The overlap between Extensive-Form Rationalizability and Subgame Perfection.*

Emiliano Catonini†

January 2017 - PRELIMINARY AND INCOMPLETE

Battigalli (1997) first proved that in dynamic games with perfect information and no relevant ties, the unique backward induction outcome is extensive-form rationalizable (Pearce, 1984). Other authors use the outcome inclusion between extensive-form rationalizability and backward induction to obtain the same result. Allowing for simultaneous moves, subgame perfect equilibrium outcomes are only a subset of the backward induction ones. Thus, showing that at least one subgame perfect equilibrium outcome distribution is extensive-form rationalizable requires a different strategy. Here I prove that in all finite dynamic games with observable actions there always exists an outcome distribution which is induced by both a subgame perfect equilibrium in behavioral strategies and a Nash equilibrium in strongly rationalizable strategies. Moreover, I show that only for a subgame perfect equilibrium path, strategic reasoning under belief in this path may yield as unique prediction the path itself.

Keywords: Extensive-form rationalizability, Strong Rationalizability, Subgame Perfect Equilibrium, Path agreements.

1 Introduction

Battigalli [1] was the first to prove that in dynamic games with perfect information and no relevant ties, the unique backward induction outcome is extensive-form rationalizable (Pearce, [16], Battigalli, [1]). Chen and Micali [9] come to the same conclusion as a corollary of the following result: extensive-form rationalizability refines backward induction. Heifetz and

*A special thanks goes to Pierpaolo Battigalli, this paper would not exist without his mentoring. Thank you also to Adam Brandenburger, Shurojit Chatterji, Yi-Chun Chen, Alfredo Di Tillio, Amanda Friedenberg, Amanda Jakobsson, Atsushi Kajii, Mattia Landoni, Xiao Luo, Elena Manzoni, Andres Perea, Burkhard Schipper, Madhav Shrihari Aney, Satoru Takahashi, Elias Tsakas, and Dimitrios Tsomocos for precious suggestions.

†Higher School of Economics, ICEF, emiliano.catonini@gmail.com

Perea [11] also prove Battigalli's theorem by showing the outcome equivalence of extensive-form-rationalizability and backward induction in the same class of games.

Unfortunately, dynamic games with perfect information are a very small class of games. In applications, players are very frequently allowed to move simultaneously — or equivalently, before observing each other's move. Repeated games with perfect monitoring are a prominent example of *games with observable actions*. In this class of games, the set of backward induction outcomes can expand dramatically. Therefore, games are usually solved with some notion of equilibrium. In absence of payoff uncertainty, subgame perfect equilibrium is by far the most widespread and accepted equilibrium concept.[1] However, the possible multiplicity calls for further refinements. A vast literature, stemming from the seminal contribution of Kohlberg and Mertens [12], refines subgame perfect equilibrium with forward induction arguments.[2] However, the forward induction arguments employed in this literature have two main shortcomings. First, they are quite opaque: it is very hard to understand what kind of strategic reasoning they actually capture beyond the simple examples analyzed in the papers. Second, they do not seem to capture all possible steps of forward induction reasoning that players may apply.[3]

Another stream of literature, stemming from the seminal contribution of Pearce [16], tackles forward induction reasoning from a completely different perspective. Abstracting away from any equilibrium concept, extensive-form rationalizability aims to capture in a transparent way all possible steps of reasoning that players may perform by forward induction. In absence of first-order-belief restrictions (which may throw equilibrium reasoning in the mix), this literature probably culminates in the notion of Strong Rationalizability [4], which is given an epistemic characterization.

The orthogonality of the two approaches, makes it very difficult to establish any connection between Strong Rationalizability and subgame perfect equilibrium However, the existence of a common prediction between the two solution concepts would be extremely valuable in terms of robustness and would, so to say, pacify the advocates of the two approaches. Such prediction would be even more valuable if obtainable under a unified approach to strategic reasoning and equilibrium reasoning, as the one offered by Selective Rationalizability [7].[4] The main result of this paper affirms that in all finite dynamic games with observable actions, there always exists a Nash equilibrium in strongly rationalizable strategies that induces a subgame perfect equilibrium outcome distribution. Extending the results of Catonini [8] to

[1] Sequential equilibrium (Kreps and Wilson [13]) coincides with subgame perfect equilibrium in this class of games.

[2] See, for instance, Govindan and Wilson [10] and Man [14].

[3] See Catonini (2017) for a critical survey.

[4] Under Strong-Δ-Rationalizability (Battigalli [3], Battigalli and Siniscalchi, [5]), the first solution concept to introduce first-order belief restrictions into the structure of Strong Rationalizability, it is easy to prove that, in games with observable actions, all subgame perfect equilibria can be obtained, hence also those that Strong Rationalizability excludes.

mixed strategies in the obvious way, it turns out that the support of the distribution can be induced by Selective Rationalizability under first-order belief restrictions on such Nash equilibrium.[5] Using the same arguments, I also show that only a subgame perfect equilibrium path can be delivered by Strong-Δ-Rationalizability/Selective Rationalizability[6] as unique prediction in absence of off-the-path first-order belief restrictions.

Section 2 provides the preliminary definitions. Section 3 states the main results and provides an intuition of their proof. Section 4 proves the key lemma.

2 Preliminaries

Primitives of the game.[7] Let I be the finite set of *players*. For any profile $(X_i)_{i\in I}$ and any $\emptyset \neq J \subseteq I$, I write $X_J := \times_{j\in J} X_j$, $X := X_I$, $X_{-i} := X_{I\setminus\{i\}}$. Let $(\overline{A}_i)_{i\in I}$ be the finite sets of *actions* potentially available to each player. Let $\overline{H} \subseteq \cup_{t=1,...,T}\overline{A}^t \cup \{\emptyset\}$ be the set of histories, where $h^0 := \{\emptyset\} \in \overline{H}$ is the root of the game and T is the finite horizon. For any $h = (a^1, ..., a^t) \in \overline{H}$ and $l < t$, it holds $h' = (a^1, ..., a^l) \in \overline{H}$, and I write $h' \prec h$.[8] Let $Z := \{z \in \overline{H} : \forall h \in \overline{H}, z \nprec h\}$ be the set of terminal histories (henceforth, *outcomes* or *paths*)[9], and $H := \overline{H}\backslash Z$ the set of non-terminal histories (henceforth, just *histories*). For each $i \in I$, let $A_i : H \rightrightarrows \overline{A}_i$ be the correspondence that assigns to each history h, always observed by player i, the set of actions $A_i(h) \neq \emptyset$[10] available at h. Thus, H has the following property: For every $h \in H$, $(h, a) \in \overline{H}$ if and only if $a \in A(h)$. Note that to simplify notation every player is required to play an action at every history: when a player is not truly active at a history, her set of feasible actions consists of just one "wait" action. For each $i \in I$, let $u_i : Z \to \mathbb{R}$ be the *payoff function*. The list $\Gamma = \langle I, \overline{H}, (u_i)_{i\in I}\rangle$ is a *finite game with complete information and observable actions*.

Derived objects. A strategy of player i is a function $s_i : h \in H \longmapsto s_i(h) \in A_i(h)$. Let S_i denote the set of all strategies of i. The set of mixed strategies of player i is denoted by $\Delta(S_i)$. Abusing notation, for a profile of mixed strategies $\sigma = (\sigma_j)_{j\in I} \in \times_{j\in I}\Delta(S_j)$, σ will be used also to denote the induced probability distribution over S, and $u_i(\sigma)$ will denote the expected payoff of player i under σ. A strategy *profile* $s \in S$ naturally induces a unique

[5] With a caveat: if the equilibrium does not provide strict incentive to remain on path, Selective Rationalizability will yield a superset of the equilibrium outcomes.

[6] The two are equivalent under path restrictions: see Catonini (2017).

[7] The main notation is almost entirely taken from Osborne and Rubinstein [15].

[8] $\overline{H}$ endowed with the precedence relation $\prec$ is a tree with root h^0.

[9] "Path" will be used with emphasis on the moves, and "outcome" with emphasis on the end-point of the game.

[10] When player i is not truly active at history h, $A_i(h)$ consists of just one "wait" action.

outcome $z \in Z$. Let $\zeta : S \to Z$ be the function that associates each strategy profile with the induced outcome. For any $h \in \overline{H}$, the set of strategies of i compatible with h is:

$$S_i(h) := \{s_i \in S_i : \exists z \succeq h, \exists s_{-i} \in S_{-i}, \zeta(s_i, s_{-i}) = z\}.$$

For any $(\overline{S}_j)_{j\in I} \subset S$, let $\overline{S}_i(h) := S_i(h) \cap \overline{S}_i$. Moreover, let:

$$\overline{S}_i[h] = \left\{a_i \in A_i(h) : \exists s_i \in \overline{S}_i(h), s_i(h) = a_i\right\}.$$

For any $J \subseteq I$, let $H(\overline{S}_J) := \left\{h \in H : \overline{S}_J(h) \neq \emptyset\right\}$ denote the set of histories compatible with $\overline{S}_J$. For any $h = (h', a) \in \overline{H}$, let $p(h)$ denote the immediate predecessor h' of h.

Since the game has observable actions, each history $h \in H$ is the root of a subgame $\Gamma(h)$. In $\Gamma(h)$, all the objects defined above will be denoted with h as superscript, except for single histories and outcomes, which will be identified with the corresponding history or outcome of the whole game, and not redefined as shorter lists of action profiles. For any $h \in H$, $s_i^h \in S_i^h$, and $\widehat{h} \succ h$, $s_i^h|\widehat{h}$ will denote the strategy $s_i^{\widehat{h}} \in S_i^{\widehat{h}}$ such that $s_i^{\widehat{h}}(\widetilde{h}) = s_i^h(\widetilde{h})$ for all $\widetilde{h} \succeq \widehat{h}$. For any $\overline{S}_i^h \subseteq S_i^h$, $\overline{S}_i^h|\widehat{h}$ will denote the set of all strategies $s_i^{\widehat{h}} \in S_i^{\widehat{h}}$ such that $s_i^{\widehat{h}} = s_i^h|\widehat{h}$ for some $s_i^h \in \overline{S}_i^h$.

Equilibria. A mixed strategy profile $\sigma = (\sigma_i)_{i\in I} \in \times_{i\in I}\Delta(S_i)$ is an *equilibrium* if, for all $i \in I$ and s_i with $\sigma_i(s_i) = 0$, $u_i(\sigma) \geq u_i(s_i, \sigma_{-i})$. Moreover, σ is SPE if it induces an equilibrium in every subgame. See Section 4 for a formal definition.

Beliefs. In this dynamic framework, beliefs are modeled as Conditional Probability Systems (Renyi, [17]; henceforth, CPS).

Definition 1 *A Conditional Probability System on* $(S_{-i}, (S_{-i}(h))_{h\in H})$ *is a mapping* $\mu(\cdot|\cdot) : 2^{S_{-i}} \times \{S_{-i}(h)\}_{h\in H} \to [0, 1]$ *satisfying the following axioms:*

CPS-1 for every $C \in (S_{-i}(h))_{h\in H}$*,* $\mu(\cdot|C)$ *is a probability measure on* S_{-i}*;*

CPS-2 for every $C \in (S_{-i}(h))_{h\in H}$*,* $\mu(C|C) = 1$*;*

CPS-3 for every $E \in 2^{S_{-i}}$ *and* $C, D \in (S_{-i}(h))_{h\in H}$*, if* $E \subseteq D \subseteq C$*, then* $\mu(E|C)=\mu(E|D)\mu(D|C)$*.*

The set of all CPS's on $(S_{-i}, (S_{-i}(h))_{h\in H})$ *is denoted by* $\Delta^H(S_{-i})$*.*

For brevity, the conditioning events will be indicated with just the information set, which represents all the information acquired by players through observation. For each set $J \subseteq I\backslash\{i\}$ of opponents of player i, and for each set of strategy sub-profiles $\overline{S}_J \subseteq S_J$, I say that a CPS $\mu_i \in \Delta^H(S_{-i})$ *strongly believes* $\overline{S}_J$ if, for all $h \in H(\overline{S}_J)$, $\mu_i(\overline{S}_J \times S_{I\backslash(J\cup\{i\})}|h) = 1$.

Rationality. I consider players who reply rationally to their conjectures. By rationality I mean that players, at every information set, choose an action that maximizes expected utility given the conjecture about how co-players will play and the expectation to reply rationally again in the continuation of the game. This is equivalent (see Battigalli [2]) to playing a *sequential best reply* to the CPS.

Definition 2 *Fix $\mu_i \in \Delta^H(S_{-i})$. A strategy $s_i \in S_i$ is a sequential best reply to μ_i if for every $h \in H(s_i)$,*[11] *s_i is a continuation best reply to $\mu_i(\cdot|h)$, i.e. for every $\widetilde{s}_i \in S_i(h)$,*

$$\sum_{s_{-i}\in S_{-i}(h)} u_i(\zeta(s_i, s_{-i}))\mu_i(s_{-i}|h) \geq \sum_{s_{-i}\in S_{-i}(h)} u_i(\zeta(\widetilde{s}_i, s_{-i}))\mu_i(s_{-i}|h).$$

I say that a strategy s_i is *rational* if it is a sequential best reply to some $\mu_i \in \Delta^H(S_{-i})$. The set of sequential best replies to μ_i is denoted by $\rho(\mu_i)$. For each $h \in H$, the set of continuation best replies to $\mu_i(\cdot|h)$ is denoted by $\widehat{r}(\mu_i, h)$. The set of best replies to a conjecture $\nu_i \in \Delta(S_{-i})$ in the normal form of the game is denoted by $r(\nu_i)$.

Elimination procedures.I provide a very general notion of elimination procedure for a subgame $\Gamma(h)$, which encompasses all the procedure I am ultimately interested in, or that will be needed for the proofs.

Definition 3 *Fix $h \in H$. An elimination procedure in $\Gamma(h)$ is a sequence $((S^h_{i,q})_{i\in I})^\infty_{q=0}$ where, for every $i \in I$,*

RP1 $S^h_{i,0} = S^h_i$;

RP2 $S^h_{i,n-1} \supseteq S^h_{i,n}$ for all $n \in \mathbb{N}$;

RP3 for every $s^h_i \in S^h_{i,\infty} = \cap_{n\in\mathbb{N}} S^h_{i,n}$, there exists μ^h_i that strongly believes $(S^h_{-i,q})^\infty_{q=0}$ such that $s^h_i \in \rho(\mu^h_i) \subseteq S^h_{i,\infty}$.

Lemma 1 *For every elimination procedure $((S^h_{i,q})_{i\in I})^\infty_{q=0}$ and every $\widehat{h} \succ h$, $((S^h_{i,q}(\widehat{h})|\widehat{h})_{i\in I})^\infty_{q=0}$ is an elimination procedure.*

Proof. See Catonini [6].

Indeed, elimination procedures have been defined purposedly to encompass the implications in the subgames of traditional elimination procedures for the whole game. In a subgame, substrategies can be eliminated "exogenously" and not because they are not sequential best replies to any valid conjecture in the subgame. On the other hand, substrategies can survive

[11] It would be totally immaterial to require s_i to be optimal also at the histories precluded by itself.

even if the opponents do not reach the subgame anymore. Note that the elimination can stop for some steps and then resume: for this reason, RP2 allows a weak inclusion at all steps

Now I can define Strong Rationalizability.

Definition 4 *Strong Rationalizability (Battigalli and Siniscalchi, [4]) is an elimination procedure* $((S_i^q)_{i\in I})_{q=0}^{\infty}$ *where for every* $n \in \mathbb{N}$ *and* $i \in I$, $s_i \in S_i^n$ *if and only if there exists* μ_i *that strongly believes* $(S_{-i}^q)_{q=0}^{n-1}$ *such that* $s_i \in \rho(\mu_i)$.

I will talk formally of first-order-belief restrictions on a path of play. To do so, I need the definition of Strong-Δ-Rationalizability.

Definition 5 *For each* $i \in I$, *fix* $\Delta_i \subseteq \Delta^H(S_{-i}^h)$. *Strong-Δ-Rationalizability (Battigalli [3], Battigalli and Siniscalchi [4]) is the elimination procedure* $((S_{i,\Delta}^q)_{i\in I})_{q=0}^{\infty}$ *such that, for every* $n \in \mathbb{N}$, $i \in I$, *and* $s_i \in S_i$, $s_i \in S_{i,n}$ *if and only if* $s_i \in \rho(\mu_i)$ *for some* $\mu_i \in \Delta_i$ *that strongly believes* $(S_{-i,\Delta}^q)_{q=0}^{n-1}$.

3 Main results

Consider an elimination procedure where no SPE is ever eliminated "exogenously". That is, for any belief that guarantees at least the expected payoff of a SPE of the game, its sequential best replies always survive the elimination step. Then, a Nash equilibrium that yields the same outcome distribution as a SPE survives the procedure.

Lemma 2 *Fix* $h \in H$ *and an elimination procedure* $((S_{i,q}^h)_{i\in I})_{q=0}^{\infty}$ *such that for each* $n \in \mathbb{N}$, $i \in I$, *SPE* $(\sigma_j^h)_{j\in I} \in \times_{j\in I}\Delta(S_j^h)$, *and* μ_i^h *that strongly believes* $(S_{-i}^h)_{q=0}^{n-1}$, *if*

$$\sum_{s_{-i}\in S_{-i}} u_i(\zeta(s_i, s_{-i}))\mu_i^h(s_{-i}|h^0) \geq u_i(\sigma^h),$$

then $\rho(\mu_i) \subseteq S_{i,n}^h$. *Thus, there exist a SPE* σ^h *and an equilibrium* $\widetilde{\sigma}^h \in \Delta(S_\infty^h)$ *such that* $\widetilde{\sigma}^h(S(z)) = \sigma^h(S(z))$ *for all* $z \in Z$.

The (very rough) intuition for this result is the following. Since a SPE can only be eliminated "endogeneously", if a SPE is eliminated at some step of the procedure, there must be a unilateral deviation from the SPE path(s) that is profitable whatever the deviator can expect thereafter. The key is to show that one of the possible post-deviation beliefs corresponds to an outcome distribution of a SPE of the subgame, which can be used to create a SPE of the whole game that substitutes the eliminated one. Note that the opponents may be surprised by the deviation and thus play any continuation best reply thereafter too. This creates a set of continuation plan profiles where all plans are best replies to some conjecture over the continuation plans of the others. In turn, this guarantees that any equilibrium in this reduced subgame is actually an equilibrium of the subgame. If an equilibrium is not a SPE, there exists a unilateral deviation from its path(s) that the deviator would take whatever SPE of the post-deviation subgame she expects thereafter. Thus, in this subgame, a SPE can only be eliminated "endogeneously", and we can iterate until subgames of depth 1 are reached. There, any equilibrium is a SPE. This intuition is refined here through an example, directly in the context of Theorem 1: the entire game Γ, and Strong Rationalizability as elimination procedure.

$\longleftarrow \quad C \quad — \quad Ann \quad — \quad D \quad \longrightarrow$

$A\backslash B$	H	I
E	$1,1$	$0,0$
F	$0,0$	$3,1$
G	$0,0$	$\cdot$ $\longrightarrow$

$A\backslash B$	N	O	P
L	$6,3$	$5,0$	$1,2$
M	$5,0$	$6,3$	$2,2$

$A\backslash B$	S	T	U
Q	$2,2$	$0,0$	$0,0$
R	$0,0$	$4,4$	$0,0$

Strong Rationalizability goes as follows. First, Bob eliminates all strategies that prescribe U. Second, Ann eliminates $C.E$. Third, Bob eliminates all strategies that prescribe H. Fourth, Ann eliminates $D.Q$. Fifth, Bob eliminates all strategies that prescribe S.

Note that at step 4, Ann eliminates a SPE path: $(D,(Q,S))$, which is sustained by the off-path equilibrium (E,H). In the proof of Lemma 9, which is more general than Lemma 2, the external recursive procedure (the one indexed by k) proceeds as follows for every elimination of a SPE path.

At step 3, Bob, while planning to play S after D, may play any surviving continuation plan $s_B^{(C)} \in S_B^3|(C)$ after C. Then, the elimination of $D.Q$ by Ann at step 4 implies that for each belief over $S_B^3|(C)$, Ann expects a higher payoff than under $(D,(Q,S))$, thus she may play any sequential best reply to such belief (Lemma 7). The two things together imply that the "reduced subgame" $S^3|(C)$ features all the sequential best replies of both players $i = Ann, Bob$ to all μ_i's that strongly believe $S_{-i}^3|(C), S_{-i}^2|(C), S_{-i}^1|(C)$. Then, every Nash equilibrium of the reduced subgame $S^3|(C)$ is a Nash equilibrium of the whole subgame. As in the internal recursive procedure in the proof of Lemma 9 (the one indexed by t), pick one of these Nash, say $(F, I.P)$. It does not induce a SPE path of $\Gamma((C))$. But then, there must exist a unilateral deviation from (F,I) which is profitable for any SPE of the subsequent subgame. Consider a deviation of Ann to G. Indeed, Ann prefers all the equilibria of $\Gamma((C,(G,I)))$ to (F,I) (for P cannot be played with probability higher than 1/2 in equilibrium). Then, Ann may play any best reply to an equilibrium conjecture of $\Gamma((C,(G,I)))$ (Lemma 8). Bob, while planning to play I, may play any surviving continuation plan $s_B^{(C,(G,I))} \in S_B^3|(C,(G,I))$ after $(C,(G,I))$ (Lemma 3). So, since $\Gamma((C,(G,I)))$ has depth 1, all its equilibria will survive until step 3. This is the basis step in the proof of Lemma 9; if the subgame had higher depth, the survival of *one* SPE path of the subgame would have been guaranteed by the Induction Hypothesis. Going back to the example, pick the surviving equilibrium (L, N) and look for a new Nash equilibrium of $\Gamma(C)$ that prescribes L and N: $(G.L, I.N)$. This time, it induces a SPE path of $\Gamma(C)$. So we found a Nash $s \in S^3|(C)$ inducing a SPE path of $\Gamma(C)$, which Ann prefers to $(D,(Q,S))$. So, if by backward induction we select $(G.L, I.N)$ after C and (Q, S) after D, we find a new SPE of Γ, which induces a different path than $(D,(Q,S))$.

The proof of Lemma 9 assumes by contradiction that all SPE paths are eliminated, and for each of them, the external recursive procedure finds as above a SPE subpath that is able to generate, by backward induction, a new SPE path. The external recursive procedure starts from (one of) the last SPE path(s) to be eliminated along the elimination procedure. So, the newly generated SPE path must have been eliminated no later than the one under consideration. Thus, the SPE subpath generated after the deviation that "killed" the old path survives until the step of elimination of the new path. This guarantees two things. First, if the deviation that kills the new path preceeds the deviation that killed the old path, the internal recursive procedure is then able to generate a Nash inducing a SPE subpath that is compatible with the SPE subpath already fixed after the old deviation. So, when the new SPE subpath will be imposed by backward induction, the resulting SPE paths are a subset of those generated by imposing the old SPE subpath, hence they are eliminated no later than the SPE path under consideration. Second, the new deviation cannot follow the old one, for after the old deviation, the fixed SPE subpath is induced by a surviving Nash. Then, the new deviations occur higher and higher in the tree, until the candidate deviations are actually exhausted and a contradiction arises.

By Lemma 2, the overlap between Strong Rationalizability and Subgame Perfect Equilibrium is immediate.

Theorem 1 *There exists a SPE $(\sigma_i)_{i\in I} \in \times_{i\in I}\Delta(S_i)$ and an equilibrium in strongly rationalizable strategies $(\widetilde{\sigma}_i)_{i\in I} \in \times_{i\in I}\Delta(S_i^{\infty})$ such that for every $z \in Z$, $\sigma(S(z)) = \widetilde{\sigma}(S(z))$.*

Proof. Lemma 2 can be applied with strong rationalizability as elimination procedure and $h := h^0$. ■

Fix now a path $z \in Z$ and, for each $i \in I$, let Δ_i be set of all CPS's μ_i such that $\mu_i(S_{-i}(z)|h) = 1$ for all $h \prec z$.[12] Suppose that Strong-Δ-Rationalizability yields z as unique prediction under $(\Delta_i)_{i\in I}$.[13] Then, no unilateral deviation from z has survived. This means that in each of the corresponding subgames, no SPE has survived. But then, by Lemma 2,

[12] I show in [6] that these restrictions are equivalent to strong belief in $S_{-i}(z)$.
[13] Or equivalently, by the results in [6], Selective Rationalizability.

there must be at least one SPE against which the deviator would not deviate. With this, I can construct a SPE with path z. Thus in absence of other belief restrictions, subgame perfection is a necessary (but not sufficient)[14] condition for players to certainly comply with the path they believe all opponents will comply with.

Theorem 2 *Fix $z \in Z$. If $\zeta(S_{\Delta}^{\infty}) = \{z\}$, then there exists a SPE*[15] $(\sigma_i)_{i\in I} \in \times_{i\in I}\Delta(S_i)$ *such that for every $i \in I$, $\sigma_i(S_i(z)) = 1$.*

Proof. Fix a history h that immediately follows a unilateral deviation by player $j \in I$ from z, i.e. $h \not\prec z$, $p(h) \prec z$, and $h \in (S_{-j}(z))$. Let $(S_n^h)_{n=0}^{\infty} = (S_{\Delta}^n(h)|h)_{n=0}^{\infty}$. Fix $m \in \mathbb{N}$, $i \neq j$, and μ_i^h t.s.b. $(S_{-i,n}^h)_{n=0}^{m-1}$. By $\zeta(S_{\Delta}^{\infty}) = \{z\}$ there exists μ_i t.s.b. $(S_{-i,\Delta}^n)_{n=0}^{m-1}$ such that $\mu_i(S_{-i}(h)|p(h)) = 0$ and $\rho(\mu_i)(h) \neq \emptyset$. Hence, there exists $\widetilde{\mu}_i$ t.s.b. $(S_{-i,\Delta}^n)_{n=0}^{m-1}$ with $\mu_i^h(s_{-i}^h|\widetilde{h}) = \widetilde{\mu}_i(\{s_{-i} \in S_{-i}(h) : s_{-i}|h = s_{-i}^h\}|\widetilde{h})$ for all $\widetilde{h} \in H^h$ such that $\emptyset \neq \rho(\widetilde{\mu}_i)(h) \subseteq S_{i,m}^h$. So the hypothesis of Lemma 2 holds for every $i \neq j$. By $\zeta(S_{\Delta}^{\infty}) = \{z\}$ there is $m \in \mathbb{N}$ such that $S_m^h = \emptyset$, so Lemma 2 cannot hold. Thus the hypothesis of Lemma 2 must be violated for j and some $v \leq m$, SPE σ^h, and μ_j^h t.s.b $(S_{-j,v}^h)_{q=0}^{v}$. Fix $s_{-j} \in S_{-j,\Delta}^{\infty} \subseteq S_{-j}(z)$. By $\zeta(S_{\Delta}^{\infty}) = \{z\}$, $r(s_{-j}) \subseteq S_j(z)$. For every s_{-j}^h with $\mu_j^h(s_{-j}^h|h) > 0$, by constructing beliefs as above, I obtain[16] $\widetilde{s}(s_{-j}^h) \in S_{-j,\Delta}^v(z)$ such that $\widetilde{s}(s_{-j}^h)(\widetilde{h}) = s_{-j}^h(\widetilde{h})$ for all $\widetilde{h} \in H^h$ and $\widetilde{s}(s_{-j}^h)(\widetilde{h}) = s_{-j}(\widetilde{h})$ for all $\widetilde{h} \notin H^h$. Fix μ_j t.s.b. $(S_{-j,v}^h)_{q=0}^{v}$ such that $\mu_j(\widetilde{s}(s_{-j}^h)|h^0) = \mu_j(\widetilde{s}(s_{-j}^h)|h) = \mu_j^h(s_{-j}^h|h)$. Then, $\mu_j \in \Delta_j$. Hence, by the violation of the hypothesis of Lemma 2, $\rho(\mu_j)(h) = \emptyset$. Thus, by $r(s_{-j}) \subseteq S_j(z)$, $\rho(\mu_j)(z) \neq \emptyset$. So, $u_j(z)$ is higher than j's expected payoff against μ_j^h, hence, by the violation of the hypothesis of Lemma 2, also than $u_j(\sigma^h)$. Repeating for all j and post-deviation h, z is a SPE path. ■

4 Proof of the main lemma.

Additional notation is needed. Fix $h \in H$, $(s_j^h)_{j\in I}, (\widetilde{s}_j^h)_{j\in I} \in S^h$, $\widehat{H} \subseteq H^h$, $i \in J \subseteq I$, $\mu_i^h \in \Delta^{H^h}(S_{-i}^h)$, $\widehat{h} \succeq h$, $(\widetilde{\sigma}_j^h)_{j\in I} \in \times_{j\in I}\Delta(S_j^h)$, and $(\sigma_j^{\widehat{h}})_{j\in I} \in \times_{j\in I}\Delta(S_j^{\widehat{h}})$.

- $s_J^h =^{\widehat{H}} s_J^{\widehat{h}}$ if for each $\widetilde{h} \in \widehat{H}$, $s_J^h(\widetilde{h}) = s_J^{\widehat{h}}(\widetilde{h})$;
- $\widetilde{\sigma}_J^h(\cdot) \in \Delta(S_J^h)$ is the product of the marginal distributions $(\widetilde{\sigma}_j^h)_{j\in J}$;[17]

[14] In [8] I show that not even when the SPE is unique, its path is necessarily delivered as unique predicition under the corresponding path restrictions.

[15] SPE in mixed strategies are formally defined in the next section, using additional notation. For every SPE in mixed strategy there is a SPE in behavioral strategies that induces the same outcome distribution, thus the theorem holds also with SPE in behavioral strategies.

[16] Formally, this is shown by Lemma 3.

[17] This is an exception to the rule of subscripts: $\widetilde{\sigma}_J^h$ is not a (sub-)profile of individual distributions but an uncorrelated joint distribution. Equilibria $(\widetilde{\sigma}_j^h)_{j\in I}$ will be represented as the joint uncorrelated distribution $\widetilde{\sigma}^h$ they induce.

- $H(\widetilde{\sigma}_J^h) := H(\mathrm{Supp}\widetilde{\sigma}_J^h)$, $\widetilde{\sigma}_J^h[\widehat{h}] := (\mathrm{Supp}\widetilde{\sigma}_J^h)[\widehat{h}]$;
- $D_i(\widetilde{\sigma}^h) := \{\widetilde{h} \in H(\widetilde{\sigma}_{-i}^h) \backslash H(\widetilde{\sigma}^h) : p(\widetilde{h}) \in H(\widetilde{\sigma}^h)\}$;[18]
- $D^{-i}(\widetilde{\sigma}^h) := \{\widetilde{h} \in H \backslash H(\widetilde{\sigma}^h) : p(\widetilde{h}) \in H(\widetilde{\sigma}^h) \wedge \widetilde{h} \in H(\widetilde{\sigma}_i^h)\}$;
- $\widetilde{\sigma}_J^h|\widehat{h}$ is the product of $(\widetilde{\sigma}_j^h|\widehat{h})_{j\in J}$ and $\widetilde{\sigma}_i^h|\widehat{h} \in \Delta(S_i^{\widehat{h}})$ is def. for every $s_i^{\widehat{h}} \in S_i^{\widehat{h}}$ as:
  - $(\widetilde{\sigma}_i^h|\widehat{h})(s_i^{\widehat{h}}) = \widetilde{\sigma}_i^h(\{s_i^h \in S_i^h(\widehat{h}) : s_i^h|\widehat{h} = s_i^{\widehat{h}}\})/\widetilde{\sigma}_i^h(S_i^h(\widehat{h}))$ if $\widehat{h} \in H(\widetilde{\sigma}_i^h)$,
  - $(\widetilde{\sigma}_i^h|\widehat{h})(s_i^{\widehat{h}}) = \widetilde{\sigma}_i^h(\{s_i^h \in S_i^h : s_i^h|\widehat{h} = s_i^{\widehat{h}}\})$ else;
- $\widetilde{\sigma}_J^h =^* \sigma_J^{\widehat{h}}$ if for every $z \in \zeta(\sigma^{\widehat{h}})$ and $j \in J$, $(\widetilde{\sigma}_j^h|\widehat{h})(S_j^{\widehat{h}}(z)) = \sigma_j^{\widehat{h}}(S_j^{\widehat{h}}(z))$;[19]
- $\widetilde{\sigma}_J^h =^{\widehat{h}} \sigma_J^{\widehat{h}}$ if for every $z \succ \widehat{h}$ and $j \in J$, $(\widetilde{\sigma}_j^h|\widehat{h})(S_j^{\widehat{h}}(z)) = \sigma_j^{\widehat{h}}(S_j^{\widehat{h}}(z))$;
- $\mu_i^h =^* \sigma_{-i}^{\widehat{h}}$ if $\mu_i^h(\cdot|\widehat{h}) =^* \sigma_{-i}^{\widehat{h}}$; $\mu_i^h =^{\widehat{h}} \sigma_{-i}^{\widehat{h}}$ if $\mu_i^h(\cdot|\widehat{h}) =^{\widehat{h}} \sigma_{-i}^{\widehat{h}}$;
- $\pi_i(\widetilde{\sigma}^h)$ is i's exp. payoff under $\widetilde{\sigma}^h$; $\pi(\widetilde{\sigma}_{-i}^h) := \max_{\widetilde{s}_i^h \in S_i^h} \sum_{s_{-i}^h \in \mathrm{Supp}\widetilde{\sigma}_{-i}^h} u_i(\zeta(\widetilde{s}_i^h, s_{-i}^h))\widetilde{\sigma}_{-i}^h(s_{-i}^h)$;
- $\widetilde{\sigma}_{-i}^h|\widehat{h} \geq \sigma_{-i}^{\widehat{h}}$ if $\pi(\widetilde{\sigma}_{-i}^h|\widehat{h}) \geq \pi(\sigma_{-i}^{\widehat{h}})$; $\mu_i^h \geq \widetilde{\sigma}_{-i}^h$ if $\mu_i^h(\cdot|h) \geq \widetilde{\sigma}_{-i}^h$;
- $\widetilde{\sigma}^h$ is a SPE of $\Gamma(h)$ if for every $\widetilde{h} \in H^h$, $\widetilde{\sigma}^h|\widetilde{h}$ is an equilibrium of $\Gamma(\widetilde{h})$;
- for any set of unordered[20] non-terminal histories $\widetilde{H} \subseteq H$ and any set of SPE $\Sigma^{\widetilde{H}} = (\sigma^{\widetilde{h}})_{\widetilde{h}\in\widetilde{H}}$ of the corresponding subgames, $E^h(\Sigma^{\widetilde{H}})$ is the set of SPE of $\Gamma(h)$ σ^h such that for every $\widetilde{h} \in \widetilde{H} \cap H^h$, $\sigma^h|\widetilde{h} = \sigma^{\widetilde{h}}$.

I will use the fact that $=^*$ and $=^{\widehat{h}}$ are transitive and that $=^{\widehat{h}}$ implies $=^*$. Moreover, when $\widetilde{\sigma}_i^h =^* \sigma_i^{\widehat{h}}$:

♡ for every $\widetilde{h} \succ \widehat{h}$ with $p(\widetilde{h}) \in H(\sigma^{\widehat{h}})$ and $\widehat{h} \preceq \overline{h} \prec \widetilde{h}$, $(\widetilde{\sigma}_i^h|\overline{h})(S_i^{\overline{h}}(\widetilde{h})) = (\sigma_i^{\widehat{h}}|\overline{h})(S_i^{\overline{h}}(\widetilde{h}))$;

♠ if $\widetilde{\sigma}_i^h|\widetilde{h} = \sigma_i^{\widehat{h}}|\widetilde{h}$ for all $\widetilde{h} \in D^{-i}(\sigma^{\widehat{h}})$, then $\widetilde{\sigma}_i^h =^{\widehat{h}} \sigma_i^{\widehat{h}}$.[21]

When $\sigma^{\widehat{h}}$ is an equilibrium and $\widetilde{\sigma}^h =^{\widehat{h}} \sigma^{\widehat{h}}$, I will often use the fact that $\widetilde{\sigma}^h|\widetilde{h}$ is an equilibrium for all $\widetilde{h} \in H(\sigma^{\widehat{h}})$. Moreover:

♦ if $\widetilde{\sigma}_{-i}^{\widehat{h}} =^* \sigma_{-i}^{\widehat{h}}$, $\pi(\widetilde{\sigma}_{-i}^{\widehat{h}}) \geq \pi(\sigma_{-i}^{\widehat{h}}) = \pi_i(\sigma^{\widehat{h}})$ and if $\widetilde{\sigma}^{\widehat{h}}$ is an equil., $\pi(\widetilde{\sigma}_{-i}^{\widehat{h}}) = \pi(\sigma_{-i}^{\widehat{h}})$;

[18] Unilateral deviations by player i from the paths induced by $\widetilde{\sigma}^h$.

[19] Notie that only the outcomes induced with positive probability by $\sigma^{\widehat{h}}$ and not all those compatible with $\sigma_J^{\widehat{h}}$ matter.

[20] For every two histories h, h' in the set, $h \not\succeq h'$ and $h' \not\succeq h$.

[21] By $\widetilde{\sigma}_i^h =^* \sigma_i^{\widehat{h}}$ and ♡, $(\widetilde{\sigma}_i^h|\widehat{h})(S_i^{\widehat{h}}(\widetilde{h})) = \sigma_i^{\widehat{h}}(S_i^{\widehat{h}}(\widetilde{h}))$. For every $z \succ \widetilde{h}$, by $\widetilde{\sigma}_i^h|\widetilde{h} = \sigma_i^{\widehat{h}}|\widetilde{h}$, $(\widetilde{\sigma}_i^h|\widetilde{h})(S_i^{\widetilde{h}}(z)) = (\sigma_i^{\widehat{h}}|\widetilde{h})(S_i^{\widetilde{h}}(z))$. Together, $(\widetilde{\sigma}_i^h|\widehat{h})(S_i^{\widehat{h}}(z)) = \sigma_i^{\widehat{h}}(S_i^{\widehat{h}}(z))$.

♣ if $\mu_i^{\widehat{h}} =^{\widehat{h}} \sigma_{-i}^{\widehat{h}}$, for every $s_i^{\widehat{h}} \in \mathrm{Supp}\sigma_i^{\widehat{h}}$, there is $\widetilde{s}_i^{\widehat{h}} \in \rho(\mu_i^{\widehat{h}})$ such that $\widetilde{s}_i^{\widehat{h}} =^{H(\sigma^{\widehat{h}})} s_i^{\widehat{h}}$.

I start with a lemma about the combination of substrategies. Consider a player who may be surprised by history $\widehat{h}$, in the sense that she may play a strategy that allows $\widehat{h}$ while believing that the co-players do not until $\widehat{h}$ is actually reached. This player can keep the same beliefs and the same strategy out of $\Gamma(\widehat{h})$, whatever she believes the co-players would do and hence however she may play after $\widehat{h}$. The proof is immediate and is here omitted. However, a formal proof can be found in [6].

Lemma 3 *Fix an elimination procedure $((S_{i,q}^h)_{i\in I})_{q\geq 0}$, $i \in I$, $n \in \mathbb{N}$, $\widehat{h} \in H^h$, and μ_i^h t.s.b. $(S_{-i,q}^h)_{q=0}^{n-1}$ such that $\mu_i^h(S_{-i}^h(\widehat{h})|p(\widehat{h})) = 0$. Fix $s_i^h \in \rho(\mu_i^h)$, $\mu_i^{\widehat{h}}$ t.s.b. $(S_{-i,q}^h(\widehat{h})|\widehat{h})_{q=0}^{n-1}$ and $s_i^{\widehat{h}} \in \rho(\mu_i^{\widehat{h}})$.*

Consider the unique $\widetilde{s}_i^h =^{\widehat{h}} s_i^{\widehat{h}}$ such that for every $\widetilde{h} \notin H^{\widehat{h}}$, $\widetilde{s}_i^h(\widetilde{h}) = s_i^h(\widetilde{h})$.

There exists $\widetilde{\mu}_i^h =^{\widehat{h}} \mu_i^{\widehat{h}}$ t.s.b. $(S_{-i,q}^h)_{q=0}^{n-1}$ such that for every $\widetilde{h} \notin H^{\widehat{h}}$, $\widetilde{\mu}_i^h(\cdot|\widetilde{h}) = \mu_i^h(\cdot|\widetilde{h})$, and $\widetilde{s}_i^h \in \rho(\widetilde{\mu}_i^h)$ (so that $\rho(\mu_i^h)(\widehat{h}) \neq \emptyset$ implies $\rho(\widetilde{\mu}_i^h)(\widehat{h}) \neq \emptyset$).

For the next five lemmata, fix $n \in \mathbb{N}$, $h \in H$, an elimination procedure $(S_q^h)_{q\geq 0}$ and $\widetilde{\sigma}^h \in \Delta(S_{n-1}^h)$. Let $H^\sigma := H(\widetilde{\sigma}^h)$. For every $i \in I$, let $D_i := D_i(\widetilde{\sigma}^h)$ and $D^{-i} := D^{-i}(\widetilde{\sigma}^h)$.

This Lemma exploits Lemma 3 to combine different reactions of player i to unexpected deviations from H^σ.

Lemma 4 *Fix $v \leq n$ and $i \in I$ such that $\widetilde{\sigma}_i^h \in \Delta(r(\widetilde{\sigma}_{-i}^h))$ and for every $\mu_i^h =^* \widetilde{\sigma}_{-i}^h$ t.s.b. $(S_{-i,q}^h)_{q=0}^{v-1}$, $\rho(\mu_i^h) \subseteq S_{i,v}^h$. For every $\widetilde{h} \in D^{-i}$, fix $\widehat{\sigma}_i^{\widetilde{h}} \in \Delta(S_{i,v}^h(\widetilde{h})|\widetilde{h})$. There exists $\widehat{\sigma}_i^h \in \Delta(S_{i,v}^h)$ such that $\widehat{\sigma}_i^h =^* \widetilde{\sigma}_i^h$ and for every $\widetilde{h} \in D^{-i}$, $\widehat{\sigma}_i^h|\widetilde{h} = \widehat{\sigma}_i^{\widetilde{h}}$.*

Proof.

I show that for every $i \in I$, $\overline{s}_i^h \in \mathrm{Supp}\widetilde{\sigma}_i^h$, and $\varsigma : \widetilde{h} \in D^{-i} \mapsto s_i^{\widetilde{h}} \in \mathrm{Supp}\widehat{\sigma}_i^{\widetilde{h}}$, there exists $\widehat{s}_i^h \in S_{i,v}^h$ such that $\widehat{s}_i^h =^{H^\sigma} \overline{s}_i^h$ and $\widehat{s}_i^h =^{\widetilde{h}} \varsigma(\widetilde{h})$ for all $\widetilde{h} \in D^{-i}$. Using all such $\widehat{s}_i^h$'s, it is easy to construct the desired $\widehat{\sigma}_i^h$.

Fix $\mu_i^h =^h \widetilde{\sigma}_{-i}^h$ t.s.b. $(S_{-i,q}^h)_{q=0}^{v-1}$. Since $\overline{s}_i^h \in r(\widetilde{\sigma}_{-i}^h)$, by ♣ there exists $s_i^h \in \rho(\mu_i^h)$ such that for every $\widetilde{h} \in H^\sigma$, $s_i^h(\widetilde{h}) = \overline{s}_i^h(\widetilde{h})$. For every $\widetilde{h} \in D^{-i}$, $\varsigma(\widetilde{h}) \in \mathrm{Supp}\widehat{\sigma}_i^{\widetilde{h}} \subseteq S_{i,v}^h(\widetilde{h})|\widetilde{h}$, so there exists $\mu_i^{\widetilde{h}}$ t.s.b. $(S_{-i,q}^h(\widetilde{h})|\widetilde{h})_{q=0}^{v-1}$ such that $\varsigma(\widetilde{h}) \in \rho(\mu_i^{\widetilde{h}})$. Thus, by repeatedly applying Lemma 3, I can find $\widetilde{\mu}_i^h =^h \widetilde{\sigma}_{-i}^h$ t.s.b. $(S_{-i,q}^h)_{q=0}^{v-1}$ and $\widehat{s}_i^h \in \rho(\widetilde{\mu}_i^h)$ such that $\widehat{s}_i^h =^{H^\sigma} \overline{s}_i^h$ and $\widehat{s}_i^h =^{\widetilde{h}} \varsigma(\widetilde{h})$ for all $\widetilde{h} \in D^{-i}$. By hypothesis of the Lemma, $\rho(\widetilde{\mu}_i^h) \subseteq S_{i,v}^h$. ■

For the next four lemmata suppose that $\widetilde{\sigma}^h$ is an equilibrium and:

A0 for every $v \leq n$, $i \in I$ and $\mu_i^h =^* \widetilde{\sigma}_{-i}^h$ t.s.b. $(S_{-i,q}^h)_{q=0}^{v-1}$, $\rho(\mu_i^h) \subseteq S_{i,v}^h$, so that every $i \in I$ satisfies the hypotheses of Lemma 4.

Lemma 5 is a characterization of equilibrium which will turn out to be useful. Since the arguments for it are standard, the proof is omitted.

Lemma 5 *Fix $(\widehat{\sigma}_i^h)_{i\in I} \in \times_{i\in I}\Delta(S_i^h)$: $\widehat{\sigma}^h$ is an equilibrium if and only if for every $\overline{h} \in H(\widehat{\sigma}^h)$, $i \in I$ and $a_i \in A_i(\overline{h})\backslash\widehat{\sigma}_i^h[\overline{h}]$, calling $H_{a_i}^{\overline{h}} := (\overline{h},(a_i,a_{-i}))_{a_{-i}\in\widehat{\sigma}_{-i}^h[\overline{h}]}$,*

$$\sum_{\widetilde{h}\in H_{a_i}^{\overline{h}}\backslash Z} \pi(\widehat{\sigma}_{-i}^h|\widetilde{h}) \cdot (\widehat{\sigma}_{-i}^h|\overline{h})(S_{-i}^{\overline{h}}(\widetilde{h})) + \sum_{z\in H_{a_i}^{\overline{h}}\cap Z} u_i(z)\cdot(\widehat{\sigma}_{-i}^h|\overline{h})(S_{-i}^{\overline{h}}(z)) \leq \pi_i(\widehat{\sigma}^h|\overline{h}). \qquad (\bigstar)$$

Lemma 6 converts a condition on CPS's into $\bigstar$ for some related conjectures.

Lemma 6 *Fix $\widehat{\sigma}^h =^* \widetilde{\sigma}^h$, $\overline{h} \in H^\sigma$, $i \in I$, $a_i \in A_i(\overline{h})\backslash\widehat{\sigma}_i^h[\overline{h}]$, $\widehat{h} \in H_{a_i}^{\overline{h}}\backslash Z$, $v \leq n$, and $\widetilde{\mu}_i^{\widehat{h}}$ t.s.b. $(S_{-i,q}^h(\widehat{h})|\widehat{h})_{q=0}^v$ such that (i) $\widehat{\sigma}_{-i}^h|\widehat{h} \leq \widetilde{\mu}_i^{\widehat{h}}$, (ii) for every $\widetilde{h} \in H_{a_i}^{\overline{h}}\backslash(Z\cup\{\widehat{h}\})$, $\widehat{\sigma}_{-i}^h|\widetilde{h} \leq \overline{\sigma}_{-i}^{\widetilde{h}}$ for some $\overline{\sigma}_{-i}^{\widetilde{h}} \in \Delta(S_{-i,v}^h(\widetilde{h})|\widetilde{h})$, and (iii) for every $\mu_i^h =^* \widetilde{\sigma}_{-i}^h$ t.s.b. $(S_{-i,q}^h)_{q=0}^{v-1}$, if $\mu_i^h =^{\widehat{h}} \widetilde{\mu}_i^{\widehat{h}}$, then $\rho(\mu_i^h)(\widehat{h}) = \emptyset$. Then $\bigstar$ holds.*

Proof. Let $\overline{\sigma}_{-i}^{\widehat{h}} := \widetilde{\mu}_i^{\widehat{h}}(\cdot|\widehat{h})$. By Lemma 4 there exists $\overline{\sigma}_{-i}^h \in \Delta(S_v^h)$ such that $\overline{\sigma}_{-i}^h =^* \widetilde{\sigma}_{-i}^h$, $\overline{\sigma}_{-i}^h|\widetilde{h} = \overline{\sigma}_{-i}^{\widetilde{h}}$ for all $\widetilde{h} \in H_{a_i}^{\overline{h}}\backslash Z$, and $\overline{\sigma}_{-i}^h|\widetilde{h} = \widetilde{\sigma}_{-i}^h|\widetilde{h}$ for all $\widetilde{h} \in D_i\backslash H_{a_i}^{\overline{h}}$. Fix $\mu_i^h =^h \overline{\sigma}_{-i}^h$ t.s.b. $(S_{-i,q}^h)_{q=0}^v$ such that $\mu_i^h =^{\widehat{h}} \widetilde{\mu}_i^{\widehat{h}}$ (one exists because $\mu_i^h(\cdot|h) = \overline{\sigma}_{-i}^h$ implies $\mu_i^h(\cdot|\widehat{h})|\widehat{h} = \overline{\sigma}_{-i}^h|\widehat{h} = \widetilde{\mu}_i^{\widehat{h}}(\cdot|\widehat{h})$).

For every $\widetilde{h} \in D_i\backslash H_{a_i}^{\overline{h}}$ and $z \succ \widetilde{h}$,

$$\overline{\sigma}_{-i}^h(S_{-i}^h(z)) = \overline{\sigma}_{-i}^h(S_{-i}^h(\widetilde{h}))\cdot(\overline{\sigma}_{-i}^h|\widetilde{h})(S_{-i}^{\widetilde{h}}(z)) = \widetilde{\sigma}_{-i}^h(S_{-i}^h(\widetilde{h}))\cdot(\widetilde{\sigma}_{-i}^h|\widetilde{h})(S_{-i}^{\widetilde{h}}(z)) = \widetilde{\sigma}_{-i}^h(S_{-i}^h(z)),$$

where the first and the last equalities are by the chain rule of probability, and the central equality is by $\heartsuit$ and by construction. For every $z \in \zeta(\mathrm{Supp}\widetilde{\sigma}^h)$, by $\overline{\sigma}^h =^* \widetilde{\sigma}^h$, $\overline{\sigma}_{-i}^h(S_{-i}^h(z)) = \widetilde{\sigma}_{-i}^h(S_{-i}^h(z))$. Hence every $s_i^h \notin S_i^h(\widehat{h}) = \cup_{\widetilde{h}\in H_{a_i}^{\overline{h}}} S_i^h(\widetilde{h})$ induces with $\overline{\sigma}_{-i}^h$ and $\widetilde{\sigma}_{-i}^h$ (1) and with $\overline{\sigma}_{-i}^h|\overline{h}$ and $\widetilde{\sigma}_{-i}^h|\overline{h}$ (2) the same distribution over outcomes. By equilibrium, $r(\widetilde{\sigma}_{-i}^h)(\overline{h}) \neq \emptyset$; by 1, $r(\overline{\sigma}_{-i}^h)(\overline{h}) \cup r(\overline{\sigma}_{-i}^h)(\widehat{h}) \neq \emptyset$; by $\overline{h} \prec \widehat{h}$, $r(\overline{\sigma}_{-i}^h)(\overline{h}) \neq \emptyset$; by $\overline{h} \in H(\overline{\sigma}_{-i}^h)$, $\rho(\mu_i^h)(\overline{h}) \neq \emptyset$; by $\rho(\mu_i^h)(\widehat{h}) = \emptyset$, $\widehat{r}(\mu_i^h,\overline{h})(\widehat{h}) = \emptyset$; by $\mu_i^h(\cdot|\overline{h})|\overline{h} = \overline{\sigma}_{-i}^h|\overline{h}$, $r(\overline{\sigma}_{-i}^h|\overline{h})(\widehat{h}) = \emptyset$; thus,

$$\sum_{\widetilde{h}\in H_{a_i}^{\overline{h}}\backslash Z} \pi(\overline{\sigma}_{-i}^{\widetilde{h}}) \cdot (\overline{\sigma}_{-i}^h|\overline{h})(S_{-i}^{\overline{h}}(\widetilde{h})) + \sum_{z\in H_{a_i}^{\overline{h}}\cap Z} u_i(z)\cdot(\overline{\sigma}_{-i}^h|\overline{h})(S_{-i}^{\overline{h}}(z)) \leq \pi_i(\overline{\sigma}^h|\overline{h}). \qquad (\blacktriangledown)$$

By equilibrium, $r(\widetilde{\sigma}_{-i}^h|\overline{h})\backslash S_i(\widehat{h}) \neq \emptyset$; by $r(\overline{\sigma}_{-i}^h|\overline{h})(\widehat{h}) = \emptyset$ and 2, $\pi(\overline{\sigma}_{-i}^h|\overline{h}) = \pi(\widetilde{\sigma}_{-i}^h|\overline{h})$; by equilibrium $\pi(\widetilde{\sigma}_{-i}^h|\overline{h}) = \pi_i(\widetilde{\sigma}^h|\overline{h}) = \pi_i(\widehat{\sigma}^h|\overline{h})$ (by $\widetilde{\sigma}^h|\overline{h} =^* \widehat{\sigma}^h|\overline{h}$) By $\overline{\sigma}_{-i}^h =^* \widetilde{\sigma}_{-i}^h =^* \widehat{\sigma}_{-i}^h$ and $\heartsuit$, for every $\widetilde{h} \in H_{a_i}^{\overline{h}}$, $(\overline{\sigma}_{-i}^h|\overline{h})(S_{-i}^{\overline{h}}(\widetilde{h})) = (\widehat{\sigma}_{-i}^h|\overline{h})(S_{-i}^{\overline{h}}(\widetilde{h}))$. For each $\widetilde{h} \in H_{a_i}^{\overline{h}}\backslash Z$, $\overline{\sigma}_{-i}^{\widetilde{h}} \geq \widehat{\sigma}_{-i}^h|\widetilde{h}$. So $\blacktriangledown$ implies $\bigstar$. $\blacksquare$

For the next two Lemmata, fix a set of unordered histories $\widehat{H} \subseteq H^h$ and a set of SPE $\Sigma^{\widehat{H}} = (\sigma^{\widetilde{h}})_{\widetilde{h}\in\widehat{H}}$ such that:

A1 for every $\widetilde{h} \in \widehat{H}$, there exists an equilibrium $\widetilde{\sigma}^{\widetilde{h}} =^* \sigma^{\widetilde{h}}$ such that $\widetilde{\sigma}^{\widetilde{h}} \in \Delta(S_n^h(\widetilde{h})|\widetilde{h})$.

Lemma 7 claims that if only until step $n-1$ an equilibrium that mimics a given SPE survives, at step n one deviation is always *strictly* preferred to continuing as the equilibrium prescribes.

Lemma 7 *Suppose that there exists $\sigma^h \in E^h(\Sigma^{\widehat{H}})$ such that $\widetilde{\sigma}^h =^* \sigma^h$ but there is no equilibrium $\widehat{\sigma}^h \in \Delta(S_n^h)$ such that $\widehat{\sigma}^h =^* \sigma^h$. Then there exist $l \in I$ and $\widehat{h} \in D_l \backslash (\cup_{\widetilde{h}\in\widehat{H}} H^{\widetilde{h}})$ such that for every $(\widehat{\sigma}_j^{\widehat{h}})_{j\neq l} \in \times_{j\neq l}\Delta(S_{j,n}^h(\widehat{h})|\widehat{h})$,[22] $v \leq n$ and $\widetilde{\mu}_l^{\widehat{h}} \geq \widehat{\sigma}_{-l}^{\widehat{h}}$ t.s.b. $(S_{-l,q}^h(\widehat{h})|\widehat{h})_{q=0}^v$, there exists $\widetilde{\mu}_l^h =^* \widetilde{\sigma}_{-l}^h$ t.s.b. $(S_{-l,q}^h)_{q=0}^{v-1}$ such that $\widetilde{\mu}_l^h =^{\widehat{h}} \widetilde{\mu}_l^{\widehat{h}}$ and $\rho(\widetilde{\mu}_l^h)(\widehat{h}) \neq \emptyset$ (so by A0 $\rho(\widetilde{\mu}_l^h) \subseteq S_{l,v}^h$).*

Proof. Suppose not. For every $i \in I$ and $\widehat{h} \in D_i \backslash (\cup_{\widetilde{h}\in\widehat{H}} H^{\widetilde{h}}) =: \overline{D}_i$, fix $\widehat{\sigma}_{-i}^{\widehat{h}}$, $v(\widehat{h})$, and $\widetilde{\mu}_i^{\widehat{h}}$ that violate the statement, and let $\widetilde{\sigma}_{-i}^{\widehat{h}} := \widehat{\sigma}_{-i}^{\widehat{h}}$. By Lemma 4 there exists $\widehat{\sigma}^h \in \Delta(S_n^h)$ such that $\widehat{\sigma}^h =^* \widetilde{\sigma}^h =^* \sigma^h$ and for every $i \in I$, $\widetilde{h} \in \widehat{H} \cup (\cup_{j\neq i}\overline{D}_j)$, and $\widehat{h} \in D^{-i} \cap H^{\widetilde{h}}$, $\widehat{\sigma}_i^h|\widehat{h} = \widetilde{\sigma}_i^{\widetilde{h}}|\widehat{h}$. I show that $\widehat{\sigma}^h$ is an equilibrium.

Fix $\overline{h} \in H^\sigma$, $i \in I$, and $a_i \in A_i(\overline{h})\backslash\widetilde{\sigma}_i^h[\overline{h}]$. If there exists $\widetilde{h} \preceq \overline{h}$ such that $\widetilde{h} \in \widehat{H}$, $\widehat{\sigma}^h =^* \sigma^h =^{\widetilde{h}} \sigma^{\widetilde{h}} =^* \widetilde{\sigma}^{\widetilde{h}}$, so, by ♠, $\widehat{\sigma}^h =^{\widetilde{h}} \widetilde{\sigma}^{\widetilde{h}}$; then $\widehat{\sigma}^h|\overline{h}$ is an equilibrium, so by Lemma 5 ("only if") ★ holds. If $H_{a_i}^{\overline{h}}\backslash Z \subseteq \widehat{H}$, for every $\widehat{h} \in H_{a_i}^{\overline{h}}\backslash Z$, $\widehat{\sigma}_{-i}^h|\widehat{h} = \widetilde{\sigma}_{-i}^{\widehat{h}}$ and by ♦, $\pi(\widetilde{\sigma}_{-i}^{\widehat{h}}) = \pi(\sigma_{-i}^h|\widehat{h})$; by $\widehat{\sigma}^h =^* \sigma^h$, $\pi_i(\widehat{\sigma}^h|\overline{h}) = \pi_i(\sigma^h|\overline{h})$ and by ♡, $(\widehat{\sigma}_{-i}^h|\overline{h})(S_{-i}^{\overline{h}}(\widehat{h})) = (\sigma_{-i}^h|\overline{h})(S_{-i}^{\overline{h}}(\widehat{h}))$; so since $\sigma^h|\overline{h}$ is an equilibrium by Lemma 5 ("only if") ★ holds. If $H_{a_i}^{\overline{h}} \cap \overline{D}_i \neq \emptyset$, fix $v := \min_{\widetilde{h}\in H_{a_i}^{\overline{h}}\cap\overline{D}_i} v(\widetilde{h})$ and $\widehat{h} := \arg\min_{\widetilde{h}\in H_{a_i}^{\overline{h}}\cap\overline{D}_i} v(\widetilde{h})$: for every $\widetilde{h} \in H_{a_i}^{\overline{h}} \cap \overline{D}_i$, $\widehat{\sigma}_{-i}^h|\widetilde{h} \leq \widetilde{\mu}_i^{\widetilde{h}}(\cdot|\widetilde{h}) \in \Delta(S_{-i,v}^h(\widetilde{h})|\widetilde{h}) \neq \emptyset$[23] and for every $\widetilde{h} \in H_{a_i}^{\overline{h}}\backslash(Z \cup \overline{D}_i)$, $\widehat{\sigma}_{-i}^h|\widetilde{h} \in \Delta(S_{-i,v}^h(\widetilde{h})|\widetilde{h})$; therefore by Lemma 6 ★ holds. Thus by Lemma 5 ("if"), $\widehat{\sigma}^h$ is an equilibrium. ■

Lemma 8 is the "dual" of Lemma 7: if an equilibrium has survived n steps but it does not mimic a SPE (within a subset), then there is a deviation from one of the equilibrium paths that the deviator could take whenever thereafter she expects at least the payoff of a SPE of the subgame (within a subset).

Lemma 8 *Suppose that $\widetilde{\sigma}^h \in \Delta(S_n^h)$ and for every $\widetilde{h} \in \widehat{H}$, $\widetilde{\sigma}^h =^{\widetilde{h}} \widetilde{\sigma}^{\widetilde{h}}$, but there is no $\sigma^h \in E^h(\Sigma^{\widehat{H}})$ with $\widetilde{\sigma}^h =^* \sigma^h$. Then there exist $p \in I$ and $\overline{h} \in D_p(\widetilde{\sigma}^h)\backslash(\cup_{\widetilde{h}\in\widehat{H}} H^{\widetilde{h}})$ such that*

[22] Note that the statement needs not hold for all $\widehat{\sigma}_{-l}^{\widehat{h}} \in \Delta(S_{-j,n}^h(\widehat{h})|\widehat{h})$: this is due to the fact that equilibria are not correlated.

[23] For every $j \neq i$, there is $\mu_j^h =^h \widetilde{\sigma}_{-j}^h$ t.s.b. $(S_{-l,q}^h)_{q=0}^{n-1}$; by ♣, $\rho(\mu_j^h)(\widetilde{h}) \neq \emptyset$; by A0 $\rho(\mu_j^h) \subseteq S_{j,n}^h$.

for every $\sigma^{\overline{h}} \in E^{\overline{h}}(\Sigma^{\widehat{H}})$, $v \leq n$ and $\widetilde{\mu}_p^{\overline{h}} \geq \sigma_{-p}^{\overline{h}}$ t.s.b. $(S_{-p,q}^h(\overline{h})|\overline{h})_{q=0}^v$, there exists $\widetilde{\mu}_p^h =^* \widetilde{\sigma}_{-p}^h$ t.s.b. $(S_{-p,q}^h)_{q=0}^{v-1}$ such that $\widetilde{\mu}_p^h =^{\overline{h}} \widetilde{\mu}_p^{\overline{h}}$ and $\rho(\widetilde{\mu}_p^h)(\overline{h}) \neq \emptyset$ (so by A0 $\rho(\widetilde{\mu}_p^h) \subseteq S_{p,v}^h$).

Proof. Suppose not. For every $i \in I$ and $\widehat{h} \in D_i \backslash (\cup_{\widetilde{h} \in \widehat{H}} H^{\widetilde{h}}) =: \overline{D}_i$, fix $\sigma^{\widehat{h}}$, $v(\widehat{h})$, and $\widetilde{\mu}_i^{\widehat{h}}$ that violate the statement. Construct $\widehat{\sigma}^h =^* \widetilde{\sigma}^h$ such that for every $\widetilde{h} \notin H^\sigma$ with $p(\widetilde{h}) \in H^\sigma$, $\widehat{\sigma}^h|\widetilde{h} \in E^{\widetilde{h}}(\Sigma^{\widehat{H}})$, and in particular, for every $i \in I$, $\widetilde{h} \in \widehat{H} \cup (\cup_{j \in I} \overline{D}_j)$, and $\widehat{h} \in D^{-i} \cap H^{\widetilde{h}}$, $\widehat{\sigma}_i^h|\widehat{h} = \sigma_i^{\widetilde{h}}|\widehat{h}$. I show that $\widehat{\sigma}^h$ is an equilibrium with $\widehat{\sigma} =^{\widetilde{h}} \sigma^{\widetilde{h}} \in E^{\widetilde{h}}(\Sigma^{\widehat{H}})$ for all $\widetilde{h} \in \widehat{H} \cap H^\sigma$. Then, by Lemma 5 ("only if") for every $\widetilde{h} \in H^\sigma$, $\widehat{\sigma}^h|\widetilde{h}$ is an equilibrium, and so $\widehat{\sigma}^h \in E^h(\Sigma^{\widehat{H}})$.

Fix $\overline{h} \in H^\sigma$, $i \in I$, and $a_i \in A_i(\overline{h}) \backslash \widetilde{\sigma}_i^h[\overline{h}]$. If there exists $\widetilde{h} \preceq \overline{h}$ such that $\widetilde{h} \in \widehat{H}$, $\widehat{\sigma}^h =^* \widetilde{\sigma}^h =^{\widetilde{h}} \widetilde{\sigma}^{\widetilde{h}} =^* \sigma^{\widetilde{h}}$, so by ♠ $\widehat{\sigma} =^{\widetilde{h}} \sigma^{\widetilde{h}}$; then $\widehat{\sigma}^h|\overline{h}$ is an equilibrium, so by Lemma 5 ("only if") ★ holds. If $H_{a_i}^{\overline{h}} \backslash Z \subseteq \widehat{H}$, for every $\widehat{h} \in H_{a_i}^{\overline{h}} \backslash Z$, $\widehat{\sigma}_{-i}^h|\widehat{h} = \sigma_{-i}^{\widehat{h}}$ and by ♦, $\pi(\sigma_{-i}^{\widehat{h}}) = \pi(\widetilde{\sigma}_{-i}^h|\widehat{h})$; by $\widehat{\sigma}^h =^* \widetilde{\sigma}^h$, $\pi_i(\widehat{\sigma}^h|\overline{h}) = \pi_i(\widetilde{\sigma}^h|\overline{h})$ and by ♡, $(\widehat{\sigma}_{-i}^h|\overline{h})(S_{-i}^{\overline{h}}(\widehat{h})) = (\widetilde{\sigma}_{-i}^h|\overline{h})(S_{-i}^{\overline{h}}(\widehat{h}))$; so since $\widetilde{\sigma}^h|\overline{h}$ is an equilibrium by Lemma 5 ("only if") ★ holds. If $H_{a_i}^{\overline{h}} \cap \overline{D}_i \neq \emptyset$, fix $v := \min_{\widetilde{h} \in H_{a_i}^{\overline{h}} \cap \overline{D}_i} v(\widetilde{h})$ and $\widehat{h} := \arg\min_{\widetilde{h} \in H_{a_i}^{\overline{h}} \cap \overline{D}_i} v(\widetilde{h})$: for every $\widetilde{h} \in H_{a_i}^{\overline{h}} \cap \overline{D}_i$, $\widehat{\sigma}_{-i}^h|\widetilde{h} \leq \widetilde{\mu}_i^{\widetilde{h}}(\cdot|\widetilde{h}) \in \Delta(S_{-i,v}^h(\widetilde{h})|\widetilde{h}) \neq \emptyset$[24] and for every $\widetilde{h} \in H_{a_i}^{\overline{h}} \backslash (Z \cup \overline{D}_i)$, $\widehat{\sigma}_{-i}^h|\widetilde{h} \leq \widetilde{\sigma}^{\widetilde{h}} \in \Delta(S_{-i,v}^h(\widetilde{h})|\widetilde{h})$; hence by Lemma 6 ★ holds. Thus by Lemma 5 ("if") $\widehat{\sigma}^h$ is an equilibrium. ■

Now I can prove the fundamental Lemma.

Lemma 9 *Fix $h \in H$, $m \in \mathbb{N}$, an elimination procedure $(S_q^h)_{q \geq 0}$, a set of unordered histories $\widehat{H} = \{h^1, ..., h^w\} \subseteq H^h \backslash \{h\}$ and a set of SPE $\Sigma^{\widehat{H}} = (\sigma^{\widetilde{h}})_{\widetilde{h} \in \widehat{H}}$ s.t. A1 holds for $n = m$ and:*

- *A2 for every $v \leq w$, there exists an equilibrium $\widehat{\sigma}^{h,v} \in \Delta(S_{m-1}^h)$ such that $h^v \in \cup_{i \in I} D_i(\widehat{\sigma}^{h,v})$ and for every $q < v$, if $h^q \in H(\widehat{\sigma}^{h,v})$, $\widehat{\sigma}^{h,v} =^* \widetilde{\sigma}^{h^q}$;*

- *A3 for each $i \in I$, $n \leq m$, $\sigma^h \in E^h(\Sigma^{\widehat{H}})$ and $\mu_i^h \geq \sigma_{-i}^h$ t.s.b. $(S_{-i}^h)_{q=0}^{n-1}$, $\rho(\mu_i^h) \subseteq S_{i,n}^h$.*

Then there exist $\sigma^h \in E^h(\Sigma^{\widehat{H}})$ and an equilibrium $\widetilde{\sigma}^h \in \Delta(S_m^h)$ such that $\widetilde{\sigma}^h =^* \sigma^h$.

Proof. The proof is by induction on the depth of $\Gamma(h)$.

Inductive hypothesis (d)

The Lemma holds for every $h \in H$ such that $\Gamma(h)$ has depth not bigger than d.

Basis step (1) For every $i \in I$, $n \leq m$ and equilibrium of $\Gamma(h)$ σ^h such that Supp$\sigma^h \subseteq S_{n-1}^h$, by A3 $r(\sigma_{-i}^h) \subseteq S_{i,n}^h$. Inductively, Supp$\sigma^h \subseteq S_m^h$.

INDUCTIVE STEP (d+1) Suppose not. I will find a contradiction through a recursive procedure. Set $k = 0$ and $\overline{H}^0 := \widehat{H}$.

Recursive step (k)

[24] See the previous footnote.

If $k > 0$, $\overline{H}^k$ and $\Sigma^{\overline{H}^k}$ are defined in step $k-1$. Let $n \le m$ be the greatest $q \in \mathbb{N}$ such that there exist $\sigma^{h,k} \in E^h(\Sigma^{\overline{H}^k})$ and an equilibrium $\widetilde{\sigma}^{h,k} \in \Delta(S^h_{q-1})$ with $\widetilde{\sigma}^{h,k} =^* \sigma^{h,k}$. If $k > 0$, by the last remark of the previous steps, $E^h(\Sigma^{\overline{H}^k}) \subseteq ... \subseteq E^h(\Sigma^{\overline{H}^0})$. Then by ♦ A3 implies A0. Moreover, n weakly decreases with k. Then $\Sigma^{\overline{H}^k}$ satisfies A1 and A2 with n in place of m.[25] Lemma 7 yields $l \in I$ and $\widehat{h} \in D_l(\widetilde{\sigma}^{h,k}) \setminus \cup_{\overline{h} \in \overline{H}^k} H^{\overline{h}}$.

Define $((S^{\widehat{h}}_{i,q})_{i\in I})^\infty_{q=0}$ as $((S^h_{i,q}(\widehat{h})|\widehat{h}))_{i\in I})^\infty_{q=0}$. By 1, it is an elimination procedure. Fix $i \neq l$ and $v \le n$. For every $\widetilde{\mu}^{\widehat{h}}_i$ t.s.b. $(S^{\widehat{h}}_{-i,q})^{v-1}_{q=0}$, since $\widehat{h} \in D_l(\widetilde{\sigma}^{h,k})$, by Lemma 3 there exists $\widetilde{\mu}^h_i =^h \widetilde{\sigma}^{h,k}_{-i}$ t.s.b. $(S^h_{-i,q})^{v-1}_{q=0}$ such that $\widetilde{\mu}^h_i =^{\widehat{h}} \widetilde{\mu}^{\widehat{h}}_i$. By A0, $\rho(\widetilde{\mu}^h_i) \subseteq S^h_{i,v}$; by ♣, $\rho(\widetilde{\mu}^h_i)(\widehat{h}) \neq \emptyset$. Hence, together with Lemma 7, for every $i \in I$, $v \le n$, $(\widehat{\sigma}^{\widehat{h}}_j)_{j\neq i} \in \times_{j\neq i}\Delta(S^{\widehat{h}}_{i,n})$ and $\widetilde{\mu}^{\widehat{h}}_i \geq \widehat{\sigma}^{\widehat{h}}_{-i}$ t.s.b. $(S^{\widehat{h}}_{-i,q})^v_{q=0}$, $\rho(\widetilde{\mu}^{\widehat{h}}_i) \subseteq S^{\widehat{h}}_{i,v} \neq \emptyset$ (**F**).

Let $\widetilde{H}^0 := \overline{H}^k \cap H^{\widehat{h}} \not\ni \widehat{h}$: I show that there exist $\sigma^{\widehat{h}} \in E^{\widehat{h}}(\Sigma^{\widetilde{H}^0})$ and an equilibrium $\widetilde{\sigma}^{\widehat{h}} =^* \sigma^{\widehat{h}}$ such that $\mathrm{Supp}\widetilde{\sigma}^{\widehat{h}} \subseteq S^{\widehat{h}}_n$. Suppose not (**G**). I will find a contradiction through a recursive procedure. For every $q \le w + k$ such that $h^q \in \widetilde{H}^0$, let $\overline{h}^q := h^q$ and $\widehat{\sigma}^{\widehat{h},q} := \widehat{\sigma}^{h,q}|\widehat{h}$, which is an equilibrium because by $\widehat{h} \notin \cup_{\overline{h}\in\overline{H}^k} H^{\overline{h}}$, $\widehat{h} \prec \overline{h}^q$, so by $\overline{h}^q \in \cup_{i\in I} D_i(\widehat{\sigma}^{h,q})$, $\widehat{h} \in H(\widehat{\sigma}^{h,q})$. Set $t = 0$.

Recursive step (t) *If $t > 0$, $\widetilde{H}^t$, and $\Sigma^{\widetilde{H}^t}$ are defined in step $t-1$, and satisfy A1 and A2 with n in place of m and $\widehat{h}$ in place of h.*[26] *Let $\tau := w + k + t$. For every $i \in I$, let $\Sigma^{\widehat{h},t}_i$ be the set of $\widehat{\sigma}^{\widehat{h}}_i \in \Delta(S^{\widehat{h}}_{i,n})$ such that for every $\widetilde{h} \in \widetilde{H}^t \cap H(\widehat{\sigma}^{\widehat{h}}_i)$, $\widehat{\sigma}^{\widehat{h}}_i =^{\widetilde{h}} \widetilde{\sigma}^{\widetilde{h}}_i$. Note that for every $i \in I$,*

$$\Sigma^{\widehat{h},t}_i = \cap_{\widetilde{h}\in\widetilde{H}^t} \cap_{z\in Z^{\widetilde{h}}\cup H^{\widetilde{h}}} \{\widehat{\sigma}^{\widehat{h}}_i \in \Delta(S^{\widehat{h}}_i) : \widehat{\sigma}^{\widehat{h}}_i(S^{\widehat{h}}_i(z)) = \widehat{\sigma}^{\widehat{h}}_i(S^{\widehat{h}}_i(\widetilde{h})) \cdot \widetilde{\sigma}^{\widetilde{h}}_i(S^{\widetilde{h}}_i(z))\} \cap \Delta(S^{\widehat{h}}_{i,n}),$$

an intersection of convex and compact sets.[27] *Hence $\Sigma^{\widehat{h},t}_i$ is convex and compact. Then, since expected utility is linear, the reduced game with strategy sets $(\Sigma^{\widehat{h},t}_i)_{i\in I}$, if non-empty, features an equilibrium $\widetilde{\sigma}^{\widehat{h},t}$. For later reference, for each $i \in I$, fix $\mu^{\widehat{h}}_i =^{\widehat{h}} \widetilde{\sigma}^{\widehat{h},t}_{-i}$ t.s.b. $(S^{\widehat{h}}_{-i,q})^n_{q=0}$ and $\widehat{\sigma}^{\widehat{h},\tau+1}_i \in \Delta(\rho(\mu^{\widehat{h}}_i)) \subseteq \Delta(r(\widetilde{\sigma}^{\widehat{h},t}_{-i}))$ such that, by $\widetilde{\sigma}^{\widehat{h},t}_{-i} =^{\widetilde{h}} \widetilde{\sigma}^{\widetilde{h}}_{-i}$ and ♣, $\widehat{\sigma}^{\widehat{h},\tau+1}_i =^* \widetilde{\sigma}^{\widetilde{h}}_i$ for all $\widetilde{h} \in \widetilde{H}^t \cap H(\widetilde{\sigma}^{\widehat{h},t}_{-i})$; by F, $\widehat{\sigma}^{\widehat{h},\tau+1}_i \in \Delta(S^{\widehat{h}}_{i,n})$. I show that $\Sigma^{\widehat{h},t}_i$ is non-empty and that $\widetilde{\sigma}^{\widehat{h},t}$ is an equilibrium of the whole $\Gamma(\widehat{h})$.*

If $t = 0$ and $\widetilde{H}^0 = \emptyset$, $\Sigma^{\widehat{h},0}_i = \Delta(S^{\widehat{h}}_{i,n}) \neq \emptyset$ (by F) and $\widehat{\sigma}^{\widehat{h},\tau+1}_i \in \Sigma^{\widehat{h},0}_i$, so $\widetilde{\sigma}^{\widehat{h},t}_i \in \Delta(r(\widetilde{\sigma}^{\widehat{h},t}_{-i}))$ too. Else, for notational convenience let $\widehat{\sigma}^{\widehat{h},\tau+1}_{-i} := \widetilde{\sigma}^{\widehat{h},t}_{-i}$ and proceed as follows.

---

[25] For every $h^q \in \overline{H}^k$, $\widetilde{\sigma}^{h^q} \in \Delta(S^h_n(h^q)|h^q)$ and $\widehat{\sigma}^{h,q} \in \Delta(S^h_{n-1})$ come from A1 and A2 if $h^q \in \widehat{H}$, from some previous step if $h^q \notin \widehat{H}$.

[26] For every $\overline{h}^q \in \widetilde{H}^t$, $\widetilde{\sigma}^{\overline{h}^q} \in \Delta(S^h_n(\overline{h}^q)|\overline{h}^q)$ and $\widehat{\sigma}^{\widehat{h},q} \in \Delta(S^{\widehat{h}}_{n-1})$ come from the outer recursive step if $\overline{h}^q \in \overline{H}^k$, from some previous step if $\overline{h}^q \notin \overline{H}^k$.

[27] Clearly, $\Delta(S^{\widehat{h}}_{i,n})$ is convex and compact. Each set of the kind $\{\widehat{\sigma}^{\widehat{h}}_i \in \Delta(S^{\widehat{h}}_i) : (\widehat{\sigma}^{\widehat{h}}_i)(S^{\widehat{h}}_i(z)) = (\widehat{\sigma}^{\widehat{h}}_i)(S^{\widehat{h}}_i(\widetilde{h}))\cdot c\}$, where c is a constant, is clearly convex and compact too. Notice that if $\widetilde{h} \notin H(\widehat{\sigma}^{\widehat{h}}_i)$, $\widehat{\sigma}^{\widehat{h}}_i$ satisfies $(\widehat{\sigma}^{\widehat{h}}_i)(S^{\widehat{h}}_i(z)) = (\widehat{\sigma}^{\widehat{h}}_i)(S^{\widehat{h}}_i(\widetilde{h})) \cdot c$ as $0 = 0$.

For every $\widetilde{h} \in H^{\widehat{h}}$ and $q \leq \tau$, let $Q^{\widetilde{h}}_q := \{g \leq q : \overline{h}^g \in \widetilde{H}^t \cap H^{\widetilde{h}}\}$ and $Q^{\widetilde{h}}_{\tau+1} := Q^{\widetilde{h}}_{\tau}$. I show that for every $q \in Q^{\widehat{h}}_{\tau} \cup \{\tau+1\}$, there exists $\overline{\sigma}^{\widehat{h},q}_i \in \Delta(S^{\widehat{h}}_{i,n})$ such that $\overline{\sigma}^{\widehat{h},q}_i =^* \widehat{\sigma}^{\widehat{h},q}_i$ and for every $g \in Q^{\widehat{h}}_q$, $\overline{\sigma}^{\widehat{h},q}_i =^{\overline{h}^g} \widetilde{\sigma}^{\overline{h}^g}_i$. Fix $q \in Q^{\widehat{h}}_{\tau} \cup \{\tau+1\}$ and suppose to have shown it already for every $g \in Q^{\widehat{h}}_q \backslash \{q\}$. Since $\widehat{\sigma}^{\widehat{h},q}_i \in \Delta(r(\widehat{\sigma}^{\widehat{h},q}_{-i}))$, by $\widehat{\sigma}^{\widehat{h},q} \in \Delta(S^{\widehat{h}}_n)$ and F*, Lemma 4 yields $\overline{\sigma}^{\widehat{h},q}_i \in \Delta(S^{\widehat{h}}_{i,n})$ such that $\overline{\sigma}^{\widehat{h},q}_i =^* \widehat{\sigma}^{\widehat{h},q}_i$ and for every $\widetilde{h} \in D^{-i}(\widehat{\sigma}^{h,q})$:*

- *if $\widetilde{h} \in \widetilde{H}^t$, $\overline{\sigma}^{\widehat{h},q}_i|\widetilde{h} = \widetilde{\sigma}^{\widetilde{h}}_i$;*

- *if $\widetilde{h} \notin \widetilde{H}^t$ but $Q^{\widetilde{h}}_q \neq \emptyset$, $\overline{\sigma}^{\widehat{h},q}_i|\widetilde{h} = \overline{\sigma}^{\widehat{h},\max Q^{\widetilde{h}}_q}_i|\widetilde{h}$ (where $\max Q^{\widetilde{h}}_q < q$ because $\overline{h}^q \in \cup_{j\in I} D_j(\widehat{\sigma}^{h,q}) \cap \widetilde{H}^t$ and $\cup_{j\in I} D_j(\widehat{\sigma}^{h,q}) \cap H^{\widetilde{h}} = \{\widetilde{h}\} \nsubseteq \widetilde{H}^t$);*

- *if $\widetilde{h} \succ \overline{h}$ for some $\overline{h} \in \widetilde{H}^t$, $\overline{\sigma}^{\widehat{h},q}_i|\widetilde{h} = \widetilde{\sigma}^{\overline{h}}_i|\widetilde{h}$, so that by $\overline{\sigma}^{\widehat{h},q}_i =^* \widehat{\sigma}^{\widehat{h},q}_i =^* \widetilde{\sigma}^{\overline{h}}_i$ and ♠, $\overline{\sigma}^{\widehat{h},q}_i =^{\overline{h}} \widetilde{\sigma}^{\overline{h}}_i$.*

Then $\overline{\sigma}^{\widehat{h},\tau}_i \in \Sigma^{\widehat{h},t}_i \neq \emptyset$.[28] *So,$\widetilde{\sigma}^{\widehat{h},t}$ and $\widehat{\sigma}^{\widehat{h},\tau+1}$ exist and $\overline{\sigma}^{\widehat{h},\tau+1}_i \in \Sigma^{\widehat{h},t}_i$. Since $\overline{\sigma}^{\widehat{h},\tau+1}_i =^* \widehat{\sigma}^{\widehat{h},\tau+1}_i$, $\overline{\sigma}^{\widehat{h},\tau+1}_i \in \Delta(r(\widetilde{\sigma}^{\widehat{h},t}_{-i}))$. Then $\widetilde{\sigma}^{\widehat{h},t}_i \in \Delta(r(\widetilde{\sigma}^{\widehat{h},t}_{-i}))$ too.*

By ♦, F *implies A0 with $\widehat{h}$ in place of h; $\widetilde{H}^t$ satifies A1 with n in place of m. By the last remark of the previous steps $E^h(\Sigma^{\widetilde{H}^t}) \subseteq E^h(\Sigma^{\widetilde{H}^0})$, so, by* G*, $\widetilde{\sigma}^{\widehat{h},t}$ satisfies the hypotheses of Lemma 8.*[29] *Lemma 8 yields $p \in I$ and $\overline{h} \in D_p(\widetilde{\sigma}^{\widehat{h},t}) \backslash \cup_{\widetilde{h}\in\widetilde{H}^t} H^{\widetilde{h}}$.*

Define the elimination procedure $((S^{\overline{h}}_{i,q})_{i\in I})^{\infty}_{q=0} := ((S^{\widehat{h}}_{i,q}(\overline{h})|\overline{h}))_{i\in I})^{\infty}_{q=0}$. Fix $i \neq p$ and $v \leq n$. For every $\widetilde{\mu}^{\overline{h}}_i$ t.s.b. $(S^{\overline{h}}_{-i,q})^{v}_{q=0}$, since $\overline{h} \in D_p(\widetilde{\sigma}^{\widehat{h},t})$, by Lemma 3 there exists $\widetilde{\mu}^{\widehat{h}}_i =^{\widehat{h}} \widetilde{\sigma}^{\widehat{h},t}_{-i}$ t.s.b. $(S^{\widehat{h}}_{-i,q})^{v-1}_{q=0}$ such that $\widetilde{\mu}^{\widehat{h}}_i =^{\overline{h}} \widetilde{\mu}^{\overline{h}}_i$. By A0, $\rho(\widetilde{\mu}^{\widehat{h}}_i) \subseteq S^{\widehat{h}}_{i,v}$; by ♣, $\rho(\widetilde{\mu}^{\widehat{h}}_i)(\overline{h}) \neq \emptyset$. Hence, together with Lemma 8, for every $i \in I$, $v \leq n$, $\sigma^{\overline{h}} \in E^{\overline{h}}(\Sigma^{\widetilde{H}^t})$ and $\widetilde{\mu}^{\overline{h}}_i \geq \sigma^{\overline{h}}_{-i}$ t.s.b. $(S^{\overline{h}}_{-i,q})^{v}_{q=0}$, $\rho(\mu^{\overline{h}}_i) \subseteq S^{\overline{h}}_{i,v}$. For every $q \in Q^{\overline{h}}_{\tau}$, let $\widehat{\sigma}^{\overline{h},q} := \widehat{\sigma}^{\widehat{h},q}|\overline{h}$, which is an equilibrium because by $\overline{h} \notin \cup_{\widetilde{h}\in\widetilde{H}^t} H^{\widetilde{h}}$, $\overline{h} \prec \overline{h}^q$, so by *$\overline{h}^q \in \cup_{i\in I} D_i(\widehat{\sigma}^{\widehat{h},q})$, $\overline{h} \in H(\widehat{\sigma}^{\widehat{h},q})$. Then A3, A2, and A1 are satisfied with $\widetilde{H}^t \cap H^{\overline{h}}$ in place of $\widehat{H}$, n in place of m, and $\overline{h}$ in place of h. So, by the Induction Hypothesis, there exist $\sigma^{\overline{h}} \in E^{\overline{h}}(\Sigma^{\widetilde{H}^t})$ and an equilibrium $\widetilde{\sigma}^{\overline{h}} =^* \sigma^{\overline{h}}$ such that* $\mathrm{Supp}\widetilde{\sigma}^{\overline{h}} \subseteq S^{\overline{h}}_n = S^{\widehat{h}}_n(\overline{h})|\overline{h}$.

Since $\overline{h} \notin \cup_{\widetilde{h}\in\widetilde{H}^t} H^{\widetilde{h}}$ and $H^{\widehat{h}}$ is finite, $\widetilde{H}^{t+1} := \{\widetilde{h} \in \widetilde{H}^t : \widetilde{h} \nsucceq \overline{h}\} \cup \{\overline{h}\}$ is a set of unordered histories that keep shortening with t, until a contradiction is obtained. Before, let $\Sigma^{\widetilde{H}^{t+1}} := \Sigma^{\widetilde{H}^t} \cup \{\sigma^{\overline{h}}\} \backslash \{\sigma^{\widetilde{h}}\}_{\widetilde{h}\in\widetilde{H}^t\backslash\widetilde{H}^{t+1}}$, $\overline{h}^{\tau+1} := \overline{h}$, $\widehat{\sigma}^{\widehat{h},\tau+1} := \widetilde{\sigma}^{\widehat{h},t}$.[30] *Then, increase t by* 1 *and run again noting what follows: for every $\widetilde{h} \in \widetilde{H}^t$ such that $\widetilde{h} \succ \overline{h}$, $\sigma^{\overline{h}}|\widetilde{h} = \sigma^{\widetilde{h}}$, so that $E^h(\Sigma^{\widetilde{H}^{t+1}}) \subseteq E^h(\Sigma^{\widetilde{H}^t})$.*

Since $\widehat{h} \notin \cup_{\overline{h}\in\overline{H}^k} H^{\overline{h}}$ and H^h is finite, $\overline{H}^{k+1} := \{\widetilde{h} \in \overline{H}^k : \widetilde{h} \nsucceq \widehat{h}\} \cup \{\widehat{h}\}$ is a set of unordered histories that keep shortening with k, until a contradiction is obtained. Before,

[28] By recursive step $t-1$, $\max Q^{\widehat{h}}_{\tau} = \tau$.

[29] Without loss of generality assume that for every $i \in I$ and $\widetilde{h} \in \widetilde{H}^t \backslash H(\widetilde{\sigma}^{\widehat{h},t}_i)$, $\widetilde{\sigma}^{\widehat{h},t}_i =^{\widetilde{h}} \widetilde{\sigma}^{\widetilde{h}}_i$.

[30] Overwrite the previous, temporary choice.

let $\Sigma^{\overline{H}^{k+1}} := \Sigma^{\overline{H}^{k}} \cup \{\sigma^{\widehat{h}}\} \backslash \{\sigma^{\overline{h}}\}_{\overline{h}\in\overline{H}^{k}\backslash\overline{H}^{k+1}}$, $h^{w+k+1} := \widehat{h}$, $\widehat{\sigma}^{h,w+k+1} := \widetilde{\sigma}^{h,k}$.[31] Increase k by 1 and run again noting what follows: for every $\overline{h} \in \overline{H}^{k}$ such that $\overline{h} \succ \widehat{h}$, $\sigma^{\widehat{h}}|\overline{h} = \sigma^{\overline{h}}$, so that $E^{h}(\Sigma^{\overline{H}^{k+1}}) \subseteq E^{h}(\Sigma^{\overline{H}^{k}})$. ■

Proof of Lemma 2. Apply Lemma 9 with empty $\widehat{H}$ and $\Sigma^{\widehat{H}}$, and $m = \infty$. ■

References

[1] Battigalli, P., "On rationalizability in extensive games", *Journal of Economic Theory*, **74**, 1997, 40-61.

[2] Battigalli, P., "Dynamic Consistency and Imperfect Recall", *Games and Economic Behavior*, **20(1)**, 1997, 31-50.

[3] Battigalli, P., "Rationalizability in Infinite, Dynamic Games of Incomplete Information," *Research in Economics,* **57**, 2003, 1-38.

[4] Battigalli, P. and M. Siniscalchi, "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory*, **106**, 2002, 356-391.

[5] Battigalli P. and M. Siniscalchi, "Rationalization and Incomplete Information," *The B.E. Journal of Theoretical Economics*, **3(1)**, 2003, 1-46.

[6] Catonini, E., "On non-monotonic strategic reasoning", working paper, 2017.

[7] Catonini, E., "Rationalizability and Epistemic Priority Orderings", working paper, 2017.

[8] Catonini, E., "Self-Enforcing Agreements and Forward Induction Reasoning", working paper, 2017.

[9] Chen, J., and S. Micali, "The order independence of iterated dominance in extensive games", *Theoretical Economics*, **8**, 2013, 125-163.

[10] Govindan, S., and R. Wilson, "On forward induction," *Econometrica*, **77**, 2009, 1-28.

[11] Heifetz, A., and A. Perea, "On the Outcome Equivalence of Backward Induction and Extensive Form Rationalizability", *International Journal of Game Theory*, **44**, 2015, 37–59.

[12] Kohlberg, E. and J.F. Mertens, "On the Strategic Stability of Equilibria", *Econometrica*, **54**, 1986, 1003-1038.

[31] Note that for each $\overline{h} \in \overline{H}^{k} \cap H(\widetilde{\sigma}^{h,k})$, $\widetilde{\sigma}^{h,k} =^{*} \sigma^{h,k} =^{*} \sigma^{\overline{h}} =^{*} \widetilde{\sigma}^{\overline{h}}$, so $\widetilde{\sigma}^{h,k}$ satisfies A2 with $\overline{H}^{k+1}$.

[13] Kreps, D.M. and R. Wilson, "Sequential equilibria", *Econometrica*, **50**, 1982, 863-94.

[14] Man, P. "Forward Induction Equilibrium", *Games and Economic Behavior*, **75(1)**, 2012, 265-276.

[15] Osborne, M. J. and A. Rubinstein, "A Course in Game Theory", 1994, Cambridge, Mass.: MIT Press.

[16] Pearce, D., "Rational Strategic Behavior and the Problem of Perfection", *Econometrica*, **52**, 1984, 1029-1050.

[17] Renyi, A., "On a New Axiomatic Theory of Probability", *Acta Mathematica Academiae Scientiarum Hungaricae*, **6**, 1955, 285-335.