.

Proof. We will prove that for every player i and every $k \geq 0$ the following hold:

1. $\mathcal{W}_{i,k}^{\mathscr{T}}(t_i) = \mathcal{W}_{i,k}(\phi_i^{\mathscr{T}}(t_i))$. The claim is trivially true for $k = 0$, so let us prove that if is true for some $k \geq 0$, then it is also true for $k + 1$. Fix player i and type t_i. For the eastward inclusion pick conjecture $\mu_i \in \mathrm{C}_i^{\mathscr{T}}(t_i)$ that displays initial belief in $\mathcal{W}_{-i,k}^{\mathscr{T}}$ (where, according to the induction hypothesis, every $\mathcal{W}_{j,k}^{\mathscr{T}}$ is type-space invariant) and $s_i \in r_i(\mu_i)$. Then, we know from Lemma 4 that there exists some $\mu_i^* \in \mathrm{C}_i^{\mathscr{T}^*}(\phi_i^{\mathscr{T}}(t_i))$ that initially believes in $\mathcal{W}_{-i,k}$ and such that $r_i(\mu_i^*) = r_i(\mu_i)$. For the westward inclusion pick conjecture $\mu_i \in \mathrm{C}_i^{\mathscr{T}^*}(\phi_i^{\mathscr{T}}(t_i))$ that displays initial belief in $\mathcal{W}_{-i,k}$ and $s_i \in r_i(\mu_i)$. Then, we know from Lemma 5 that there exists some $\mu_i' \in \mathrm{C}_i^{\mathscr{T}}((t_i))$ that displays initial believes in $\mathcal{W}_{-i,k}^{\mathscr{T}}$ and such that $r_i(\mu_i') = r_i(\mu_i)$. Hence, $\mathcal{W}_{i,k+1}^{\mathscr{T}}(t_i) = \mathcal{W}_{i,k+1}(\phi_i^{\mathscr{T}}(t_i))$.
2. $\mathcal{S}_i^{\mathscr{T}}(t_i) = \mathcal{S}_i(\phi_i^{\mathscr{T}}(t_i))$. Simply proceed as in the previous case, but substitute $\mathcal{W}$ for $\mathcal{S}$, and "initial belief" for "strong belief."
3. $\mathcal{B}_i^{\mathscr{T}}(t_i) = \mathcal{B}_i(\phi_i^{\mathscr{T}}(t_i))$. Simply proceed as in the previous case, but substitute $\mathcal{S}$ for $\mathcal{B}$, and "initial belief" for "future belief."

The claims of the proposition follow directly. ■

B Proof of Theorem 1

Throughout the following proof, for every player i, strategy s_i and history $h \in H$ we denote by $[s_i]_h$ the set of strategies of player i that are equivalent to s_i for every history h' of player i's weakly following h. That is:

$$[s_i]_h := \left\{ s_i' \in S_i \,\middle|\, s_i'(h') = s_i(h') \text{ for every } h' \in H_i \cup \{h^0\} \text{ s.t. } S_i(h') \subseteq S_i(h) \right\},$$

and for any $s_{-i} \in S_{-i}$ we denote:

$$[s_{-i}]_h := \prod_{j \neq i} [s_j]_h.$$

Then, the proof of Theorem 1 is simple:

Theorem 1. *Let $(\mathscr{E}, \mathscr{T}^*)$ be a dynamic Bayesian game. Then, for any player i the following two hold:*

1. $\mathcal{W}_i : T_i^* \rightrightarrows S_i$ *is upper-hemicontinuous.*

2. $\mathcal{B}_i : T_i^* \rightrightarrows S_i$ *is upper-hemicontinuous.*

Proof. We will prove the two slightly more general claims that, for every playaer i and every $k \geq 0$, both $\mathcal{W}_{i,k}$ and $\mathcal{B}_{i,k}$ are upper-hemicontinuous. We proceed by induction on k. The initial case $(k = 0)$ is trivially true, so we can focus on the proof of the inductive step. To this end, fix $k \geq 0$ for which the claims hold, and let us verify that then, they also hold for $k + 1$. Fix player i, sequence of types $(t_i^n)_{n \in \mathbb{N}}$ with limit t_i, and:

- Strategy $s_i \in \bigcap_{n \in \mathbb{N}} \mathcal{W}_{i,k+1}(t_i^n)$ and sequence of conjectures $(\mu_i^n)_{n \in \mathbb{N}}$ where, for each $n \in \mathbb{N}$, μ_i^n justifies the inclusion of s_i in $\mathcal{W}_{i,k+1}(t_i^n)$. Then, notice first that $\mathrm{Graph}(\mathcal{W}_{-i,k})$ is equal to the following set:

$$\left\{ (s_{-i}, t_{-i}) \in S_{-i}(h^0) \times T_{-i}^* \,\middle|\, [s_{-i}]_{h^0} \cap \mathcal{W}_{-i,k}(t_{-i}) \neq \emptyset \right\},$$

and that we know from the induction hypothesis and Lemma 6 that the latter is closed. In consequence, we have that:

$$\forall n \in \mathbb{N},\ \mu_i^n(h^0)\left[\Theta^* \times \mathrm{Graph}(\mathcal{W}_{-i,k})\right] = 1 \implies \mu_i(h^0)\left[\Theta^* \times \mathrm{Graph}(\mathcal{W}_{-i,k})\right] = 1$$

Since, obviously, s_i is a sequential best-reply to μ_i, we conclude the latter justifies the inclusion of s_i in $\mathcal{W}_{i,k+1}(t_i)$.

- Strategy $s_i \in \bigcap_{n\in\mathbb{N}} \mathcal{B}_{i,k+1}(t_i^n)$ and sequence of conjectures $(\mu_i^n)_{n\in\mathbb{N}}$ where, for each $n \in \mathbb{N}$, μ_i^n justifies the inclusion of s_i in $\mathcal{B}_{i,k+1}(t_i^n)$. Then, we know from the induction hypothesis and Lemma 6 that the following set is closed for every $h \in H_i \cup \{h^0\}$

$$E_{-i}(h) := \left\{(s_{-i}, t_{-i}) \in S_{-i}(h) \times T_{-i}^* \,|[s_{-i}]_h \cap \mathcal{B}_{-i,k}(t_{-i}) \neq \emptyset\right\},$$

and hence, we have that:

$$\forall n \in \mathbb{N},\ \mu_i^n(h)\left[\Theta^* \times E_{-i}(h)\right] = 1 \implies \mu_i(h)\left[\Theta^* \times E_{-i}(h)\right] = 1$$

Since, obviously, s_i is a sequential best-reply to μ_i, we conclude the latter justifies the inclusion of s_i in $\mathcal{B}_{i,k+1}(t_i)$.

It follows from the above that both $\mathcal{W}_i$ and $\mathcal{B}_{i,k}$ are upper-hemicontinuous. ■

Lemma 6. *Let $(\mathscr{E}, \mathscr{T}^*)$ be a dynamic Bayesian game, let i be a player, and let $\mathcal{I}_i : T_i^* \rightrightarrows S_i$ be an upper-hemicontinuous correspondence. Then, the following set is closed for every $h \in H$:*

$$E_i(h) = \{(s_i, t_i) \in S_i(h) \times T_i^* \,|[s_i]_h \cap \mathcal{I}_i(t_i) \neq \emptyset\},$$

Proof. Fix $h \in H$ and pick as convergent sequence $(s_i^n, t_i^n)_{n\in\mathbb{N}}$ in $E(h)$ with limit (s_i, t_i). Since S_i is finite, we know that there exists some $N \in \mathbb{N}$ such that $s_i^n = s_i$ for every $n \geq N$. Now, since $[s_i]_h \cap \mathcal{I}_i(t_i^n) \neq \emptyset$, we know that for every $n \in \mathbb{N}$ where exists some $\tilde{s}_i^n \in [s_i]_h \cap \mathcal{I}_i(t_i^n)$, and since S_i is finite, we also know that there exist some $\tilde{N} \geq N$ and some $\tilde{s}_i \in S_i$ such that $\tilde{s}_i^n = \tilde{s}_i$ for every $n \geq \tilde{N}$. Now, since $\mathcal{I}$ is upper-hemicontinuous we conclude $\tilde{s}_i \in \mathcal{I}_i(t_i)$ too. Thus, we have that $[s_i]_h \cap \mathcal{I}_i(t_i) \neq \emptyset$ and hence, that $(s_i, t_i) \in E_i(h)$. In consequence, the latter is closed. ■

C Proof of Theorem 2

C.1 Auxiliary notation

$$\mathrm{H}_{i,k}(\Delta_i^*(t_i)) := \left\{h \in H_i \cup \{h^0\} \,|S_{-i}(h) \times \Delta_i^*(t_i) \cap \mathrm{Graph}(\mathcal{S}_{-i,k}) \times \Theta^* \neq \emptyset\right\}$$

$$B_i(E_{-i}) := \{t_i \in T_i^* \,|\Delta_i^*(t_i) \subseteq E_{-i}\}.$$

$$L_i^k := \{t_i \in T_i^* \,|\mathcal{S}_{i,k}(t_i) = \mathcal{S}_i(t_i)\}$$

C.2 Auxiliary lemmas

Lemma 7. *Let $(\mathscr{E}, \mathscr{T}^*)$ be a dynamic Bayesian game. Then, for every player i and every open and dense set $E_{-i} \subseteq T_{-i}^*$ the set $B_i(\Theta^* \times E_{-i})$ is open and dense in T_i^*.*

Proof. Fix arbitrary type $t_i \in T_i^*$ and open and dense $E_{-i} \subseteq T_{-i}^*$. Since $\Delta_i^*(t_i)$ is compact we know that, for every $n \in \mathbb{N}$, it contains finitely many elements $x_1^n, \ldots, x_{N_n}^n$ such that:

$$\Delta_i^*(t_i) \subseteq \bigcup_{x_i \in \Delta_i^*(t_i)} \mathring{B}_{1/n}(x_i) = \bigcup_{\ell=1}^{N_n} \mathring{B}_{1/n}(x_i^{n,\ell}).$$

Now, since E_{-i} is open and dense then, for every $n \in \mathbb{N}$ and every $\ell = 1, \ldots, N_n$:

$$E_{-i}^{n,\ell} := \mathring{B}_{1/n}(x_i^{n,\ell}) \cap \Theta^* \times E_{-i}$$

is open and nonempty and thus, we know that there exists some $\varepsilon^n \in (0,1)$ such that, for every $\ell = 1, \ldots, N_n$,

$$Y_i^{n,\ell} := \overline{B}_{-\varepsilon^n/n}\left(E_{-i}^{n,\ell}\right)$$

is nonempty and, of course, contained in $\Theta^* \times E_{-i}$. Now, for each $\ell = 1, \ldots, N_n$ pick arbitrary $y_i^{n,\ell} \in Y_i^{n,\ell}$ and notice that by setting:

$$\tau_i^n[y_i^{n,\ell}] := \tau_i^*(t_i)\left[B_{1/n}(x_i^{n,\ell})\right],$$

we obtain a well-defined probability measure $\tau_i^n \in \Delta(\Theta^* \times T_{-i}^*)$. Finally, define t_i^n as follows:

$$t_i^n := h_i^*\left(Y_i^n, \tau_i^n\right),$$

where $Y_i^n := \bigcup_{\ell=1}^{N_m} Y_i^{n,\ell}$. Obviously, for every $n \in \mathbb{N}$ it holds that $\Delta_i^*(t_i) \subseteq \Theta^* \times E_{-i}$, and obviously too, $(t_i^n)_{n \in \mathbb{N}}$ converges to t_i. ■

LEMMA 8. *Let $(\mathscr{E}, \mathscr{T}^*)$ be a dynamic Bayesian game. Then, for every player i the following set is open and dense in T_i^*:*

$$\bigcup_{k \in \mathbb{N}} \mathring{L}_i^k.$$

Proof. For each player i set first:

$$Y_i^1 := \{t_i \in T_i^* \,|\mathcal{S}_{i,1}(t_i) = [s_i] \text{ for some } s_i \in S_i\},$$

and then, define iteratively, for every $k \in \mathbb{N}$,

$$Y_i^{k+1} := B_i\left(\Theta^* \times \prod_{j \neq i} Y_j^k\right).$$

We claim that, for every $k \in \mathbb{N}$, Y_i^k is open and included in $Y_i^k \subseteq \mathring{L}_i^k$. The claim is trivially true for $k = 1$ and easily extends to every k by a standard iterative argument. Thus, we have that:

$$U_i := \bigcup_{k \in \mathbb{N}} Y_i^k \subseteq \bigcup_{k \in \mathbb{N}} \mathring{L}_i^k$$

is an open set that, furthermore is also dense, as every type t_i admits being approximated at increasingly many lower orders by elements in U_i. Hence, $\bigcup_{k \in \mathbb{N}} \mathring{L}_i^k$ is dense too. ■

C.3 Proof of the result

Theorem 2. *Let $(\mathscr{E}, \mathscr{T}^*)$ be a dynamic Bayesian game. Then, for any player i the following three hold:*

1. *$\mathcal{S}_i : T_i^* \rightrightarrows S_i$ is generically upper-hemicontinuous.*
2. *$\mathcal{S}_{i|T_i^0} : T_i^0 \rightrightarrows S_i$ is upper-hemicontinuous.*

Proof. We begin with the first claim. First, for each player i define set $X_i^0 := T_i^*$ which is, obviously, some generic set in which $S_{i,0}$ is upper-hemicontinuous everywhere. Next, pick $k \geq 0$ such that, for each player $j \neq i$ there exists some generic set $X_j^k \subseteq T_j^*$ in which $\mathcal{S}_{j,\ell}$ is upper-hemicontinuous everywhere for every $\ell = 0, 1, \ldots, k$ (we know that for $k = 0$ these sets exist), and define:

$$Z_i^{k+1} := \left\{ t_i' \in T_i^* \,\middle|\, \begin{array}{ll} (1) & t_i' \in B_i\left(\Theta^* \times \prod_{j \neq i} X_j^k\right) \\ (2) & \text{There exist } t_i' \in T_i \text{ and } n \in \mathbb{N} \text{ s.t. } \overline{B}_{1/n}(\Delta_i^*(t_i')) = \Delta_i^*(t_i) \end{array} \right\}.$$

Pick arbitrary $t_i \in Z_i^{k+1}$. Because of (1) above, we know that for every $t_{-i} \in \mathrm{Proj}_{T_j^*}(\Delta_i^*(t_i))$ there exists some open set $U_{-i}(t_{-i}) \subseteq T_{-i}^*$ such that $\mathcal{S}_{-i,\ell}(t_{-i}') \subseteq \mathcal{S}_{-i,\ell}(t_{-i})$ for every $t_{-i}' \in U_{-i}(t_{-i})$ and every $\ell = 0, 1, \ldots, k$. Set then:

$$U_{-i}(t_i) = \bigcup \left\{ U_{-i}(t_{-i}) \,\middle|\, t_{-i} \in \mathrm{Proj}_{T_{-i}} \Delta_i^*(t_i) \right\}.$$

Since $U_{-i}(t_i)$ is open and it contains $\mathrm{Proj}_{T_{-i}^*} \Delta_i^*(t_i)$ then we know that there exists some $\delta(t_i) > 0$ such that, for every $t_i' \in \mathring{B}_{\delta_i(t_i)}(t_i)$, $t_i' \in B_i(\Theta^* \times U_{-i}(t_i))$ and, consequently:

$$\mathrm{H}_{i,\ell}(\Delta_i^*(t_i')) \subseteq \mathrm{H}_{i,\ell}(\Delta_i^*(t_i))$$

for every $\ell = 0, 1, \ldots, k$. Now, for each $n \in \mathbb{N}$ define type:

$$\tilde{t}_i^n(t_i) := h_i^*(\overline{B}_{1/n}(\Delta_i^*(t_i)), \tau_i(t_i)).$$

Obviously, sequence $(\tilde{t}_i^n(t_i))_{n \in \mathbb{N}}$ converges to t_i and thus, it follows from the above that there exists some $N(t_i)$ such that $\tilde{t}_i^n(t_i) \in \mathring{B}_{\delta(t_i)}(t_i)$ for every $n \geq N(t_i)$. Given this, fix $t_i^n(t_i) := \tilde{t}_i^{N(t_i)+n}(t_i)$ for every $n \in \mathbb{N}$. Then, for every $n \in \mathbb{N}$ there exists some $\delta_n(t_i) > 0$ such that $\mathring{B}_{\delta_n(t_i)}(t_i^n(t_i)) \subseteq \mathring{B}_{\delta(t_i)}(t_i)$ so that for every two $t_i', t_i'' \in \mathring{B}_{\delta_n(t_i)}(t_i^n(t_i))$ we have that:

$$\mathrm{H}_{i,\ell}(\Delta_i^*(t_i')) = \mathrm{H}_{i,\ell}(\Delta_i^*(t_i'')))$$

for every $\ell = 0, 1, \ldots, k$. We can then define:

$$X_i^{k+1} := \bigcup \left\{ \bigcup \left\{ \mathring{B}_{\delta_n(t_i)}(t_i^n(t_i)) \,\middle|\, n \in \mathbb{N} \right\} \,\middle|\, t_i \in Z_i^{k+1} \right\},$$

which is clearly open. In addition, we have that:

- X_i^{k+1} is dense. To prove this, let us prove first that Z_i^{k+1} is dense. Since $\Theta^* \times \prod_{j \neq i} X_j^k$ is dense, we know by Lemma 7 that the following set is dense too:

$$\tilde{Z}_i^{k+1} := B_i\left(\Theta^* \times \prod_{j \neq i} X_j^k\right).$$

  Now, notice that Z_i^{k+1} is dense in $\tilde{Z}_i^{k+1}$ (and hence in T_i^*): it is a subset of the latter, and we can approximate every $t_i \in \tilde{Z}_i^{k+1} \backslash Z_i^{k+1}$ using sequence of types $(h_i^*(\overline{B}_{1/n}(\Delta_i^*(t_i)), \tau_i^*(t_i)))_{n\in\mathbb{N}}$ whose tail is contained in Z_i^{k+1}. Now, obviously, for every $t_i \in Z_i^{k+1}$ it is the case that sequence $(t_i^n(t_i))_{n\in\mathbb{N}}$ converges to t_i. Since the tail of this sequence is contained in X_i^{k+1}, we conclude that this set is dense in T_i^*.

- $\mathcal{S}_{i,k+1}$ is upper-hemicontinuous everywhere in X_i^{k+1}. To see this, fix $t_i \in X_i^{k+1}$ and notice first that there exists some open $U_i \subseteq T_i^*$ such that $\mathrm{H}_{i,k}(\Delta_i^*(t_i)) = \mathrm{H}_{i,k}(\Delta_i^*(t_i'))$ for every $t_i' \in U_i$. Then, pick sequence $(t_i^n)_{n\in\mathbb{N}}$ converging to t_i, strategy $s_i \in \bigcap_{n\in\mathbb{N}} \mathcal{S}_{i,k+1}(t_i^n)$, and, for each $n \in \mathbb{N}$, conjecture μ_i that justifies the inclusion of s_i in $\mathcal{S}_{i,k+1}(t_i^n)$. Pick some μ_i in the set of cluster points of the previous sequence of conjectures. Obviously, μ_i is consistent with t_i and s_i is a best-reply to it. Now, fix $N \in \mathbb{N}$ such that $t_i^n \in U_i$ for every $n \geq N$. Then, we have that, for every $\ell = 0, 1, \ldots, k$,

$$\begin{aligned}
&\forall n \geq N, \forall h \in \mathrm{H}_{i,\ell}(\Delta_i^*(t_i^n)),\ \mu_i^n(h)\,[\Theta^* \times \mathrm{Graph}(\mathcal{S}_{-i,k})] = 1 \\
&\quad\Longrightarrow\ \forall n \geq N, \forall h \in \mathrm{H}_{i,\ell}(\Delta_i^*(t_i)),\ \mu_i^n(h)\left[\Theta^* \times \mathrm{Graph}(\mathcal{S}_{-i,\ell}) \cap S_{-i}(h) \times \overline{\bigcup_{m\geq n} \Delta_i^*(t_i^m)}\right] = 1 \\
&\quad\Longrightarrow\ \forall n \geq N, \forall h \in \mathrm{H}_{i,\ell}(\Delta_i^*(t_i)),\ \mu_i(h)\left[\Theta^* \times \mathrm{Graph}(\mathcal{S}_{-i,\ell}) \cap S_{-i}(h) \times \overline{\bigcup_{m\geq n} \Delta_i^*(t_i^m)}\right] = 1 \\
&\quad\Longrightarrow\ \forall h \in \mathrm{H}_{i,\ell}(\Delta_i^*(t_i)),\ \mu_i(h)\left[\Theta^* \times \mathrm{Graph}(\mathcal{S}_{-i,\ell}) \cap S_{-i}(h) \times \bigcap_{n\geq N} \overline{\bigcup_{m\geq n} \Delta_i^*(t_i^m)}\right] = 1 \\
&\quad\Longrightarrow\ \forall h \in \mathrm{H}_{i,\ell}(\Delta_i^*(t_i)),\ \mu_i(h)\,[\Theta^* \times \mathrm{Graph}(\mathcal{S}_{-i,\ell}) \cap S_{-i}(h) \times \Delta_i^*(t_i)] = 1,
\end{aligned}$$

where: (a) the second implication is a consequence of $\mathcal{S}_{-i,\ell}$ being upper-hemicontinuous everywhere in $U_{-i}(t_i)$ and, as a result, everywhere in the projection of T_{-i}^* of the closure of $\bigcap_{m\geq n} \Delta_i^*(t_i^m)$ (remember that T_i^* is metrizable)—this implies that the graph of $\mathcal{S}_{-i,\ell}$ is closed when restricted to those types, and (b) the third implication is a consequence of $(t_i^n)_{n\in\mathbb{N}}$ and Δ_i^* being continuous. Thus, we conclude that μ_i strongly believes in $\mathcal{S}_{-i,\ell}$ for every $\ell = 0, 1 \ldots, k$ and hence, that it justifies the inclusion of s_i in $\mathcal{S}_{i,k+1}$.

We have thus concluded that, for every $k \in \mathbb{N}$ there exists some generic set $X_i^k \subseteq T_i^*$ in which $\mathcal{S}_{i,k}$ is upper-hemicontinuous everywhere. Finally, for each player i and each $k \in \mathbb{N}$ define:

$$G_i := \bigcup_{k\in\mathbb{N}} X_i^k \cap \mathring{L}_i^k$$

Obviously, G_i is open and satisfies that $\mathcal{S}_i$ is upper-hemicontinuous everywhere on it.[40] To see that it is also dense, notice that we have:

$$\begin{aligned}
\bigcup_{k\in\mathbb{N}} \overline{X_i^k \cap \mathring{L}_i^k} \subseteq \overline{G}_i = \overline{\bigcup_{k\in\mathbb{N}} X_i^k \cap \mathring{L}_i^k} \ &\Longrightarrow\ \overline{\bigcup_{k\in\mathbb{N}} \overline{X_i^k \cap \mathring{L}_i^k}} \subseteq \overline{G}_i \\
&\Longrightarrow\ \overline{\bigcup_{k\in\mathbb{N}} \overline{\mathring{L}_i^k}} \subseteq \overline{G}_i
\end{aligned}$$

[40] Simply notice that if $t_i \in X_i^k$ then there exist some open set U_i such that $\mathcal{S}_{i,k}(t_i') \subseteq \mathcal{S}_{i,k}(t_i)$ for every $t_i' \in U_i$ and if $t_i \in L_i^k$ we also have that $\mathcal{S}_{i,k}(t_i) = \mathcal{S}_i(t_i')$. Obviously, it follows that $\mathcal{S}_i(t_i') \subseteq \mathcal{S}_i(t_i)$ for every $t_i' \in U_i$.

$$\implies \overline{\bigcup_{k\in\mathbb{N}} \mathring{L}_i^k} \subseteq \overline{G}_i \implies \overline{G}_i = T_i^*$$

where the first inclusion and the first and third implication follow from elementary properties of topological closures, the second implication follows from the fact that each X_i^k is dense and each $\mathring{L}_i^k$ is open, and the last implication follows from Lemma 8.

We proceed now to the proof of the second claim. Fix a sequence of consistent profiles of types $(t^n)_{n\in\mathbb{N}}$ converging to another consistent profile of information-based types t and pick arbitrary $s \in \mathcal{S}(t^n)$. Then, we can denote:

$$z(s|h^0) = (h^0, a^1, \dots, a^L),$$

where, for each $\ell = 1, \dots, L$, a^ℓ describes the actions chosen by the players active at history $h^{\ell-1} := (h^0, a^1, \dots, a^{\ell-1})$, denoted by $I(h^{\ell-1})$. Then, we will prove the following claim: for every $\ell = 0, 1, \dots, L-1$ and every $i \in I(h^\ell)$ the following two hold:

(1) $h^\ell \in \bigcap_{k\geq 0} \mathrm{H}_{i,k}(\Delta_i^*(t_i))$

(2) There exists some $\bar{s}_i \in S_i(h^\ell) \cap \mathcal{S}_i(t_i)$ such that $\bar{s}_i(h^\ell) = a_i^{\ell+1}$.

We proceed by induction on ℓ. The claim holds trivially at the initial case ($\ell = 0$), so we can focus on the proof of the inductive step. Suppose that ℓ is such that the claim holds; let us verify then that the claims also hold for $\ell + 1$. Fix $i \in I(h^{\ell+1})$. That (1) is satisfied is an immediate consequence of part (2) of the inductive hypothesis and the fact that t is consistent (and hence $t_{-i} \in \mathrm{Proj}_{T_{-i}^*} \Delta_i^*(t_i)$). To see (2), first, for each $n \in \mathbb{N}$ pick a conjecture μ_i^n that justifies the inclusion of s_i in $\mathcal{S}_i(t_i^n)$,[41] and pick μ_i to be a cluster-point of sequence $(\mu_i^n)_{n\in\mathbb{N}}$. Obviously, μ_i is consistent with t_i. Now, pick arbitrary conjecture μ_i' that is consistent with t_i and strongly believes in $\mathcal{S}_{-i,k}$ for every $k \geq 0$, and define conjecture $\bar{\mu}_i$ by setting, for each $h \in H_i \cup \{h^0\}$,

$$\bar{\mu}_i(h) := \begin{cases} \mu_i(h) & \text{if } \mathrm{marg}_{S_{-i}} \mu_i(h')[S_{-i}(h)] > 0 \\ & \quad \text{for some } h' \in H_i \cup \{h^0\} \text{ s.t. } S_{-i}(h') \subseteq S_{-i}(h) \cap S_{-i}(h^{\ell+1}), \\ \mu_i'(h) & \text{otherwise,} \end{cases}$$

and, pick arbitrary $s_i' \in r_i(\bar{\mu}_i)$ and define strategy $\bar{s}_i$ by setting, for each $h \in H_i$,

$$\bar{s}_i(h) := \begin{cases} s_i(h) & \text{if } \mathrm{marg}_{S_{-i}} \mu_i(h')[S_{-i}(h)] > 0 \\ & \quad \text{for some } h' \in H_i \cup \{h^0\} \text{ s.t. } S_{-i}(h') \subseteq S_{-i}(h) \cap S_{-i}(h^{\ell+1}), \\ s_i'(h) & \text{otherwise.} \end{cases}$$

Then, we have that:

- $\bar{\mu}_i$ is a well-defined conjecture that is consistent with t_i. This is immediate.
- $\bar{s}_i \in r_i(\bar{\mu}_i)$. This follows from the choice of s_i' and the fact that, for every history $h \in H_i \cup \{h^0\}$ that obtains positive probability under some other history $h' \in H_i \cup \{h^0\}$ weakly preceding $h^{\ell+1}$, we have that, for every $s_i'' \in S_i$,

$$\int_{S_{-i}\times\Theta^*} \theta_i(z(s_{-i}, s_i''|h)) \mathrm{d}(\mathrm{marg}_{S_{-i}\times\Theta^*} \bar{\mu}_i(h)) = \int_{S_{-i}\times\Theta^*} \theta_i(z(s_{-i}, s_i''|h)) \mathrm{d}(\mathrm{marg}_{S_{-i}\times\Theta^*} \mu_i(h)).$$

[41] Just notice the there is some k such that $\mathcal{S}_i(t_i^n) = \mathcal{S}_{i,k}(t_i^n)$; then, pick a conjecture μ_i^n that justifies the inclusion of s_i in $\mathcal{S}_{i,k}(t_i^n.)$

- $\bar{\mu}_i$ strongly believes in $\mathcal{S}_{-i,k}$ for every $k \geq 0$. Clearly, it suffices with checking that, for every $k \geq 0$ and every history h that weakly precedes h, $\mu_i(h)$ puts probability 1 on $\Theta^* \times \text{Graph}(\mathcal{S}_{-i,k})$.

Thus, we conclude that $\bar{s}_i \in \mathcal{S}_i(t_i)$. Now, it is clear by construction of $\bar{s}_i$ that $\bar{s}_i \in S_i(h^{\ell+1})$ and $\bar{s}_i(h^{\ell+1}) = s_i(h^{\ell+1})$; hence, the proof is complete. ■

D Proof of Proposition 2

Proposition 2. *Let $(\mathscr{E}, \mathscr{T})$ be a dynamic Bayesian game where every type t_i of every player i basic. Then, for every player i the correspondence $\mathcal{S}_i^{\mathscr{T}} : T_i \rightrightarrows S_i$ is upper-hemicontinuous.*

Proof. We will prove that, for every player i and every $k \geq 0$, $\mathcal{S}_{i,k}^{\mathscr{T}} : T_i \rightrightarrows S_i$ is upper-hemicontinuous. We proceed by induction on k. The claim is trivially true for the initial case ($k = 0$), so we can focus in the proof of the inductive step: suppose that the claim holds for $k \geq 0$, and let us verify that it also does then for $k+1$. To this end, fix player i, convergent sequence of types $(t_i^n)_{n\in\mathbb{N}}$ with limit t_i, strategy $s_i \in \cap_{n\in\mathbb{N}} \mathcal{S}_{i,k+1}^{\mathscr{T}}(t_i^n)$ and sequence of conjecture $(\mu_i^n)_{n\in\mathbb{N}}$ where, for each $n \in \mathbb{N}$, μ_i^n justifies that $s_i \in \mathcal{S}_{i,k+1}^{\mathscr{T}}(t_i^n)$. Let μ_i be a cluster-point off this last sequence. Clearly, s_i is a sequential best-reply to μ_i and μ_i is consistent with t_i. Now, fix $\ell = 0, 1, \ldots, k$ and to see that μ_i strongly believes in $\mathcal{S}_{-i,\ell}$, notice first that, as $\mathscr{T}$ only consists of basic types, we have that $\text{Proj}_{T_j^*}(t_j) = \text{Proj}_{T_j^*}(t_j')$ for very player $j \in I$ and every pair of players $t_j, t_j' \in T_j$. Thus, it clearly follows that:

$$\text{H}_{i,\ell}(\Delta_i(t_i^n)) = \text{H}_{i,\ell}(\Delta_i(t_i))$$

Furthermore, notice that we know from the induction hypothesis that $\text{Graph}(\mathcal{S}_{-i,\ell}^{\mathscr{T}})$ is closed for every. Then for each $h \in \text{H}_{i,k}(\Delta_i(t_i))$ we have that:

$$\mu_i^n(h)\left[\Theta \times \text{Graph}(\mathcal{S}_{-i,\ell}^{\mathscr{T}})\right] = 1$$

what implies that:

$$\mu_i^n(h)\left[\Theta \times \text{Graph}(\mathcal{S}_{-i,\ell}^{\mathscr{T}})\right] \geq \limsup_{n\to\infty} \mu_i^n(h)\left[\Theta \times \text{Graph}(\mathcal{S}_{-i,\ell}^{\mathscr{T}})\right] = 1.$$

Hence, we conclude that μ_i strongly believes in $\mathcal{S}_{-i,\ell}^{\mathscr{T}}$ and from here, together with the above, that is justifies the inclusion of $s_i \in \mathcal{S}_{i,k+1}^{\mathscr{T}}(t_i)$. It follows then that $\mathcal{S}_i^{\mathscr{T}}$ is upper-hemicontinuous. ■

E Proof of Theorem 3

E.1 Auxiliary notation

E.1.1 For strict rationalizability

Let $(\mathscr{E}, \mathscr{T}^*)$ be dynamic Bayesian game. Then, for each player i and strategy s_i, we define the set of strategies that are *behaviorally equivalent* to s_i as:

$$[s_i] := \{s_i' \in S_i \,|\, s_i'(h) = s_i(h) \text{ for every } h \in H_i(s_i)\}$$

Given this, the *strictly rationalizable* strategies of player i are given by $\mathcal{S}_i^0 : T_i^* \rightrightarrows S_i$ where, for each type t_i we have $\mathcal{S}_i^0(t_i) := \bigcap_{k\geq 0} \mathcal{S}_{i,k}^0(t_i)$ with $\mathcal{S}_{i,0}^0(t_i) := S_i$, and, for every $k \geq 0$,

$$\mathcal{S}_{i,k+1}^0(t_i) := \left\{ s_i \in S_i \;\middle|\; \begin{array}{ll} \multicolumn{2}{l}{\text{There exists some } \mu_i \in \mathrm{C}_i^{\mathscr{T}^*}(t_i) \text{ such that:}} \\ (1) & r_i(\mu_i) = [s_i], \\ (2) & \mu_i \text{ displays strong belief in } \mathcal{S}_{-i,\ell}^0 \text{ for every } \ell = 0, 1, \dots, k \end{array} \right\}.$$

For each the set of histories of player i in which type t_i can interpret others to play according to kth-order strictly rationalizable strategies is:

$$\mathrm{H}_{i,k}^0(\Delta_i^*(t_i)) := \left\{ h \in H_i \cup \{h^0\} \,\middle|\, S_{-i}(h) \times \Delta_i^*(t_i) \cap \mathrm{Graph}(\mathcal{S}_{-i,k}^0) \times \Theta^* \neq \emptyset. \right\}.$$

E.1.2 For the first perturbation

For $n \in \mathbb{N}$, each payoff-state $\theta \in \Theta^*$, each player i and each strategy s_i let payoff-state $\theta^n[\theta, s_i] \in \Theta^*$ be defined by setting, for each player j,

$$(\theta^n[\theta, s_i])_j(z) := \begin{cases} \theta_i(z) + \frac{1}{n} & \text{if } j = 1 \text{ and } z = z((s_{-i}, s_i)|h^0) \text{ for some } s_{-i} \in S_{-i}, \\ \theta_j(z) & \text{otherwise.} \end{cases}$$

E.1.3 For the second perturbation

$$\mathcal{T}_{i,k}(s_i) := \left\{ t_i \in T_i^* \;\middle|\; \begin{array}{ll} \multicolumn{2}{l}{\text{There exists some } \mu_i \in \mathrm{C}_i^{\mathscr{T}^*}(t_i) \text{ such that:}} \\ (1) & r_i(\mu_i) = [s_i] \\ (2) & \mu_i \text{ displays strong belief in } \mathcal{S}_{-i,\ell}^0 \text{ for every } \ell = 0, 1, \dots, k \\ (3) & \mu_i(h^0)[S_{-i}(h) \times \Theta^* \times T_{-i}^*] > 0 \text{ for every } h \in \mathrm{H}_{i,k}(\Delta_i^*(t_i)) \end{array} \right\}$$

E.1.4 For the third perturbation

Based on the above, for each the set of histories of player i in which type t_i can interpret others to play according to kth-order weakly rationalizable strategies, is:

$$\mathrm{H}_{i,k}^{\mathcal{W}}(\Delta_i^*(t_i)) := \left\{ h \in H_i \cup \{h^0\} \,|S_{-i}(h) \times \Delta_i^*(t_i) \cap \mathrm{Graph}(\mathcal{W}_{-i,k}) \times \Theta^* \times \neq \emptyset. \right\}.$$

Based on this, we consider the set of strategies that are behaviorally equivalent to some strategy s_i only at those histories that a type t_i considers reachable by $(k-1)$th-order weakly rationalizable strategies of others:

$$[s_i|t_i]_k := \left\{ s_i' \in S_i \,\middle|\, s_i'(h) = s_i(h) \text{ for every } h \in \mathrm{H}_{i,k}^{\mathcal{W}}(\Delta_i^*(t_i)) \cap H_i(s_i) \right\}.$$

E.2 Auxiliary lemmas

E.2.1 Availability of finite conjectures

Lemma 9. *Let $(\mathscr{E}, \mathscr{T}^*)$ be a dynamic Bayesian game. Then for any $k \geq 0$, any player i, any finite type $t_i \in X_i^k$ and every strategy $s_i \in \mathcal{S}_{i,k}(t_i)$ there exists a finite conjecture μ_i that justifies the inclusion of s_i in $\mathcal{S}_{i,k}(t_i)$.*

Proof. Fix $k \geq 0$, player i, finite type X_i^{k+1} and strategy $s_i \in \mathcal{S}_{i,k+1}(t_i)$ and pick conjecture μ_i that justifies the inclusion of s_i in $\mathcal{S}_{i,k+1}(t_i)$. For convenience, let us denote $Z_i := \Theta_0 \times T_{-i}^*$. Since

Z_i is separable, we can pick countable $(x_i^m)_{m\in\mathbb{N}}$ where $\{x_i^m\}_{m\in\mathbb{N}}$ is dense in Z_i. Now, for each $m, n \in \mathbb{N}$ let $B_{1/n}(x_i^m)$ denote the ball of radius $1/n$ around x_i^m. We know because of the upper hemicontinuity of $\mathcal{S}_{-i,\ell}$ in X_{-i}^k for each $\ell = 0, 1, \ldots, k$ that for each $n, m \in \mathbb{N}$ there exists some open set $W_i^{\ell,n,m} \subseteq B_{1/n}(x_i^m)$ such that $\mathcal{S}_{-i,\ell}(t_{-i}) \subseteq \mathcal{S}_{-i,\ell}(t_{-i}^m)$ for every t_{-i} is in the projection of $W_i^{\ell,n,m}$ on T_{-i}^*, t_{-i}^m being the projection on T_{-i}^* of x_i^m. Furthermore, for each $\ell = 0, 1, \ldots, k$ and $n \in \mathbb{N}$, the family $\{W_i^{\ell,n,m}\}_{m\in\mathbb{N}}$ is an open cover of Z_i, which, due to Z_i being compact, we can assume as finite: $\{W_i^{\ell,n,m}\}_{m=1}^{M_{\ell,n}}$. Now, for each $\ell = 0, 1, \ldots, k$ and $n \in \mathbb{N}$ set:

$$V_i^{\ell,n,m} := W_i^{\ell,n,m} \setminus \bigcup_{r=1}^{m-1} W_i^{\ell,n,r}.$$

Notice that for each $\ell = 0, 1, \ldots, k$ and $n \in \mathbb{N}$ family $\{V_i^{\ell,n,m}\}_{m=1}^{M_{\ell,n}}$ is a partition of Z_i consisting of measurable sets, contained in a ball of radius $1/n$. Since the set of finite types is dense in Z_i for each $\ell = 0, 1, \ldots, k$ and $n \in \mathbb{N}$, there exists some list $(y_i^{\ell,n,m})_{m=1}^{M_{\ell,n}}$ such that, for each $m = 1, \ldots, M_{\ell,n}$, the projection on T_{-i}^* of $y_i^{\ell,n,m}$ is finite and $y_i^{\ell,n,m} \in V_i^{\ell,n,m}$.

We turn now back to μ_i. For every $\ell = 0, 1, \ldots, k$ set $\mathrm{H}_{i,\ell}(t_i, \mu_i) := \mathrm{H}_i(\mu_i) \cap \mathrm{H}_{i,\ell}(\Delta_i^*(t_i))$ and, for notational convenience, set $\mathrm{H}_{i,k+1}(t_i, \mu_i) = \emptyset$. Next, we will construct a conditional probability system μ_i^n for each $n \in \mathbb{N}$. First, for each each $\ell = 0, 1, \ldots, k$ and each $h \in \mathrm{H}_{i,\ell}(t_i, \mu_i) \setminus \mathrm{H}_{i,\ell+1}(t_i, \mu_i)$ define:

$$\mu_i^n(h)[(s_{-i}, y_i^{\ell,m})] := \mu_i(h)[\{s_{-i}\} \times V_i^{\ell,n,m}],$$

for every $s_{-i} \in S_{-i}$ and every $m = 1, \ldots, M_{\ell,n}$. Second, set $\mu_i^n(h^0) := \mu_i(h^0)$. Finally, for each $h \notin \mathrm{H}_i(\mu_i)$ define $\mu_i^n(h)$ via conditional probability. Notice that the marginals on S_{-i} of $\mu_i(h)$ and each $\mu_i^n(h)$ coincide for every history h, and this guarantees that μ_i^n is (or, has been) well-defined. Now, notice also that, for each $\ell = 0, 1, \ldots, k$, $\mu_i^n(h)$ assigns probability one to the graph of $\mathcal{S}_{-i,\ell}$. Obviously, every μ_i^m is consistent with type t_i, and sequence $(\mu_i^n)_{n\in\mathbb{N}}$ converges to μ_i. Thus, the upper hemicontinuity of r_i ensures the existence of some $N \in \mathbb{N}$ such that $s_i \in r_i(\mu_i^n)$ for every $n \geq N$. Hence, every μ_i^n where $n \geq N$ is a *finite* conjecture that justifies the inclusion of s_i in $\mathcal{S}_{i,k+1}(t_i)$. ∎

E.2.2 First perturbation

LEMMA 10. *Let $(\mathscr{E}, \mathscr{T}^*)$ be a dynamic Bayesian game. Then, for any player i, any finite type t_i, any $k \in \mathbb{N}$ and any $s_i \in \mathcal{S}_{i,k}(t_i)$ there exists a sequence of finite types $(t_i^n)_{n\in\mathbb{N}}$ converging to t_i such that, for every $n \in \mathbb{N}$, the following hold:*

(1) $\mathrm{H}_{i,k-1}^0(\Delta_i^*(t_i^n)) = \mathrm{H}_{i,k-1}(\Delta_i^*(t_i^n)) = \mathrm{H}_{i,k-1}(\Delta_i^*(t_i))$

(2) $s_i \in \mathcal{S}_{i,k}^0(t_i^n)$.

Proof. Let us begin proving the following claim: for any $k \in \mathbb{N}$ and any finite $t_i \in X_i^k$ (as defined in the proof of part 1 of Theorem 2) and any $s \in \mathcal{S}_{i,k}(t_i)$ there exists a sequence of finite types $(t_i^n(t_i, k, s_i))_{n\in\mathbb{N}}$ converging to t_i such that the following hold:

(1) $\mathrm{H}_{i,k-1}^0(\Delta_i^*(t_i^n(t_i, k, s_i))) = \mathrm{H}_{i,k-1}(\Delta_i^*(t_i^n(t_i, k, s_i))) = \mathrm{H}_{i,k-1}(\Delta_i^*(t_i))$

(2) $s_i \in \mathcal{S}_{i,k}^0(t_i^n(t_i, k, s_i))$.

(3) $\mathcal{S}_{i,k}(t_i^n(t_i, k, s_i)) \subseteq \mathcal{S}_{i,k}(t_i)$.

(4) $\mathcal{S}_{i,k}^0(t_i^n(t_i, k, s_i)) \subseteq \mathcal{S}_{i,k+1}(t_i^n(t_i, k, s_i))$.

We proceed by induction on k. To verify the claim for the initial case ($k = 1$), fix player i and finite type $t_i \in X_i^1$, and pick strategy $s_i \in \mathcal{S}_{i,1}(t_i)$ and *finite* conjecture μ_i that justifies this inclusion.[42] Notice that, as $t_i \in X_i^1$ there exists some open set $U_i(t_i) \subseteq X_i^1$ such that $\Delta_i^*(t_i') \subseteq X_{-i}^0$, $\mathrm{H}_{i,0}(\Delta_i^*(t_i')) = \mathrm{H}_{i,0}(\Delta_i^*(t_i))$ and $\mathcal{S}_{i,1}(t_i') \subseteq \mathcal{S}_{i,1}(t_i)$ for every $t_i' \in U_i(t_i)$. Then, for every $n \in \mathbb{N}$ define:

- $\Delta^n(t_i, 1) := \overline{\bigcup\{B_{1/n}(\theta) \times \{t_{-i}\} | (\theta, t_{-i}) \in \Delta_i^*(t_i)\}}$.
- For every $h \in \mathrm{H}_i(\mu_i)$ and every (s_{-i}, θ, t_{-i}) in the support of $\mu_i(h)$ set:

$$\hat{\mu}_i^n(h)[(s_{-i}, \theta^n[\theta, s_{-i}], t_{-i})] := \mu_i(h)[(s_{-i}, \theta, t_{-i})],$$

  The finiteness of $\hat{\mu}_i^n$ guarantees that the above induces a well-defined measure $\hat{\mu}_i^n(h) \in \Delta(S_{-i} \times \Theta^* \times T_{-i}^*)$. For every $h \notin \mathrm{H}_{i,k}(\mu_i)$ define $\hat{\mu}_i^n(h)$ via conditional probability. Obviously, it holds that $r_i(\hat{\mu}_i^n) = [s_i]$ for every $n \in \mathbb{N}$.
- $\hat{t}_i^n(t_i, 1, s_i) := h_i^*(\Delta^n(t_i, 1), \mathrm{marg}_{\Theta^* \times T_{-i}^*} \hat{\mu}_i^n(h^0))$ which is, clearly, finite.

It is immediate that $(\hat{t}_i^n(t_i, 1, s_i))_{n \in \mathbb{N}}$ converges to t_i and thus, that there exists some $N \in \mathbb{N}$ such that $\hat{t}^n(t_i, 1, s_i) \in U_i(t_i)$ for every $n \geq N$. Redefine then by setting: $t^n(t_i, 1, s_i) := \hat{t}^{n+N}(t_i, 1, s_i)$ and $\mu_i^n := \hat{\mu}_i^{n+N}$ for every $n \in \mathbb{N}$. We know verify that (1), (2), (3), and (4) are satisfied for every $n \in \mathbb{N}$:

(1) Since $\mathcal{S}_{j,0}^0(t_j) = \mathcal{S}_{j,0}(t_j) = S_j$ for every $j \neq i$ and every $t_j \in T_j^*$,

$$\mathrm{H}_{i,0}^0(\Delta_i^*(t_i^n(t_i, 1, s_i))) = \mathrm{H}_{i,0}(\Delta_i^*(t_i^n(t_i, 1, s_i))) = \mathrm{H}_{i,0}(\Delta_i^*(t_i)).$$

(3) $\mu_i^n \in \mathrm{C}_i^{\mathscr{T}^*}(t_i^n(t_i, 1, s_i))$ is a conjecture that strongly believes in $\mathcal{S}_{-i,0}^0$. Since $r_i(\mu_i^n) = [s_i]$, we conclude that $s_i \in \mathcal{S}_{i,1}^0(t^n(t_i, 1, s_i))$.

(4) Since $t_i^n(t_i, 1, s_i) \in U_i(t_i)$ we know that $\mathcal{S}_{i,1}(t^n(t_i, 1, s_i)) \subseteq \mathcal{S}_{i,1}(t_i)$.

(5) Pick arbitrary $s_i' \in \mathcal{S}_{i,1}^0(t_i^n(t_i, 1, s_i))$ and conjecture μ_i' that justifies this inclusion. Obviously, μ_i' strongly believes in $\mathcal{S}_{-i,0}$ and thus, we conclude that it also justifies the inclusion of s_i' in $\mathcal{S}_{i,1}(t_i^n(t_i, 1, s_i))$. Hence, $\mathcal{S}_{i,1}^0(t_i^n(t_i, 1, s_i)) \subseteq \mathcal{S}_{i,1}(t^n(t_i, 1, s_i))$.

We proceed now to the proof of the inductive step. Let $k \geq 1$ be such that the claim holds, and let us verify that, then, it also does for $k+1$. Fix player i and finite type $t_i \in X_i^{k+1}$, and pick strategy $s_i \in \mathcal{S}_{i,k+1}(t_i)$ and *finite* conjecture μ_i that justifies this inclusion. Notice that, as $t_i \in X_i^{k+1}$ there exists some open set $U_i(t_i) \subseteq X_i^{k+1}$ such that $\Delta_i^*(t_i') \subseteq X_{-i}^k$, $\mathrm{H}_{i,\ell}(\Delta_i^*(t_i')) = \mathrm{H}_{i,\ell}(\Delta_i^*(t_i))$ for every $\ell = 0, 1, \ldots, k$, and $\mathcal{S}_{i,k+1}(t_i') \subseteq \mathcal{S}_{i,k+1}(t_i)$ for every $t_i' \in U_i(t_i)$. Then, for every $n \in \mathbb{N}$ define:

- $\Delta^n(t_i, k+1) := \overline{\{(\theta, t_{-i}^n(t_{-i}, k, s_{-i})) | (\theta, t_{-i}) \in \Delta_i^*(t_i) \text{ and } s_{-i} \in \mathcal{S}_{-i,k}(t_{-i})\}}$.
- For every $\ell = 0, 1, \ldots, k$ set $H_i^\ell := \mathrm{H}_{i,\ell}(\Delta_i^*(t_i)) \cap \mathrm{H}_i(\mu_i)$ (and, for convenience, $H_i^{k+1} = \emptyset$) and, for each $h \in H_i^\ell \setminus H_i^{\ell+1}$ define:

$$\hat{\mu}_i^n(h)[E] := \mu_i(h)\left[\left\{(s_{-i}, \theta, t_{-i}) \in \mathrm{supp}(\mu_i(h)) \;\middle|\; \begin{array}{ll} \multicolumn{2}{l}{(s_{-i}, \theta', t_{-i}') \in E \text{ for:}} \\ (a) & \theta' = \theta^n[\theta, s_i], \\ (b) & t_{-i}' = t_{-i}^n(s_{-i}, \ell, t_{-i}) \end{array}\right\}\right],$$

  for every measurable $E \subseteq S_{-i} \times \Theta^* \times T_{-i}^*$. The finiteness of $\hat{\mu}_i^n$ guarantees that the above induces a well-defined measure $\hat{\mu}_i^n(h) \in \Delta(S_{-i} \times \Theta^* \times T_{-i}^*)$. For every $h \notin \mathrm{H}_{i,k}(\mu_i)$ define $\hat{\mu}_i^n(h)$ via conditional probability. Obviously, it holds that $r_i(\hat{\mu}_i^n) = [s_i]$ for every $n \in \mathbb{N}$.

[42] The existence of this finite conjecture is ensured by Lemma 9.

- $\hat{t}_i^n(t_i, k+1, s_i) := h_i^*(\Delta^n(t_i, k+1), \mathrm{marg}_{\Theta^* \times T_{-i}^*} \hat{\mu}_i^n(h^0))$, which is, clearly, finite.

It is immediate that $(\hat{t}_i^n(t_i, k+1, s_i))_{n\in\mathbb{N}}$ converges to t_i and thus, that there exists some $N \in \mathbb{N}$ such that $\hat{t}^n(t_i, k+1, s_i) \in U_i(t_i)$ for every $n \geq N$. Redefine then by setting: $t^n(t_i, k+1, s_i) := \hat{t}^{n+N}(t_i, k+1, s_i)$ and $\mu_i^n := \hat{\mu}_i^{n+N}$ for every $n \in \mathbb{N}$. We know verify that (1), (2), (3), and (4) are satisfied for every $n \in \mathbb{N}$:

(1) Notice first that, for every $j \neq i$, as $\mathrm{Proj}_{T_j^*} \Delta_i^n(t_i, k+1) \subseteq X_j^k$, we know that $\mathcal{S}_{j,k}$ is upper-hemicontinuous at every $t_j \in \mathrm{Proj}_{T_j^*} \Delta_i^n(t_i, k+1)$, and thus, we have that:

$$\bigcup \left\{ \mathcal{S}_{j,k}(t_j) \,\middle|\, t_j \in \mathrm{Proj}_{T_j^*} \Delta_i^n(t_i, k+1) \right\} = \bigcup \left\{ \mathcal{S}_{j,k}(t_j) \,\middle|\, t_j \in \mathrm{Proj}_{T_j^*} \overbrace{\Delta_i^n(t_i, k+1)}^{\circ} \right\},$$

and part (4) of the induction hypothesis implies that:

$$\bigcup \left\{ \mathcal{S}_{j,k}(t_j) \,\middle|\, t_j \in \mathrm{Proj}_{T_j^*} \overbrace{\Delta_i^n(t_i, k+1)}^{\circ} \right\} \subseteq \bigcup \left\{ \mathcal{S}_{j,k}(t_j) \,\middle|\, t_j \in \mathrm{Proj}_{T_j^*} \overbrace{\Delta_i^*(t_i)}^{\circ} \right\}.$$

In addition, it follows from part (3) of the induction hypothesis that:

$$\bigcup \left\{ \mathcal{S}_{j,k}(t_j) \,\middle|\, t_j \in \mathrm{Proj}_{T_j^*} \overbrace{\Delta_i^*(t_i)}^{\circ} \right\} \subseteq \bigcup \left\{ \mathcal{S}_{j,k}^0(t_j) \,\middle|\, t_j \in \mathrm{Proj}_{T_j^*} \overbrace{\Delta_i^n(t_i, k+1)}^{\circ} \right\},$$

and finally, it is a trivial truth that:

$$\bigcup \left\{ \mathcal{S}_{j,k}^0(t_j) \,\middle|\, t_j \in \mathrm{Proj}_{T_j^*} \overbrace{\Delta_i^n(t_i, k+1)}^{\circ} \right\} \subseteq \bigcup \left\{ \mathcal{S}_{j,k}^0(t_j) \,\middle|\, t_j \in \mathrm{Proj}_{T_j^*} \Delta_i^n(t_i, k+1) \right\}.$$

Obviously, it follows that:

$$\mathrm{H}_{i,k}(\Delta_i^*(t_i^n(t_i, k+1, s_i)) \subseteq \mathrm{H}_{i,k}^0(\Delta_i^*(t_i^n(t_i, k+1, s_i)),$$

and since $t_i^n(t_i, k+1, s_i) \in U_i(t_i)$, we know that:

$$\mathrm{H}_{i,k}(\Delta_i^*(t_i)) = \mathrm{H}_{i,k}(\Delta_i^*(t_i^n(t_i, k+1, s_i))).$$

Now, before continuing with the argument, we need to make first a technical observation: for any sequence of types $(t_j^m)_{m\in\mathbb{N}}$ that are consistent with $\Delta_i^*(t_i)$ and converge to some t_j we have that $(t_j^n(t_j^m, k, s_j))_{m\in\mathbb{N}}$, where $s_j \in \bigcap_{m\in\mathbb{N}} \mathcal{S}_{j,k}(t_j^m)$, converges to $t_j^n(t_j, k, s_j)$—this is an artifact of t_j being consistent with $\Delta_i^*(t_i)$ (as the latter is closed) of $\mathcal{S}_{j,k}$ being upper-hemicontinuous at every t_j (as $\mathrm{Proj}_{T_j^*} \Delta_j^*(t_j)$ is contained in X_j^k), and the construction of each $t^n(\,\cdot\,, k, s_j)$. In turn, this technical feature, together with the compactness of $\Delta_i^*(t_i)$ implies that, if $(t_j^n(t_j^m, k, s_j^m))_{m\in\mathbb{N}}$ is a sequence where every t_j^m is consistent with $\Delta_i^*(t_i)$ and $s_j^m \in \mathcal{S}_{j,k}(t_j^m)$, that converges to some type type t_j, when there exists some t_j' consistent with $\Delta_i^*(t_i)$ and some $s_j \in \mathcal{S}_{j,k}(t_j')$ such that $t_j = t_j^n(t_j', k, s_j)$.

It turns out that this last observation implies that:

$$\bigcup \left\{ \mathcal{S}_{j,k}^0(t_j) \,\middle|\, t_j \in \mathrm{Proj}_{T_j^*} \Delta_i^n(t_i, k+1) \right\} = \bigcup \left\{ \mathcal{S}_{j,k}^0(t_j) \,\middle|\, t_j \in \mathrm{Proj}_{T_j^*} \overbrace{\Delta_i^n(t_i, k+1)}^{\circ} \right\},$$

and thus, it follows from part (5) of the induction hypothesis and the upper-hemicontinuity on X^k_{-i} of $\mathcal{S}_{-i,k}$ that:

$$\bigcup\left\{\mathcal{S}^0_{j,k}(t_j)\,\middle|\,t_j \in \mathrm{Proj}_{T^*_j}\Delta^n_i(t_i,k+1)\right\} \subseteq \bigcup\left\{\mathcal{S}_{j,k}(t_j)\,\middle|\,t_j \in \mathrm{Proj}_{T^*_j}\Delta^n_i(t_i,k+1)\right\}.$$

Together with the above, the latter clearly implies that:

$$\mathrm{H}^0_{i,k}(\Delta^*_i(t^n_i(t_i,k+1,s_i)) = \mathrm{H}_{i,k}(\Delta^*_i(t^n_i(t_i,k+1,s_i))) = \mathrm{H}_{i,k}(\Delta^*_i(t_i)).$$

(2) It follows from (1) that μ^n_i strongly believes in $\mathcal{S}^0_{-i,\ell}$ for every $\ell = 0,1,\ldots,k$. Since $\mu^n_i \in \mathrm{C}^{\mathscr{T}^*}_i(t^n_i(t_i,k+1,s_i))$ and $r_i(\mu^n_i) = [s_i]$, we conclude that $s_i \in \mathcal{S}^0_{i,k+1}(t^n(t_i,k+1,s_i))$.

(3) Since $t^n_i(t_i,k+1,s_i) \in U_i(t_i)$ we know that $\mathcal{S}_{i,k+1}(t^n(t_i,k+1,s_i)) \subseteq \mathcal{S}_{i,k+1}(t_i)$.

(4) Pick arbitrary $s'_i \in \mathcal{S}^0_{i,k+1}(t^n_i(t_i,k+1,s_i))$ and conjecture μ'_i that justifies this inclusion. Notice that the technical observation of the previous step implies that μ'_i only assigns positive probability to types $t^n_j(t_j,\ell,s_j)$ where t_j is in X^ℓ_j and s_j is in $\mathcal{S}_{j,k}(t_j)$ and thus, it follows form part (1) above and part (4) of the induction hypothesis that μ'_i strongly believes in $\mathcal{S}_{-i,\ell}$ for every $\ell = 0,1,\ldots,k$. Thus, we conclude that $s'_i \in \mathcal{S}_{i,k}(t^n_i(t_i,k+1),s_i)$. Hence, $\mathcal{S}^0_{i,k}(t^n_i(t_i,k+1),s_i) \subseteq \mathcal{S}_{i,k}(t^n_i(t_i,k+1),s_i)$.

Now, to prove the claim of the lemma notice first that we know from the proof of part 1 of Theorem 2 that X^k_i is dense for every $k \geq 0$. Then, for any finite $t_i \in T^*_i$ and any $k \geq 0$ there exists some sequence of finite types $(t^{k,n}_i)_{n\in\mathbb{N}}$ converging to t_i such that $t^{k,n}_i \in X^k_i$ for every $n \in \mathbb{N}$.[43] Now, as seen above, we know that for every $n \in \mathbb{N}$ there exists a sequence of finite types $(t^{n,m}_i)_{m\in\mathbb{N}}$ converging to $t^{k,n}_i$ and such that the following hold:

(1) $\mathrm{H}^0_{i,k-1}(\Delta^*_i(t^{n,m}_i)) = \mathrm{H}_{i,k-1}(\Delta^*_i(t^{n,m}_i)) = \mathrm{H}_{i,k-1}(\Delta^*_i(t_i))$

(2) $s_i \in \mathcal{S}^0_{i,k}(t^{n,m}_i)$

for every $m \in \mathbb{N}$. Thus, if for every $m \in \mathbb{N}$ we set $t^n_i := t^{n,n}_i$ then we have that $(t^n_i)_{n\in\mathbb{N}}$ is a sequence of finite types converging to t_i and such that the following hold:

(1) $\mathrm{H}^0_{i,k-1}(\Delta^*_i(t^n_i)) = \mathrm{H}_{i,k-1}(\Delta^*_i(t^n_i)) = \mathrm{H}_{i,k-1}(\Delta^*_i(t_i))$

(2) $s_i \in \mathcal{S}^0_{i,k}(t^n_i)$

for every $n \in \mathbb{N}$. ■

E.2.3 Second perturbation

Lemma 11. *Let $(\mathscr{E},\mathscr{T}^*)$ be a dynamic Bayesian game. Then, for any $k \in \mathbb{N}$, any player i, any finite type $t_i \in T^*_i$ and any strategy $s_i \in \mathcal{S}^0_{i,k}(t_i)$ there exists a sequence of finite types $(t^n_i)_{n\in\mathbb{N}}$ converging to t_i such that $\Delta^*_i(t^n_i) = \Delta^*_i(t_i)$ and $t^n_i \in \mathcal{T}_{i,k}(s_i)$ for every $n \in \mathbb{N}$.*

Proof. Fix $k \in \mathbb{N}$, player i, finite type $t_i \in T^*_i$ and strategy $s_i \in \mathcal{S}^0_{i,k}(t_i)$, and pick conjecture μ_i that justifies the inclusion of s_i in $\mathcal{S}^0_{i,k}(t_i)$. Then, for each $m \in \mathbb{N}$ set:[44]

$$\mu^m_i(h^0) = \left(1-\frac{1}{m}\right)\mu_i(h^0) + \left(\frac{1}{m}\right)\sum_{h\in\mathrm{H}^0_{i,k-1}(t_i,\mu_i)}\left(\frac{1}{|\mathrm{H}^0_{i,k-1}(t_i,\mu_i)|}\right)\mu_i(h).$$

[43] That the components of the sequence can be assumed to be finite is a consequence of the set of finite types being dense in X^k_i

[44] Remember that $\mathrm{H}^0_{i,k-1}(t_i,\mu_i) = \mathrm{H}^0_{i,k-1}(\Delta_i(t^*_i)) \cap \mathrm{H}_i(\mu_i)$.

Clearly, $\mu_i^m(h^0)$ is well-defined element of $\Delta(S_{-i} \times \Theta^* \times T^*_{-i})$ with finite support. Now, for each $m \in \mathbb{N}$ denote:

$$H_i^m = \left\{ h \in H_i \cup \{h^0\} \,\middle|\, \left(\mathrm{marg}_{S_{-i}} \mu_i^m(h^0)\right) [S_{-i}(h)] > 0 \right\},$$

and, for any $h \in H_i^m$ define $\mu_i^m(h) \in \Delta(S_{-i} \times \Theta^* \times T^*_{-i})$ by extending $\mu_i^m(h^0)$ via conditional probability. Finally, let conjecture μ_i^m and type $\hat{t}_i^m$ be respectively defined as follows:

$$\mu_i^m(h) := \begin{cases} \mu_i^m(h) & \text{for every } h \in H_i^m, \\ \mu_i(h) & \text{otherwise,} \end{cases} \quad \text{and} \quad \hat{t}_i^m := h_i^* \left(\Delta_i^*(t_i), \mathrm{marg}_{\Theta^* \times T^*_{-i}} \mu_i^m(h^0)\right).$$

Obviously, $(\hat{t}_i^m)_{m \in \mathbb{N}}$ is a sequence of finite types converging to t_i. We claim that the following four properties are satisfied:

(*a*) For any $m \in \mathbb{N}$ the marginal on S_{-i} of $\mu_i^m(h^0)$ assigns positive probability to $S_{-i}(h)$ for every history $h \in \mathrm{H}^0_{i,k-1}(\Delta_i^*(t_i))$. To see it, fix such h and distinguish two cases. First, if $h \in \mathrm{H}^0_{i,k-1}(t_i, \mu_i)$, then:

$$\begin{aligned}(\mathrm{marg}_{S_{-i}} \mu_i^m(h^0)) [S_{-i}(h)] &\geq \left(\frac{1}{m}\right) \cdot \left(\frac{1}{|\mathrm{H}^0_{i,k-1}(t_i, \mu_i)|}\right) \cdot (\mathrm{marg}_{S_{-i}} \mu_i(h)) [S_{-i}(h)] \\ &= \left(\frac{1}{m}\right) \cdot \left(\frac{1}{|\mathrm{H}^0_{i,k-1}(t_i, \mu_i)|}\right) > 0.\end{aligned}$$

Second, if $h \notin \mathrm{H}^0_{i,k-1}(t_i, \mu_i)$, then there must exist some history $h' \in \mathrm{H}^0_{i,k-1}(t_i, \mu_i)$ that precedes h and such that $(\mathrm{marg}_{S_{-i}} \mu_i^m(h')) [S_{-i}(h)] > 0$, and thus:

$$(\mathrm{marg}_{S_{-i}} \mu_i^m(h^0)) [S_{-i}(h)] \geq \left(\frac{1}{m}\right) \cdot \left(\frac{1}{|\mathrm{H}^0_{i,k-1}(t_i, \mu_i)|}\right) \cdot (\mathrm{marg}_{S_{-i}} \mu_i(h')) [S_{-i}(h)] > 0.$$

(*b*) For any $m \in \mathbb{N}$, μ_i^m is a well-defined conjecture consistent with $\hat{t}_i^m$. Consistency with t_i^m holds by construction (or the mere fact that t_i and $\hat{t}_i^m$ have the same information); thus, all we need to check is that μ_i^m does not violate conditional update. This is easy to see by simply noticing that for any pair of different histories $h, h' \in \mathrm{H}^0_{i,k-1}(t_i, \mu_i)$ the marginal on S_{-i} of $\mu_i(h)$ puts zero probability on $S_{-i}(h')$. This implies that for any $h, h' \in H_i^m$ there are no inconsistency issues arising from belief update.[45] Clearly, there are no problems either for any pair $h, h' \notin H_i^m$ (due to μ_i being a conditional probability system). Since for any $h \in H_i^m$ the marginal on S_{-i} of $\mu_i(h)$ puts zero probability on $S_{-i}(h')$ for any $h' \notin H_i^m$, it follows that pairs $h \in H_i^m$ and $h' \notin H_i^m$ are not problematic either.

(*c*) μ_i strongly believes in $\mathcal{S}^0_{-i,\ell}$, for every $\ell = 0, \ldots, k-1$. For $\ell = k-1$, notice that it holds by construction that:

$$\mathrm{supp}\ \mu_i^m(h^0) \subseteq \mathrm{Graph}\left(\mathcal{S}^0_{-i,\ell}\right),$$

and thus we have that, for every $h \in \mathrm{H}^0_{i,\ell}(t_i, \mu_i)$,

$$\mathrm{supp}\ \mu_i^m(h) \subseteq \mathrm{supp}\ \mu_i^m(h^0) \subseteq \mathrm{Graph}\left(\mathcal{S}^0_{-i,\ell}\right).$$

[45] Because if $h, h' \in H_i^m$ then there exist $\bar{h}, \bar{h}' \in \mathrm{H}^0_{i,k-1}(t_i, \mu_i)$ such that $\mu_i(\bar{h})$ induces $\mu_i^m(h)$ and $\mu_i(\bar{h}')$ induces $\mu_i^m(h')$. Since the marginal on S_{-i} of $\mu_i(\bar{h})$ (resp. $\mu_i(\bar{h}')$) puts zero probability on $S_{-i}(\bar{h}')$ (resp. $S_{-i}(\bar{h})$), it follows that the marginal on S_{-i} of $\mu_i^m(h)$ (resp. $\mu_i^m(h')$) puts zero probability on $S_{-i}(h')$ (resp. $S_{-i}(h)$).

Now, for every $\ell < k-1$ we have that, if $h \in \mathrm{H}^0_{i,\ell}(t_i,\mu_i) \setminus \mathrm{H}^0_{i,k-1}(t_i,\mu_i)$ that:

$$\text{supp } \mu_i^m(h) = \text{supp } \mu_i(h) \subseteq \text{Graph}\left(\mathcal{S}^0_{-i,\ell}\right).$$

(*d*) There exists some $m_0 \in \mathbb{N}$ such that $r_i(\mu_i^m) = [s_i]$ for every $m \geq m_0$. This follows trivially from the finiteness of S_i, the continuity of the conditional expected utilities and the fact that $(\mu_i^m)_{m\in\mathbb{N}}$ converges to μ_i.

This way, we conclude that $s_i \in \mathcal{S}^0_{i,k}(\hat{t}_i^m)$ for every $m \geq m_0$. Hence, if for each $n \in \mathbb{N}$ we relabel as $t_i^n := \hat{t}_i^{n+m_0}$, $(t_i^n)_{n\in\mathbb{N}}$ is the sequence we are looking for. ■

E.2.4 Third perturbation

LEMMA 12. *Let $(\mathscr{E},\mathscr{T}^*)$ be a dynamic Bayesian game. Then for any $k \geq 2$, any player i, any mildly consistent, finite type t_i, and any strategy $s_i \in \mathcal{S}^0_{i,k}(t_i)$ such that $t_i \in \mathcal{T}^0_{i,k}(s_i)$ there exists a finite type t_i^k such that:*

(*i*) $\mu_i^k(t_i^k) = \mu_i^k(t_i)$.

(*ii*) $\mathcal{W}_{i,k+1}(t_i^k) \subseteq [s_i|t_i]_{k-1}$.

Proof. We proceed by induction on k. To verify that the claims hold for the initial case ($k=1$) fix player i, mildly consistent, finite type $\bar{t}_i$, strategy $\bar{s}_i \in \mathcal{S}^0_{i,1}(\bar{t}_i)$ such that $\bar{t}_i \in \mathcal{T}_{i,1}(\bar{s}_i)$, and conjecture $\bar{\mu}_i$ that justifies the inclusion of $\bar{t}_i$ in $\mathcal{T}_{i,1}(\bar{s}_i)$ (and thus the inclusion of $\bar{s}_i$ in $\mathcal{S}^0_{i,1}(\bar{t}_i)$).

Now, for each $j \neq i$ and each $s_j \in S_j$ consider state $\theta^{s_j} \in \Theta^*$ in which s_j is conditionally dominant for player j, and type $t_j^{s_j}$ that persistently (and thus also initially) believes in θs_j—i.e., such that $\Delta^1_j(t_j^{s_j}) = \{\theta^{s_j}\}$. Obviously, it holds that $\mathcal{S}_{j,1}(t_j^{s_j}) = [s_j]$. Define then conjecture μ_i^1 by setting:

$$\bar{\mu}_i^1(h^0)[(s_{-i},\theta,t_{-i}^{s_{-i}})] := \bar{\mu}_i(h^0)[\{(s_{-i},\theta)\}\times T^*_{-i}]$$

for every $(s_{-i},\theta) \in S_{-i}\times\Theta^*$ (the finiteness of $\bar{t}_i$ guarantees that $\bar{\mu}_i^{k+1}$ is well-defined). For every $h \in \mathrm{H}^0_{i,0}(\Delta_i(\bar{t}_i)) \setminus \{h^0\}$ define $\bar{\mu}_i^1(h) \in \Delta(S_{-i}\times\Theta^*\times T^*_{-i})$ via conditional probability; notice then that, for every $h \in \mathrm{H}^0_{i,0}(\Delta_i(\bar{t}_i))$ and every measurable $E \subseteq S_{-i}\times\Theta^*$, we have that $\bar{\mu}_i^1(h)[E\times T^*_{-i}] = \bar{\mu}_i(h)[E\times T^*_{-i}]$, and thus, for every $h \in \mathrm{H}^0_{i,0}(\Delta_i(\bar{t}_i))$ and any every s_i it holds that:

$$\int_{S_{-i}\times\Theta^*}\theta_i(z(s_{-i},s_i|h))\mathrm{d}(\mathrm{marg}_{S_{-i}\times\Theta^*}\bar{\mu}_i^1(h)) = \int_{S_{-i}\times\Theta^*}\theta_i(z(s_{-i},s_i|h))\mathrm{d}(\mathrm{marg}_{S_{-i}\times\Theta^*}\bar{\mu}_i(h)).$$

Finally, define compact set:

$$\Delta_i^1 := \Delta_i^*(\bar{t}_i) \cup \left\{(\theta,t_{-i}^{s_{-i}})\,\middle|\,(\theta,s_{-i}) \in \mathrm{supp}(\mathrm{marg}_{S_{-i}\times\Theta^*}\bar{\mu}_i(h^0))\right\},$$

and then, type $t_i^1 := h_i^*(\Delta_i^1, \mathrm{marg}_{\Theta^*\times T^*_{-i}}\bar{\mu}_i^1(h^0))$. Clearly, t_i^1 is finite, and such that $\mu_i^1(t_i^1) = \mu_i^1(\bar{t}_i)$.[46] Furthermore, it is immediate that for every $\mu_i \in \mathrm{C}_i^{\mathscr{T}^*}(t_i^1)$ that strongly believes in $\mathcal{S}_{-i,1}$ the marginals of $\mu_i(h)$ and $\bar{\mu}_i^1(h)$—and hence that of $\bar{\mu}_i^1(h)$ too—coincide for every $h \in \mathrm{H}^0_{i,0}(\Delta_i^*(\bar{t}_i)) = H_i \cup \{h^0\}$, and hence, that $r_i(\mu_i) = [\bar{s}_i]$. Obviously, it follows that $\mathcal{S}_{i,2}(t_i^1) = [\bar{s}_i]$.

We continue now with the (arguably more tedious) proof of the inductive step. Suppose that $k \geq 2$ is such that the claims hold; we verify next that they also hold for $k+1$. To this end, fix player i, mildly consistent, finite type $\bar{t}_i$, strategy $\bar{s}_i \in \mathcal{S}^0_{i,k+1}(\bar{t}_i)$ such that $\bar{t}_i \in \mathcal{T}_{i,k+1}(\bar{s}_i)$, and conjecture

[46] Finiteness is immediate. To see that lower-order models are maintained, simply notice that the probability assigned by $\bar{\mu}_i(h^0)$ to θ is the same as the probability assigned by $\bar{\mu}_i^1(h^0)$.

$\bar{\mu}_i$ that justifies the inclusion of $\bar{t}_i$ in $\mathcal{T}_{i,k+1}(\bar{s}_i)$ (and thus the inclusion of $\bar{s}_i$ in $\mathcal{S}^0_{i,k+1}(\bar{t}_i)$). Let us now go claim by claim:

Claim (i). First, we know from the induction hypothesis that for any pair (s_{-i}, t_{-i}) in the support of $\bar{\mu}_i(h^0)$ there exists some $t^k_{-i}(s_{-i}, t_{-i}) = (t^k_j(s_j, t_j))_{j \neq i}$ such that, for every $j \neq i$, $t^k_j(s_j, t_j)$ is finite and such that $\mu^k_j(t^k_j(s_j, t_j)) = \mu^k_j(t_j)$ and $\mathcal{W}_{j,k+1}(t^k_j(s_j, t_j)) \subseteq [s_j|t_j]_{k-1}$. Define then $\bar{\mu}^{k+1}_i(h^0) \in \Delta(S_{-i} \times \Theta^* \times T^*_{-i})$ by setting:

$$\bar{\mu}^{k+1}_i(h^0)[E] := \bar{\mu}_i(h^0)\left[\left\{(s_{-i}, \theta, t_{-i}) \in S_{-i} \times \Theta^* \times T^*_{-i} \,\middle|\, (s_{-i}, \theta, t^k_{-i}(s_{-i}, t_{-i})) \in E\right\}\right],$$

for every measurable $E \subseteq S_{-i} \times \Theta^* \times T^*_{-i}$ (the finiteness of $\bar{t}_i$ guarantees that $\bar{\mu}^{k+1}_i$ is well-defined). For every $h \in \mathrm{H}^0_{i,k}(\Delta_i(\bar{t}_i)) \backslash \{h^0\}$ define $\bar{\mu}^{k+1}_i(h) \in \Delta(S_{-i} \times \Theta^* \times T^*_{-i})$ via conditional probability (this is possible because $\bar{\mu}_i(h^0)$ assigns positive probability to $S_{-i}(h)$ for every such h); notice then that, for every $h \in \mathrm{H}^0_{i,k}(\Delta_i(\bar{t}_i))$ and every measurable $E \subseteq S_{-i} \times \Theta^*$, we have that $\bar{\mu}^{k+1}_i(h)[E \times T^*_{-i}] = \bar{\mu}_i(h)[E \times T^*_{-i}]$, and thus, for every $h \in \mathrm{H}^0_{i,k}(\Delta_i(\bar{t}_i))$ and every $s_i \in S_i$ it holds that:

$$\int_{S_{-i} \times \Theta^*} \theta_i(z((s_{-i}, s_i)|h))\mathrm{d}(\mathrm{marg}_{S_{-i} \times \Theta^*} \bar{\mu}^{k+1}_i(h)) =$$
$$= \int_{S_{-i} \times \Theta^*} \theta_i(z((s_{-i}, s_i)|h))\mathrm{d}(\mathrm{marg}_{S_{-i} \times \Theta^*} \bar{\mu}_i(h)). \tag{U}$$

Then, define compact set:

$$\Delta^{k+1}_i := \Delta^*_i(\bar{t}_i) \cup \left\{(\theta, t^k_{-i}(s_{-i}, t_{-i})) \,\middle|\, (s_{-i}, \theta, t_{-i}) \in \mathrm{supp}\, \bar{\mu}_i(h^0)\right\},$$

and type:

$$t^{k+1}_i := h^*_i(\Delta^{k+1}_i, \mathrm{marg}_{\Theta^* \times T^*_{-i}} \bar{\mu}^{k+1}_i(h^0)).$$

Clearly, t^{k+1}_i is finite and such that $\mu^{k+1}_i(t^{k+1}_i) = \mu^{k+1}_i(\bar{t}_i)$.[47]

Claim (ii). Let us check now that:

$$\mathcal{W}_{i,k+2}(t^{k+1}_i) \subseteq [\bar{s}_i|\bar{t}_i]_k.$$

To see it fix strategy $s_i \in \mathcal{W}_{i,k+2}(t^{k+1}_i)$ and conjecture μ_i that justifies the inclusion of s_i in $\mathcal{W}_{i,k+2}(t^{k+1}_i)$. Throughout the remainder of the proof we will make use of some additional notation. First, for every $s_{-i} \in S_{-i}$ for which there is some t_{-i} consistent with $\bar{t}_i$ where $s_{-i} \in \mathcal{S}^0_{-i,k}(t_{-i})$,

$$\langle s_{-i} \rangle := \left\{ s''_{-i} \in S_{-i} \,\middle|\, \begin{array}{l} \text{There exist } t'_{-i}, t''_{-i} \in \mathrm{Proj}_{T^*_{-i}} \Delta^*(\bar{t}_i) \text{ and } s'_{-i} \in S_i \text{ such that:} \\ (1) \quad s'_{-i} \in [s_{-i}|t'_{-i}]_{k-1} \\ (2) \quad s''_{-i} \in [s''_{-i}|t''_{-i}]_{k-1} \end{array} \right\}.$$

Second, for every $h \in H_i \cup \{h^0\}$, every belief $\mu'_i \in \Delta(S_{-i} \times \Theta^* \times T^*_{-i})$ and every $s'_i \in S_i$, denote:

$$U_i(s'_i, \mu'_i|h) := \int_{S_{-i} \times \Theta^*} \theta_i(z((s_{-i}, s'_i)|h))\mathrm{d}(\mathrm{marg}_{S_{-i} \times \Theta^*} \mu'_i).$$

[47] Finiteness is immediate. To see that lower-order models are maintained, simply notice that the probability assigned by $\bar{\mu}_i(h^0)$ to (s_{-i}, θ, t_{-i}) is the same as the probability assigned by $\bar{\mu}^{k+1}_i(h^0)$ to $(s_{-i}, \theta, t^k_{-i}(s_{-i}, t_{-i}))$, and that the $(k-1)$th-order model of $t^k_{-i}(s_{-i}, t_{-i})$ exactly coincides with that of t^k_{-i} (by construction).

Then, we verify now the the following six facts hold:

(F1) *For any* $s_{-i} \in S_{-i}$ *for which there is some* t_{-i} *consistent with* $\bar{t}_i$ *where* $s_{-i} \in \mathcal{S}^0_{-i,k}(t_{-i})$*, any* $s''_{-i} \in \langle s_{-i} \rangle$*, and any* $s''_i \in \{\bar{s}_i, s_i\}$*,*

$$z((s_{-i}, s'_i)|h^0) = z((s'_{-i}, s'_i)|h^0). \tag{S1}$$

Fix $s_{-i} \in S_{-i}$, t_{-i} consistent with $\Delta^*_i(\bar{t}_i)$ where $s_{-i} \in \mathcal{S}^0_{-i,k}(t_{-i})$, and $s''_{-i} \in \langle s_{-i} \rangle$. Then, we know that there exist $s'_{-i} \in S_{-i}$ and t'_{-i}, t''_{-i} consistent with $\bar{t}_i$ such that $s'_{-i} \in [s_{-i}|t'_{-i}]_{k-1}$ and $s''_{-i} \in [s'_{-i}|t''_{-i}]_{k-1}$—case in which it holds that, for every $j \neq i$ and every $h \in \mathrm{H}^{\mathcal{W}}_{j,k-1}(\Delta^*_j(t'_j)) \cap \mathrm{H}^{\mathcal{W}}_{j,k-1}(\Delta^*_j(t''_j))$, $s''_j(h) = s_j(h)$. Fix also $s''_i \in \{\bar{s}_i, s_i\}$ and pick player $j \neq i$ such that there exists some history $h \in H_j$ that precedes $z((s_{-i}, s''_i)|h^0)$ and is reached by s''_j.[48] We now claim that the inclusion $h \in \mathrm{H}^{\mathcal{W}}_{j,k-1}(\Delta^*_j(t'_j)) \cap \mathrm{H}^{\mathcal{W}}_{j,k-1}(\Delta^*_j(t''_j))$ holds; to see it notice that the following four hold for $t'''_j \in \{t'_j, t''_j\}$:

(*a*) $((t_\ell)_{\ell \neq i,j}; \bar{t}_i)$ is consistent with t'''_j.[49]

(*b*) $s_\ell \in \mathcal{S}_{\ell,k}(t_\ell) \subseteq \mathcal{W}^0_{\ell,k-1}(t_\ell)$ for every $\ell \neq i, j$.

(*c*) $\bar{s}_i \in \mathcal{S}^0_{i,k+1}(\bar{t}_i) \subseteq \mathcal{W}_{i,k-1}(\bar{t}_i)$.

(*d*) $s_i \in \mathcal{W}_{i,k+2}(t^{k+1}_i) \subseteq \mathcal{W}_{i,k-1}(\bar{t}_i)$.

We thus conclude that $s''_j(h) = s_j(h)$ and hence the claim is proved.

(F2) *For any* $s_{-i} \in S_{-i}$ *for which there is some* t_{-i} *consistent with* $\bar{t}_i$ *where* $s_{-i} \in \mathcal{S}^0_{-i,k}(t_{-i})$*, and for every* $\theta \in \Theta^*$ *we have:*

$$\mu_i(h^0)\left[\langle s_{-i} \rangle \times \{\theta\} \times T^*_{-i}\right] = \bar{\mu}_i(h^0)\left[\langle s_{-i} \rangle \times \{\theta\} \times T^*_{-i}\right]. \tag{S2}$$

Fix (s_{-i}, θ) and develop:

$$\mu_i(h^0)\left[[\langle s_{-i} \rangle \times \{\theta\} \times T^*_{-i}\right] =$$

$$= \mu_i(h^0)\left[\langle s_{-i} \rangle \times \{\theta\} \times \left\{ t^k_{-i}(s'_{-i}, t'_{-i}) \;\middle|\; \begin{array}{ll} (1) & t'_{-i} \in \mathrm{Proj}_{T^*_{-i}} \Delta^*_i(\bar{t}_i) \\ (2) & (s'_{-i}, t'_{-i}) \in \mathrm{Graph}(\mathcal{S}^0_{-i,k}) \end{array} \right\}\right]$$

$$\overset{(i)}{=} \mu_i(h^0)\left[S_{-i} \times \{\theta\} \times \left\{ t^k_{-i}(s'_{-i}, t'_{-i}) \;\middle|\; \begin{array}{ll} (1) & t'_{-i} \in \mathrm{Proj}_{T^*_{-i}} \Delta^*_i(\bar{t}_i) \\ (2) & (s'_{-i}, t'_{-i}) \in \mathrm{Graph}(\mathcal{S}^0_{-i,k}) \\ (3) & \langle s_{-i} \rangle \cap [s'_{-i}|t'_{-i}]_{k-1} \neq \emptyset \end{array} \right\}\right]$$

$$= \bar{\mu}^{k+1}_i(h^0)\left[S_{-i} \times \{\theta\} \times \left\{ t^k_{-i}(s'_{-i}, t'_{-i}) \;\middle|\; \begin{array}{ll} (1) & t'_{-i} \in \mathrm{Proj}_{T^*_{-i}} \Delta^*_i(\bar{t}_i) \\ (2) & (s'_{-i}, t'_{-i}) \in \mathrm{Graph}(\mathcal{S}^0_{-i,k}) \\ (3) & \langle s_{-i} \rangle \cap [s'_{-i}|t'_{-i}]_{k-1} \neq \emptyset \end{array} \right\}\right]$$

$$\overset{(ii)}{=} \bar{\mu}^{k+1}_i(h^0)\left[\left\{ (s'_{-i}, \theta, t^k_{-i}(s'_{-i}, t'_{-i})) \;\middle|\; \begin{array}{ll} (1) & t'_{-i} \in \mathrm{Proj}_{T^*_{-i}} \Delta^*_i(\bar{t}_i) \\ (2) & (s'_{-i}, t'_{-i}) \in \mathrm{Graph}(\mathcal{S}^0_{-i,k}) \\ (3) & s'_{-i} \in \langle s_{-i} \rangle \end{array} \right\}\right]$$

[48] I.e., such that $(s_{-i}, s''_i) \in S(h)$ and $s''_j \in S_j(h)$. We do not have to worry about players j that lack such history h: they do not make any choice along the path leading to $z((s_{-i}, s''_i)|h^0)$.

[49] Remember that $\bar{t}_i$ is mildly consistent and thus, if t'''_j and t_ℓ are consistent with $\bar{t}_i$ then $\bar{t}_i$ and t_ℓ are consistent with t'''_j.

$$\stackrel{(iii)}{=} \bar{\mu}_i(h^0)\left[\left\{(s'_{-i},\theta,t'_{-i}) \;\middle|\; \begin{array}{ll} (1) & t'_{-i} \in \mathrm{Proj}_{T^*_{-i}} \Delta^*_i(\bar{t}_i) \\ (2) & (s'_{-i},t'_{-i}) \in \mathrm{Graph}(\mathcal{S}^0_{-i,k}) \\ (3) & s'_{-i} \in \langle s_{-i}\rangle \end{array}\right\}\right]$$

$$= \bar{\mu}_i(h^0)\left[\langle s_{-i}\rangle \times \{\theta\} \times T^*_{-i}\right].$$

(F3) *For any $h \in \mathrm{H}^{\mathcal{W}}_{i,k}(\Delta^*_i(\bar{t}_i)) \cap H_i(\bar{s}_i)$ and any $s_{-i} \in S_{-i}$ for which there is some t_{-i} consistent with $\bar{t}_i$ where $s_{-i} \in \mathcal{S}^0_{-i,k}(t_{-i})$, and any $\theta \in \Theta^*$ we have:*

$$\mu_i(h)\left[\langle s_{-i}\rangle \times \{\theta\} \times T^*_{-i}\right] = \bar{\mu}_i(h)\left[\langle s_{-i}\rangle \times \{\theta\} \times T^*_{-i}\right]. \tag{S3}$$

Fix history $h \in \mathrm{H}^{\mathcal{W}}_{i,k}(\Delta^*_i(t_i)) \cap H_i(\bar{s}_i)$ and pick $s_{-i} \in S_{-i}(h)$ and t'_{-i} consistent with $\bar{t}_i$ such that $s_{-i} \in \mathcal{S}^0_{-i,k}(t'_{-i})$. Then, for every $s'_{-i} \in [s_{-i}|t'_{-i}]_{k-1}$ we have that $[s'_{-i}|t'''_{-i}]_{k-1} \subseteq S_{-i}(h)$ for every t''_{-i} consistent with $\bar{t}_{-i}$ where $s'_{-i} \in \mathcal{S}^0_{-i,k}(t''_{-i})$. To see it, pick $s''_{-i} \in [s'_{-i}|t''_{-i}]_{k-1}$, and player $j \neq i$ with some history $h' \in H_j$ that precedes h. Notice then that $h' \in \mathrm{H}^{\mathcal{W}}_{j,k-1}(\Delta^*_i(t_j)) \cap \mathrm{H}^{\mathcal{W}}_{j,k-1}(\Delta^*_i(t''_j))$ is a consequence of these four facts:

(*a*) $((t'_\ell)_{\ell \neq i,j}, \bar{t}_i)$ is consistent with both t'_j and t''_j.

(*b*) $s_\ell \in \mathcal{S}^0_{\ell,k}(t'_\ell) \cap S_\ell(h) \subseteq \mathcal{W}_{\ell,k-1}(t'_\ell) \cap S_\ell(h)$ for every $\ell \neq i, j$.

(*c*) $\bar{s}_i \in \mathcal{S}^0_{i,k+1}(\bar{t}_i) \cap S_i(h) \subseteq \mathcal{W}_{i,k-1}(\bar{t}_i) \cap S_i(h)$.

(*d*) h' precedes h.

It follows then that $s''_j(h') = s'_j(h') = s_j(h')$ and thus, that $s''_{-i} \in S_{-i}(h)$. Hence, $\langle s_{-i}\rangle \subseteq S_{-i}(h)$ and, in consequence, there must exist some $s^1_{-i}, \ldots, s^M_{-i} \in S_{-i}(h)$ such that the family $\{\langle s^1_{-i}\rangle, \ldots, \langle s^M_{-i}\rangle\}$ is a partition of $S_{-i}(h)$. As a result:

$$\begin{aligned}\mu_i(h^0)[S_{-i}(h) \times \Theta^* \times T^*_{-i}] &= \sum_{m=1}^{M} \mu_i(h^0)[\langle s^m_{-i}\rangle \times \Theta^* \times T^*_{-i}] \\ &= \sum_{m=1}^{M} \bar{\mu}^{k+1}_i(h^0)[\langle s^m_{-i}\rangle \times \Theta^* \times T^*_{-i}] \\ &= \bar{\mu}^{k+1}_i(h^0)[S_{-i}(h) \times \Theta^* \times T^*_{-i}].\end{aligned}$$

Remember now that $\bar{\mu}^{k+1}_i(h^0)$ puts positive probability on $S_{-i}(h) \times \Theta^* \times T^*_{-i}$ and thus, we have that, for every (s_{-i}, θ, t_{-i}) in the support of $\bar{\mu}_i(h^0)$,

$$\begin{aligned}\bar{\mu}^{k+1}_i(h)[\langle s_{-i}\rangle \times \{\theta\} \times T^*_{-i}] &= \frac{\bar{\mu}^{k+1}_i(h^0)[(S_{-i}(h) \cap \langle s_{-i}\rangle \times \{\theta\} \times T^*_{-i}]}{\bar{\mu}^{k+1}_i(h^0)[S_{-i}(h) \times \Theta^* \times T^*_{-i}]} \\ &= \frac{\bar{\mu}_i(h^0)[(S_{-i}(h) \cap \langle s_{-i}\rangle \times \{\theta\} \times T^*_{-i}]}{\bar{\mu}_i(h^0)[S_{-i}(h) \times \Theta^* \times T^*_{-i}]} \\ &= \bar{\mu}_i(h)[\langle s_{-i}\rangle \times \{\theta\} \times T^*_{-i}].\end{aligned}$$

(F4) *It holds that:*

$$\mathrm{H}^{\mathcal{W}}_{i,k}(\Delta^*_i(t^{k+1}_i)) \cap H_i(\bar{s}_i) = \mathrm{H}^{0}_{i,k}(\Delta^*_i(\bar{t}_i)) \cap H_i(\bar{s}_i). \tag{S4}$$

For the westwards inclusion, notice first that we know that $\mathrm{marg}_{S_{-i}}\mu_i(h^0)[S_{-i}(h)] = 0$ for every $h \notin \mathrm{H}^{\mathcal{W}}_{i,k+1}(\Delta^*_i(t^{k+1}_i))$. Now, we also know that $\mathrm{marg}_{S_{-i}}\mu_i(h^0)[S_{-i}(h)] > 0$ for every

$h \in \mathrm{H}^{0}_{i,k}(\Delta^*_i(\bar{t}_i))$. Hence, $\mathrm{H}^{0}_{i,k}(\Delta^*_i(\bar{t}_i)) \subseteq \mathrm{H}^{\mathcal{W}}_{i,k}(\Delta^*_i(t^{k+1}_i))$. For the eastwards inclusion, we have that:

$$\begin{aligned}
&\mathrm{H}^{\mathcal{W}}_{i,k+1}(\Delta^*_i(t^{k+1}_i)) \cap H_i(\bar{s}_i) = \\
&\quad = H_i(\bar{s}_i) \cap \left\{ h \in H_i \cup \{h^0\} \,\middle|\, S_{-i}(h) \cap \bigcup_{t_{-i} \in \mathrm{Proj}_{T^*_{-i}} \Delta^*_i(t^{k+1}_i)} \mathcal{W}_{-i,k+1}(t_{-i}) \neq \emptyset \right\} \\
&\quad \subseteq H_i(\bar{s}_i) \cap \left\{ h \in H_i \cup \{h^0\} \,\middle|\, S_{-i}(h) \cap \bigcup_{t_{-i} \in \mathrm{Proj}_{T^*_{-i}} \Delta^*_i(\bar{t}_i)} \bigcup_{s_{-i} \in \mathcal{S}^0_{-i,k}(t_{-i})} [s_{-i}|t_{-i}]_{k-1} \neq \emptyset \right\}
\end{aligned}$$

Now, fix history h in the last set above. We know that there exist some $t_{-i} \in \mathrm{Proj}_{T^*_{-i}} \Delta^*_i(\bar{t}_i)$, some $s_{-i} \in \mathcal{S}^0_{-i,k}(t_{-i})$ and some $s'_{-i} \in [s_{-i}|t_{-i}]_{k-1}$ such that $s'_{-i} \in S_{-i}(h)$. We also know that $\bar{s}_i \in S_i(h)$. Now, we can write:

$$h = (h^0, a^1, a^2, \ldots, a^n),$$

where, for each $k = 1, \ldots, n$, a^k is a description of the actions chosen by the players active at $h^{k-1} = (h^0, a^1, \ldots, a^{k-1})$. Then, we proceed inductively. For every $j \neq i$ such that $h^0 \in H_j$ we have that $h^0 \in \mathrm{H}^{\mathcal{W}}_{j,k-1}(\Delta^*_j(t_j)) \cap H_j(s_j)$ (and, in case that $h \in H_i$, also that $h^0 \in \mathrm{H}^{\mathcal{W}}_{i,k-1}(\Delta^*_i(t_i)) \cap H_i(\bar{s}_i)$) and thus, that $s_j(h^0) = s'_j(h^0)$. Obviously, it follows that, for every $j \neq i$ such that $h^1 \in H_j$, we have that $h^1 \in \mathrm{H}^{\mathcal{W}}_{j,k-1}(\Delta^*_j(t_j)) \cap H_j(s_j)$ (and, in case that $h \in H_i$, also that $h^0 \in \mathrm{H}^{\mathcal{W}}_{i,k-1}(\Delta^*_i(t_i)) \cap H_i(\bar{s}_i)$) and thus, again, that $s_j(h^1) = s'_j(h^1)$. Continuing inductively establishes thus that s_{-i} and s'_{-i} choose identically at every h' preceding h and thus, that $s_{-i} \in S_{-i}(h)$. This implies then that:

$$\begin{aligned}
&H_i(\bar{s}_i) \cap \left\{ h \in H_i \cup \{h^0\} \,\middle|\, S_{-i}(h) \cap \bigcup_{t_{-i} \in \mathrm{Proj}_{T^*_{-i}} \Delta^*_i(\bar{t}_i)} \bigcup_{s_{-i} \in \mathcal{S}^0_{-i,k}(t_{-i})} [s_{-i}|t_{-i}]_{k-1} \neq \emptyset \right\} \subseteq \\
&\quad \subseteq H_i(\bar{s}_i) \cap \left\{ h \in H_i \cup \{h^0\} \,\middle|\, S_{-i}(h) \cap \bigcup_{t_{-i} \in \mathrm{Proj}_{T^*_{-i}} \Delta^*_i(\bar{t}_i)} \mathcal{S}^0_{-i,k}(t_{-i}) \neq \emptyset \right\} \\
&\quad = H_i(\bar{s}_i) \cap \mathrm{H}^{0}_{i,k}(\Delta^*_i(\bar{t}_i)).
\end{aligned}$$

(F5) *For any $h \in \mathrm{H}^{\mathcal{W}}_{i,k}(\Delta^*_i(t^{k+1}_i)) \cap H_i(\bar{s}_i)$ and any $s'_i \in \{\bar{s}_i, s_i\}$ we have that:*

$$U_i(\mu_i, s'_i|h) = U_i(\bar{\mu}^{k+1}_i, s'_i|h) \tag{S5}$$

Fix history $h \in \mathrm{H}^{\mathcal{W}}_{i,k}(\Delta^*_i(t^{k+1}_i)) \cap H_i(\bar{s}_i) = \mathrm{H}^{0}_{i,k}(\Delta^*_i(\bar{t}_i)) \cap H_i(\bar{s}_i)$ (notice the use of (S4)) and pick again $s^1_{-i}, \ldots, s^M_{-i} \in S_{-i}(h)$ such that the family $\{\langle s^1_{-i}\rangle, \ldots, \langle s^M_{-i}\rangle\}$ is a partition of $S_{-i}(h)$. Since $\bar{\mu}^{k+1}_i(h^0)$ has finite support on Θ^* (and hence also does $\mu_i(h^0)$) we have that, for any $s'_i \in \{\bar{s}_i, s_i\}$,

$$U_i(\mu_i, s'_i|h) = \sum_{m=1}^{M} \sum_{(s_{-i},\theta) \in \langle s^m_{-i}\rangle \times \Theta^*} \mu_i(h)[\{(s_{-i}, \theta)\} \times T^*_{-i}] \theta_i(z((s_{-i}, s'_i)|h)) =$$

$$
\begin{aligned}
&= \sum_{m=1}^{M} \sum_{\theta\in\Theta^*} \mu_i(h)[\langle s_{-i}^m\rangle \times \{\theta\} \times T_{-i}^*]\theta_i(z((s_{-i}^m, s_i')|h)) \\
&= \sum_{m=1}^{M} \sum_{\theta\in\Theta^*} \bar{\mu}_i^{k+1}(h)[\langle s_{-i}^m\rangle \times \{\theta\} \times T_{-i}^*]\theta_i(z((s_{-i}^m, s_i')|h)) \\
&= \sum_{m=1}^{M} \sum_{(s_{-i},\theta)\in\langle s_{-i}^m\rangle\times\Theta^*} \bar{\mu}_i^{k+1}(h)[\{(s_{-i},\theta)\} \times T_{-i}^*]\theta_i(z((s_{-i}, s_i')|h)) \\
&= U_i(\bar{\mu}_i^{k+1}, s_i'|h),
\end{aligned}
$$

the second and fourth equalities following from (S1) and the third, from (S3).

(F6) *It holds that:*

$$s_i \in [\bar{s}_i|\bar{t}_i]_{k-1} \tag{S6}$$

Proceed by contradiction and suppose that there exists some history $\bar{h} \in \mathrm{H}_{i,k}^{\mathcal{W}}(\Delta_i^*(t_i^{k+1})) \cap H_i(\bar{s}_i) = \mathrm{H}_{i,k}^0(\Delta_i^*(\bar{t}_i)) \cap H_i(\bar{s}_i)$ such that $s_i(\bar{h}) \neq \bar{s}_i(\bar{h})$ (notice the use of (S4)). Pick then strategy $\hat{s}_i$ that maximizes $U_i(\,\cdot\,, \bar{\mu}_i|h)$ at every h that weakly follows $\bar{h}$. Then, define a new strategy s_i^0 by setting:

$$
s_i^0(h') := \begin{cases} \hat{s}_i(h) & \text{if } S_{-i}(h) \subsetneq S_{-i}(\bar{h}), \\ s_i(h) & h = \bar{h}, \\ \bar{s}_i(h) & \text{otherwise}, \end{cases}
$$

for every $h \in H_i$. Let's check next that $s_i^0 \in r_i(\bar{\mu}_i)$. To this end, notice that every $h \in H_i \cup \{h^0\}$ falls in some of the following categories:

- $S_{-i}(h) \subsetneq S_{-i}(\bar{h})$. By construction of s_i^0, we have that $U_i(s_i^0, \bar{\mu}_i|h) = U_i(\hat{s}_i, \bar{\mu}_i|h)$. Thus, s_i^0 maximizes $U_i(\,\cdot\,, \bar{\mu}_i|h)$.
- $h = \bar{h}$. The above, together with $s_i^0(h) = s_i(h)$, implies that:[50]

  $$U_i(s_i^0, \bar{\mu}_i|h) \geq U_i(s_i, \bar{\mu}_i|h) = U_i(s_i, \bar{\mu}_i^{k+1}|h).$$

  Now, we also have that:[51]

  $$U_i(s_i, \bar{\mu}_i^{k+1}|h) = U_i(s_i, \mu_i|h) \geq U_i(\bar{s}_i, \mu_i|h) = U_i(\bar{s}_i, \bar{\mu}_i^{k+1}|h) = U_i(\bar{s}_i, \bar{\mu}_i|h),$$

  and remember that $\bar{s}_i$ maximizes $U_i(\,\cdot\,, \bar{\mu}_i|h)$. Consequently, s_i^0 maximizes $U_i(\,\cdot\,, \bar{\mu}_i|h)$.
- $S_{-i}(\bar{h}) \subsetneq S_{-i}(h)$. By construction of s_i^0, and because of the facts that $U_i(s_i^0, \bar{\mu}_i|\bar{h}) = U_i(\bar{s}_i, \bar{\mu}_i|\bar{h})$ and s_i^0 and $\bar{s}_i$ coincide at every history h preceding $\bar{h}$, we have $U_i(s_i^0, \bar{\mu}_i|h) = U_i(\bar{s}_i, \bar{\mu}_i|h)$. In consequence, since $\bar{s}_i$ maximizes $U_i(\,\cdot\,, \bar{\mu}_i|h)$ it follows that s_i^0 must maximize $U_i(\,\cdot\,, \bar{\mu}_i|h)$ as well.
- For any other h, clearly, we have that $U_i(s_i^0, \bar{\mu}_i|h) = U_i(\bar{s}_i, \bar{\mu}_i|h)$.

We thus reach a contradiction: $s_i^0 \notin [\bar{s}_i|\bar{t}_i]_k$ but $s_i^0 \in r_i(\bar{\mu}_i) = [\bar{s}_i] \subseteq [\bar{s}_i|\bar{t}_i]_k$.

Consequently, $\mathcal{W}_{i,k+2}(t_i^{k+1}) \subseteq [\bar{s}_i|\bar{t}_i]_k$. ■

[50] For the first inequality notice that s_i^0 and s_i coincide at $h = \bar{h}$ and that s_i^0 maximizes $U_i(\,\cdot\,, \bar{\mu}_i|h')$ for every h' strictly following h. The equality follows from (U) above.

[51] The first two equalities follow from (S4) and the last, from equation (U) above. The inequality is a consequence of s_i being a best-reply to μ_i.

E.3 Proof of the result

THEOREM 3. *Let $(\mathscr{E}, \mathscr{T}^*)$ be a dynamic Bayesian game and let t be a consistent profile of finite, information-based types. Then, the predictions of strong rationalizability for t admit unique selections via weak rationalizability.*

Proof. Fix a profile of information-based consistent types $t \in T^*$ and a profile of strategies $s \in \mathcal{S}(t)$. We know that for every $k \geq 2$ and every player i:

(a) From Lemma 10, that there exists some sequence of finite, mildly consistent types $(t_i^{k,n})_{n\in\mathbb{N}}$ converging to t_i such that, for every $n \in \mathbb{N}$, $s_i \in \mathcal{S}^0_{i,k}(t_i^{k,n})$ and:

$$\mathrm{H}_{i,k-1}(\Delta_i^*(t_i^{k,n})) = \mathrm{H}_{i,k-1}(\Delta_i^*(t_i)).$$

(b) Then, from Lemma 11, that for every $n \in \mathbb{N}$ there exists a sequence of finite, mildly consistent types $(t_i^{k,n,\ell})_{\ell\in\mathbb{N}}$ converging to $t_i^{k,n}$ such that, for any $\ell \in \mathbb{N}$, $t_i^{k,n,\ell} \in \mathcal{T}_{i,k}(s_i)$ and:

$$\mathrm{H}_{i,k-1}(\Delta_i^*(t_i^{k,n,\ell})) = \mathrm{H}_{i,k-1}(\Delta_i^*(t_i)).$$

For each player i and $m \in \mathbb{N}$ set $\bar{t}_i^m := t_i^{n_m,n_m,m}$. Then, we know then that for any player i:

(c) From Lemma 12, that there exists a finite type $t_i^{n_m}$ such that $\mu_i^{n_m}(t_i^{n_m}) = \mu_i^{n_m}(\bar{t}_i^m)$ and:

$$\mathcal{W}_i(t_i^{n_m}) \subseteq \mathcal{W}_{i,n_m+1}(t_i^{n_m}) \subseteq [s_i|t_i^{n_m}]_{n_m-1},$$

and:

$$\mathrm{H}^{\mathcal{W}}_{i,n_m-1}(\Delta_i^*(t_i^{n_m})) \cap H_i(s_i) = \mathrm{H}_{i,n_m-1}(\Delta_i^*(t_i)) \cap H_i(s_i).$$

Clearly, it also holds that:

$$\begin{aligned}
[s_i|t_i^{n_m}]_{n_m-1} &= \left\{ s_i' \in S_i \,\middle|\, s_i'(h) = s_i(h) \text{ for every } h \in \mathrm{H}^{\mathcal{W}}_{i,n_{m-1}}(\Delta_i^*(t_i^{n_m})) \cap H_i(s_i) \right\} \\
&= \left\{ s_i' \in S_i \,\middle|\, s_i'(h) = s_i(h) \text{ for every } h \in \mathrm{H}_{i,n_{m-1}}(\Delta_i^*(t_i)) \cap H_i(s_i) \right\} \\
&\subseteq \left\{ s_i' \in S_i \,|\, s_i'(h) = s_i(h) \text{ for every } h \in \mathrm{H}_i(\Delta_i^*(t_i)) \cap H_i(s_i) \right\}.
\end{aligned}$$

Thus, if for each $m \in \mathbb{N}$ and player i we set $\tilde{t}^m := t_i^{n_m}$ we have found a sequence of profiles of types $(\tilde{t}^m)_{m\in\mathbb{N}}$ converging to t and such that, for any player i and any $m \in \mathbb{N}$,

$$\mathcal{W}_i(\tilde{t}^m) \subseteq \left\{ s_i' \in S_i \,|\, s_i'(h) = s_i(h) \text{ for every } h \in \mathrm{H}_i(\Delta_i^*(t_i)) \cap H_i(s_i) \right\}.$$

Now, fix $m \in \mathbb{N}$ and $s^m \in \mathcal{W}(\tilde{t}^m)$. Let $J \subseteq I$ be the set of those players i for which $h^0 \in H_i$. Clearly, $h^0 \in \mathrm{H}_i^{\mathcal{W}}(\Delta_i^*(t_i)) \cap H_i(s_i)$ and thus, it must hold that $s_i^m(h^0) = s_i(h^0)$. Then, obviously, for every player j such that $(h^0, (s_i^m(h^0))_{i\in J}) \in H_j$, it holds that $(h^0, (s_i^m(h^0))_{i\in J}) \in \mathrm{H}_j^{\mathcal{W}}(\Delta_j^*(t_j)) \cap H_i(s_i)$ as well (remember that t is consistent). Hence, an easy inductive argument enables to conclude that $z(s^m|h^0) = z(s|h^0)$. ∎