The Epistemic Spirit of Divinity*

Pierpaolo Battigalli,† Emiliano Catonini‡

February 15, 2022

Abstract: We study strategic reasoning in a signaling game where players have common belief in an outcome distribution and in the event that the receiver believes that the sender's first-order beliefs are independent of her payoff-type. We characterize the behavioral implications of these epistemic hypotheses through a rationalizability procedure with second-order belief restrictions. Our solution concept is related to, but weaker than Divine Equilibrium (Banks and Sobel, 1987). First, we do not obtain sequential equilibrium, but just Perfect Bayesian Equilibrium with heterogeneous off-path beliefs (Fudenberg and He, 2018). Second, when we model how the receiver may rationalize a particular deviation, we take into account that some types could have preferred a different deviation, and we show this is natural and relevant via an economic example.

1 Introduction

We investigate the strategic interaction between an informed first mover (sender, *she*) and an uninformed second mover (receiver, *he*) under the following hypotheses. At the beginning of such signalling game, the sender and the receiver have a belief about each other's moves that is consistent with the same distribution over terminal nodes, i.e., type-message-action triples. Furthermore, the receiver believes that the sender's beliefs are independent of her type. There is initial common belief that players are rational (i.e., subjective expected utility maximizers), and that their beliefs satisfy the properties above. For brevity, we call "on-path (off-path) messages" those that have positive (zero) marginal probability according to the given distribution. Thus, the receiver initially initially assigns probability 0 to off-path messages. Of course, whether the sender, given her type, wants to choose on-path messages depends on her conjecture about the receiver's reaction to off-path messages, which should be compared to the expected reaction to on-path messages specified by the given outcome distribution. With this, we also assume that, when the receiver observes an off-path message, he tries to interpret it under the view that the sender's beliefs satisfy the aforementioned properties. This means that the receiver tries to rationalize the deviation with a theory about the sender that is consistent with his prior on types, such that every

*We thank Nicodemo De Vito and Julien Manili for helpful suggestions.

†Università Bocconi, pierpaolo.battigalli@unibocconi.it

‡New York University in Shanghai, Department of Economics, emiliano.catonini@gmail.com

type has the same belief about the receiver, and this belief is consistent with the given outcome distribution, but induces at least one type to deviate.[1]

The existence of a commonly expected outcome distribution can arise, for instance, from the observation of long-run data in a situation of recurrent interaction. The receiver's hypothesis that the sender's beliefs are independent of her type can be motivated in two ways. The first is that there is an ex-ante stage of the game where the sender forms beliefs about the receiver before learning her type and these beliefs do not change at the interim stage. However, this interpretation cannot apply to a context where the sender already knows her type before facing the game, as, for instance, in the context of a population "meta-game" where (heterogeneous) senders and receivers are randomly matched in each period. The second way is a principle of insufficient reason for the receiver: If there is no clear direction in which the sender's type would influence her beliefs, it seems natural to reason about the sender's beliefs independently of her type.

To characterize the behavioral implications of these hypotheses, we employ a variant of Strong-Δ-Rationalizability (Battigalli 2003, Battigalli and Siniscalchi 2003), a notion of rationalizability with belief restrictions for sequential games, which has well-understood epistemic foundations capturing forward-induction reasoning (Battigalli and Siniscalchi 2007, Battigalli and Prestipino 2013). While the baseline notion of Strong-Δ-Rationalizability only features restrictions on first-order beliefs, we put restrictions on the receiver's second-order beliefs to represent the independence hypothesis.

The seminal paper of Kohlberg and Mertens (1986) introduced in equilibrium analysis the idea of "forward induction". They show that their notion of strategic stability refines off-path beliefs with the view that the deviator is trying to move on a preferred path with respect to the underlying equilibrium. Following Kohlberg and Mertens (1986), other equilibrium refinements were introduced to capture the forward induction implications of strategic stability in a simpler and clearer way. The Intuitive Criterion (Cho and Kreps, 1987) and Divine Equilibrium (Banks and Sobel, 1987) address this issue for signaling games, which combine a simple structure with widespread relevance. On top of the rationalization of deviations in light of the equilibrium outcome distribution, which is a feature of both refinements, Divine Equilibrium is inspired also by the idea that the sender's beliefs are independent of her type. We aim to capture the spirit of Divine Equilibrium through the approach of epistemic game theory. So, we formulate primitive hypotheses on players' hierarchies of beliefs and we calculate their behavioral implications with an iterated elimination procedure. We find the following: Every outcome distribution of a Divine Equilibrium is consistent with our hypotheses, but not the other way round. There are two reasons for this. The first reason is that, differently from Divine Equilibrium, we do not require the sender to assign positive probability only to off-path reactions of the receiver that best respond to the same updated belief about types. As a consequence, while Divine Equilibrium refines sequential equilibrium (Kreps and Wilson, 1982), the outcome distributions that are consistent with our hypotheses are induced by some Perfect Bayesian Equilibrium with

[1] The receiver continues to be certain that every type has the same belief about him, but he may well be uncertain about what such belief is.

(possibly) heterogeneous off-path beliefs (Fudenberg and He, 2018). Requiring the sender to be certain of the receiver's belief after an unexpected message is in line with the spirit of sequential equilibrium and trembling-hand perfection.[2] However, the idea that the sender has purposedly deviated *despite* her belief in the receiver's equilibrium on-path behavior implies that the sender does not believe in the equilibrium reaction to the deviation. This destroys the possible source of certainty about off-path beliefs.[3] The second reason why we cannot rule out every non-divine equilibrium outcome is that Divine Equilibrium refines beliefs after each off-path message without taking into account the (possible) existence of other off-path messages. This induces the receiver to raise the relative probability of type θ over type θ' whenever θ finds that particular deviation profitable for a larger set of beliefs than θ'. But θ may find another deviation even more profitable, and choose it under a belief that induces θ' to stick to the first deviation. As we will show in Section 2 through an example, this difference has relevant and intuitive implications in meaningful games.

Our analysis is closely related to that of Sobel, Stole and Zapater (1990). Sobel et al. consider the complete-information game where the sender forms a belief at the ex-ante stage, before observing the chance move that determines her "type". Then, given a sequential equilibrium of the signaling game, they substitute the equilibrium messages with one message m^* that directly yields the (type-dependent) equilibrium expected payoff. Finally, they apply to the modified game a version of extensive-form rationalizability assuming type-independence of the sender's conjecture about the receiver (unlike the original version due to Pearce 1984). With this, they obtain a "Fixed-Equilibrium Rationalizable Outcome" (FERO) of the original game if the pseudo-message m^* survives such iterated elimination procedure. In the appendix, we consider a complete-information scenario and show that the implications of FERO (extended to non-sequential equilibria) are weaker than those of our hypotheses. The reason is that FERO allows the receiver to change his theory of the sender after different deviations in incoherent ways: for instance, the receiver can believe after an off-path message m that some type θ would have sent a different off-path message m', and after m' that θ would have sent m, so he uses different theories after m and m' although both theories are able to explain both messages (see the appendix for an example). We adopt a notion of "belief consistency" that avoids this. In particular, we model the receiver's first-order beliefs as *complete conditional probability systems* (henceforth, CCPSs): one conditional belief for every nonempty subset of the relevant space of uncertainty.[4] This has two advantages. First, it provides the language to formulate and restrict the theories about the sender that the receiver uses to rationalize deviations, i.e., from which he derives his beliefs conditional on the received messages.[5] Second, a CCPS induces a *completely consistent belief system* (Battigalli et al., 2021, Siniscalchi, 2020) over information sets; as shown by Battigalli et al. (2021) and Catonini (2022), complete consistency translates into natural properties for an agent's beliefs, which can be expressed in terms of coherence

[2] Kreps and Wilson (1982) show that sequential equilibrium is obtained by simultaneously imposing robustness to trembling-like perturbations as well as perturbations of payoffs at terminal nodes.

[3] See Catonini (2021) for a related criticism of subgame perfection.

[4] Conditional probability systems were introduced by Renyi (1995) for arbitrary collections of conditioning events. Myerson (1986) introduced *complete* CPSs in the analysis of finite games.

[5] The notion of CPS defined on the collection of observable events (corresponding to information sets, i.e., the root and the messages) feature only the latter beliefs.

between different odds ratios at different information sets, and interpreted as a matter of introspection or a "wired-in" process of belief formation.[6]

Our analysis is also related to Battigalli and Siniscalchi (2003), who consider a commonly believed outcome distribution, but do not make the independence hypothesis. With this, they show that non-emptiness of the resulting version of Strong-Δ-Rationalizability: (i) is equivalent to passing the Iterated Intuitive Criterion; (ii) implies that the outcome distribution is induced by a self-confirming equilibrium (Fudenberg and Levine, 1993). Cho (1987) strengthens the Intuitive Criterion by requiring that *one* distribution over the reactions of the receiver that are consistent with the criterion induces *all* types of the sender to stay on path, coherently with the spirit of Nash equilibrium. We directly obtain Nash equilibrium because of the independence hypothesis: The receiver must be able to believe that all types stay on-path for the same first-order belief.

The paper is structured as follows. In Section 2 we illustrate similarities and differences between our approach and Divine Equilibrium through an example. In Section 3 we formalize "path rationalizability with second-order independence", which characterizes in a transparent way the behavioral implications of our epistemic hypotheses — a less transparent but operationally simpler equivalent procedure is provided at the end of the section. In Section 4 we formalize the relation between our solution concept and Divine Equilibrium, and the solution of the example of Section 2. In the Appendix, we analyze the complete-information scenario and the relation with FERO, and we collect the proofs that are omitted from the main body of the paper.

2 Job market example

A potential employee can be of two types, good (θ^h) and bad (θ^ℓ). She can stop studying after graduating from the BSc (m_1), or she can continue to an MSc (m_2) or to a PhD program (m_3). The employer can hire the employee in three different positions, a_1, a_2, a_3, with increasing responsibilities and salaries. The employer prefers to hire a good employee in the position she is best qualified for: a BSc graduate in position a_1, a MSc graduate in position a_2, and a PhD graduate in position a_3. The reason is that there is a productivity boost from education to a good employee, but overqualification does not carry any additional benefit. There is no productivity boost from education to a bad employee, so the employer always prefers to hire her in position a_1. (Any additional benefit of hiring a good type rather than a bad type independently of education is immaterial for the analysis.) Education has an increasing cost, which is higher for the bad type, but worth paying for both types if the employee does not end up overqualified for the position. The following table summarizes

[6] It is worth noting that the classical notions of structural consistency (Kreps and Wilson, 1982), adopted by FERO, and of CPS on the collection of observable events do not induce this coherence. Catonini (2022) shows that this lack of coherence exposes an agent to the possibility of being Dutch-booked in objective expected terms, across a collection of counterfactual contingencies.

players' payoffs — the first entry in each box is the payoff of the employee.

m_1	a_1	a_2	a_3
θ^h	$0,3$	$4,2$	$9,0$
θ^ℓ	$0,3$	$4,2$	$9,0$

m_2	a_1	a_2	a_3
θ^h	$-2,3$	$2,5$	$7,3$
θ^ℓ	$-3,3$	$1,2$	$6,0$

m_3	a_1	a_2	a_3
θ^h	$-5,3$	$-1,5$	$4,6$
θ^ℓ	$-8,3$	$-4,2$	$1,0$

Suppose that the two types are a priori equally likely, and this is commonly believed. Consider the following pooling equilibrium, where getting an MSc or a PhD leads to overqualification: both types choose m_1, and the employer chooses a_1 after m_1 and m_2, and a_2 after m_3. This equilibrium is not divine. The reason is the following. Under the belief that m_1 will lead to position a_1, all the beliefs that induce θ^ℓ to prefer m_2 to m_1 (also after eliminating the dominated response a_3) also induce θ^h to strictly prefer m_2 to m_1. Hence, according to divinity, after m_2 the employer must assign to θ^h at least the prior probability, and for every belief where θ^h is not less likely than θ^ℓ, the optimal response is a_2.

However, the pooling equilibrium is consistent with our epistemic hypotheses. Suppose the employer interprets a deviation to m_2 or m_3 with the following theory: the employee expects to get the position she is qualified for, that is, position a_i after m_i for each $i = 1, 2, 3$. Given this belief, θ^h strictly prefers m_3, while θ^ℓ is indifferent between m_2 and m_3.[7] Then, whenever the employer believes that θ^ℓ picks m_2 with positive probability, he must assign probability one to θ^ℓ after m_2, and this justifies a_1. Moreover, if the employer believes that θ^ℓ breaks her tie at random, she must assign probability 1/3 to θ^ℓ after m_3 and this justifies a_2.

3 Main analysis

Primitives of the game We consider the following signaling game. There is a payoff-relevant parameter θ, and it is common knowledge that θ belongs to a finite set Θ. The sender ($i = 1$) knows the true value of θ (henceforth, the sender's "type"), and chooses a message m from the finite set M (we assume that M does not depend on θ). The receiver ($i = 2$), who does not know the true value of θ, observes m and then chooses an action a from the finite set A (we assume that A does not depend on m).[8] The payoffs of sender and receiver are given by

$$u_i : \Theta \times M \times A \to \mathbb{R}, \quad i = 1, 2.$$

Derived objects We let A^M denote the set of strategies of the receiver, i.e., maps from M to A, and we let M^Θ denote the set of choice functions of the sender, .i.e., maps from Θ to M. In the main body, we will use the choice functions of the sender only to formulate conditional statements about the sender's behavior in the mind of the receiver. Given that we do not posit an ex-ante stage where the sender can reason about the game before learning her type, such choice functions shall not be interpreted as plans of the sender. With this, it

[7] The tie between the payoffs of θ^ℓ after (m_2, a_2) and (m_3, a_3) is only for simplicity: the employer could give positive probability to two different beliefs of the employee that induce θ^ℓ to choose, respectively, m_2 and m_3.

[8] We assume that every type of the sender has the same set of available messages, and that the receiver has the same set of available actions after every message only to simplify notation.

is useful to introduce notation for the choice functions of the sender where a specific type θ chooses a given message m and the set of choice functions allowing for message m:

$$M^{\Theta}(\theta, m) = \left\{s_1 \in M^{\Theta} : s_1(\theta) = m\right\},$$

$$M^{\Theta}(m) = \left\{s_1 \in M^{\Theta} : \exists\theta \in \Theta, s_1(\theta) = m\right\}.$$

Beliefs The primitive space of uncertainty for the sender is A^M, the set of strategies of the receiver. Since the sender only makes one choice at the beginning of the game, we can model her first-order beliefs with just one probability measure $\mu_1 \in \Delta(A^M)$.

The primitive space of uncertainty for the receiver consists of the sender's type and choice functions: $\Theta \times M^{\Theta}$.[9] The receiver moves after observing the message, but we require him to derive his revised belief from a more general theory about the sender, where different types may send different messages. Therefore, we represent his first-order beliefs as a complete Conditional Probability System over $\Theta \times M^{\Theta}$. Actually, since we also want to restrict his second-order beliefs in compliance with the independence hypothesis, for each $C \in 2^{\Theta \times M^{\Theta}} \setminus \{\emptyset\}$, we directly introduce a second-order belief conditional on the event $C \times \Delta(A^M)$. Thus, we model the receiver's *system of second-order beliefs* as an array $\mu_2 = (\mu_2(\cdot|C))_{C \in 2^{\Theta \times M^{\Theta}} \setminus \{\emptyset\}}$ of probability measures over $\Theta \times M^{\Theta} \times \Delta(A^M)$ that satisfies the following two conditions:[10]

1. for every $C \in 2^{\Theta \times M^{\Theta}} \setminus \{\emptyset\}$, $\mu_2\left(C \times \Delta(A^M)|C\right) = 1$;

2. (chain rule) for every $C, D \in 2^{\Theta \times M^{\Theta}}$, if $D \subseteq C$, then

$$\forall E \subseteq D \times \Delta(A^M), \quad \mu_2(E|C) = \mu_2(E|D) \cdot \mu_2\left(D \times \Delta(A^M)|C\right).$$

Belief restrictions Suppose there exists a commonly believed distribution over terminal nodes $\mu \in \Delta(\Theta \times M \times A)$ with strictly positive marginal (prior) belief over types $p = \text{marg}_{\Theta}\mu \in \Delta^{\circ}(\Theta)$. For each type $\theta \in \Theta$, we let $\nu^{\theta} = \text{marg}_M \mu(\cdot|\theta) \in \Delta(M)$ denote the probability over messages conditional on θ, and we let

$$\begin{aligned} M^*(\theta) &= \left\{m \in M : \nu^{\theta}(m) > 0\right\}, \\ M^* &= \text{suppmarg}_M \mu = \cup_{\theta \in \Theta} M^*(\theta) \end{aligned}$$

respectively denote the set of messages sent with positive probability by type θ and the set of messages sent with positive probability. For each message $m \in M^*$, we let $\nu^m = \text{marg}_A \mu(\cdot|m) \in \Delta(A)$ denote the probability over actions of the receiver conditional on m. We assume that μ factorizes as $\mu(\theta, m, a) = p(\theta)\, \nu^{\theta}(m)\, \nu^m(a)$: as customary, the actions of the receiver are not correlated with the (unobserved) sender's type. Furthermore, to

[9] One may wonder why we do not take just $\Theta \times M$ as the receiver's relevant uncertainty space. The reason is that the receiver will rationalize a message m with a theory of the sender where *some* types choose m and other types choose different messages. With this, the corresponding conditioning event cannot be expressed in the $\Theta \times M$ space.

[10] With a slight abuse of terminology and notation, we call belief conditional on C, and write $\mu_2(\cdot|C)$, the receiver's belief conditional on event $C \times \Delta(A^M)$.

avoid uninteresting cases, we also assume that each ν^m assigns positive probability only to optimal actions of the receiver, given his posterior belief about the sender's type.

Technically, $(\nu^\theta)_{\theta\in\Theta} \in \Delta(M)^\Theta$ is a behavior strategy of the sender, but we interpret it as a conjecture of the receiver about the sender. Indeed, (i) we are silent about existence of an ex ante stage at which the still ignorant sender supposedly plans, (ii) and we do not assume that the sender randomizes when indifferent. Similarly, $(\nu^m)_{m\in M^*}$ is interpreted as part of a conjecture of the sender about the receiver, not as a partial behavior strategy of the receiver.[11]

Given this interpretation, we consider the restricted set of sender's first-order beliefs consistent with the commonly believed distribution over terminal nodes:

$$\Delta_1 = \left\{\mu_1 \in \Delta(A^M) : \forall m \in M^*, \forall a \in A, \mu_1\left(\left\{s_2 \in A^M : s_2(m) = a\right\}\right) = \nu^m(a)\right\}.$$

Note that, in principle, the sender can have correlated beliefs about the actions of the receiver after different messages, but these correlations are immaterial for her choice problem.[12]

To restrict the beliefs of the receiver, we start by letting P_2 denote the set of all *finite-support* probability measures ν over $\Theta \times M^\Theta \times \Delta(A^M)$ that conform to the prior on types and feature no correlation between types and choice function-belief pairs:

B1 for every $\theta \in \Theta$, $\nu(\{\theta\} \times M^\Theta \times \Delta(A^M)) = p(\theta)$;

B2 for every $(\theta, s_1, \mu_1) \in \Theta \times M^\Theta \times \Delta(A^M)$,

$$\nu(\theta, s_1, \mu_1) = \nu(\{\theta\} \times M^\Theta \times \Delta(A^M)) \cdot \nu(\Theta \times \{(s_1, \mu_1)\}).$$

We impose independence between types and choice function-belief pairs, not just sender's beliefs, because the belief about types and the belief about how types determine messages (that is, about the choice function) are naturally uncorrelated, as long as one conditions on a Cartesian event. In particular, conditional on an event $\Theta \times E \times \Delta(A^M)$ with $E \subseteq M^\Theta$, the receiver's belief will satisfy independence — of course, correlations will appear conditional on the observation of a message m, i.e., on the non-Cartesian event $\cup_{\theta\in\Theta}\{\theta\} \times M^\Theta(\theta, m)$ (different types may need to be paired with different choice functions to induce m).[13] Conditions B1 and B2 boil down to saying that P_2 is the set of finite-support product measures $p \times \eta$ (with $\eta \in \Delta(M^\Theta \times \Delta(A^M))$.

Let $P_2^* \subseteq P_2$ denote the set of all $\nu \in P_2$ that are consistent with the commonly believed distribution on terminal nodes, i.e., that satisfy the following additional condition:[14]

[11] "Partial", because M^* may be a strict subset of M.

[12] This is not the case if the sender also has the incentive to learn about the receiver's reactions, in a repeated interaction setting. We abstract away from these learning incentives.

[13] A paradoxical consequence of allowing for correlated beliefs over types and choice functions is that, even conditioning on $\Theta \times M^\Theta(m)$ for some message m, and even deeming every type possible, the belief could still give probability zero to type-functions pairs where the message assigned by the function to the type is m. This would imply that the theory of the sender conditional on $\Theta \times M^\Theta(m)$ fails to rationalize m.

It is worth noting that independence between types and choice functions also follows from independence between types and first-order beliefs whenever each type best replies to the first-order belief, and no first-order belief that leaves some type indifferent between different messages is assigned positive probability.

[14] Recall, all beliefs in P_2 agree with the prior p on types.

B3 for every $(\theta, m) \in \Theta \times M$,

$$\nu\left(\Theta \times M^{\Theta}(\theta, m) \times \Delta(A^M)\right) = \nu^{\theta}(m).$$

Since the sender has finite sets of types and messages, the focus on finite-support beliefs is without loss of generality for the justifiable behaviors of the receiver, for the following reason: from every $\nu = p \times \eta$ we can derive a finite-support probability measure $\nu' = p \times \eta'$ with the same marginal over type-function pairs by associating each choice function s_1 in the marginal support of η with just one belief μ_1 such that $(s_1, \mu_1) \in \text{supp}\eta$.[15]

With this, we restrict the beliefs of the receiver so that they agree with the prior on types, with the expected behavior of the sender, and with the independence hypothesis; conditional on contemplating some unexpected messages, we still require the beliefs to agree with the prior on types and with the independence hypothesis. Thus, we consider systems of second-order beliefs μ_2 such that:

D1 $\mu_2(\cdot|\Theta \times M^{\Theta}) \in P_2^*$;

D2 for each $E \subset 2^{M^{\Theta}}$, $\mu_2(\cdot|\Theta \times E) \in P_2$.[16]

We let Δ_2 denote the set of systems of second-order beliefs that satisfy D1 and D2. To ease notation, we let $\mu_2(\cdot|\varnothing) = \mu_2(\cdot|\Theta \times M^{\Theta})$ denote the receiver's ex ante belief on types and choice functions; for each $m \in M$, we let

$$\mu_2(\cdot|m) = \mu_2(\cdot| \cup_{\theta\in\Theta} \{\theta\} \times M^{\Theta}(\theta, m))$$

denote the receiver's belief conditional on receiving message m.

After an on-path message $m^* \in M^*$, the receiver can update his initial theory: by D1 and B3, $\mu_2(\cdot|\varnothing)$ assigns positive probability to some type-choice function pair such that the type is associated with m^* by the choice function. Now consider an off-path message $m \in M\backslash M^*$; by endowing the receiver with a system of second-order beliefs, we can decompose his belief-revision process in two steps. First, the receiver revises his initial theory by conditioning on $\Theta \times M^{\Theta}(m)$, an event concerning only the sender's choice function. Second, he updates the revised theory $\mu_2(\cdot|\Theta \times M^{\Theta}(m))$ by considering that m gives joint information about type and choice function, i.e., by conditioning on the non-Cartesian event $\cup_{\theta\in\Theta} \{\theta\} \times M^{\Theta}(\theta, m)$: this is possible because, by D2 and the definition of P_2, the support of $\mu_2(\cdot|\Theta \times M^{\Theta}(m))$ is a cross-product of the form $\Theta \times E$, so every choice function $s_1 \in E$ is associated with every type, including the subset of types that pick m according to s_1.[17] Deriving $\mu_2(\cdot|m)$ from some $\nu \in P_2$ has bite when ν is also required to assign probability one to a subset of $\Theta \times M^{\Theta} \times \Delta(A^M)$ — Example 1 will clarify this point. Conditional on each on-path message $m \in M^*$, every $\nu \in P_2^*$ induces the same belief p^m over the sender's types:

$$p^m(\theta) = \frac{p(\theta)\, \nu^{\theta}(m)}{\sum_{\theta'\in\Theta} p(\theta')\, \nu^{\theta'}(m)}.$$

[15] In technical terms, fix a map ψ :suppmarg$_{M^{\Theta}}\eta \longrightarrow \Delta(A^M)$ such that $(s_1, \psi(s_1)) \in \text{supp}\eta$ for every $s_1 \in \text{suppmarg}_{M^{\Theta}}\eta$, let ψ' :supp$\eta \longrightarrow M^{\Theta} \times \Delta(A^M)$ be the map that associates each $(s_1, \mu_1) \in$suppη with $(s_1, \psi(s_1))$, and finally let η' be the pushforward of η through ψ'.

[16] Of course, conditional on events that do not rule out any choice function consistent with $(\nu^{\theta})_{\theta\in\Theta}$, the receiver could keep a belief in P_2^*. However, this is irrelevant for our analysis, so it is not formalized.

[17] Recall that, by definition, for each $s_1 \in M^{\Theta}(m)$ there is some θ such that $s_1(\theta) = m$.

So, for every $\mu_2 \in \Delta_2$, we have $\text{marg}_\Theta \mu_2(\cdot|m) = p^m$ for each $m \in M^*$.

A message m could be explained also by a more general theory than $\mu_2(\cdot|\Theta \times M^\Theta(m))$. It will be useful to start from the theory μ_2^1 conditional on the event that some type would not send an on-path message, i.e., event $\cup_{m\in M\backslash M^*} M^\Theta(m)$. Take note of the set M^1 of off-path messages that have strictly positive probability according to μ_2^1, then move to the theory μ_2^2 obtained by conditioning on the residual off-path-message event $\cup_{m\in M\backslash (M^*\cup M^1)} M^\Theta(m)$, and so on.

Remark 1 *Fix $\mu_2 \in \Delta_2$. There exists a partition $M^1, ..., M^n$ of $M\backslash M^*$ such that, for each $k = 1, ..., n$ and $m \in M^k$, the revised belief $\mu_2(\cdot|m)$ is derived from $\mu_2(\cdot|\Theta\times(\cup_{\bar{m}\in M^k \cup ... \cup M^n} M^\Theta(\bar{m})))$. Moreover, all these conditional beliefs have disjoint supports.*

As a consequence of Remark 1, either two beliefs $\mu_2(\cdot|m)$ and $\mu_2(\cdot|m')$ are derived from the same theory, or one is derived from a theory that cannot explain the other. This induces the kind of coherence among conditional beliefs anticipated in the Introduction.

Epistemic hypotheses We want to characterize the behavioral implications of the following epistemic hypotheses. We assume each player i is rational, i.e., maximizes subjective expected utility, and that i's (system of) beliefs satisfy the restrictions explained above, that is, they belong to Δ_i. Moreover, we assume that there is *common strong belief* (Battigalli and Siniscalchi 2002) of this. For the receiver, strong belief in an event E means that he assigns probability 1 to E conditional on every event that is consistent with E. For example, the receiver will assign probability 1 to the event that the sender is rational and has a first-order belief in her restricted set Δ_1 conditional on receiving a message that is consistent with this, i.e., that is optimal for at least one type under a belief that conforms to $(\nu^m)_{m\in M^*}$. For the sender, we only consider the belief at the beginning of the game, therefore belief and strong belief coincide. Common strong belief in rationality and the belief restrictions can be formally defined with the language of epistemic game theory, which provides a complete description of players' hierarchies of conditional beliefs. Here we only provide an informal description of our epistemic conditions, before characterizing their behavioral implications, step by step, with an elimination procedure:

1.S: Whatever her type, the sender is rational and her first-order belief belongs to Δ_1;

1.R: the receiver is rational and his system of second-order beliefs belong to Δ_2.

n.S: Whatever her type, the sender satisfies **n-1.S** and believes that the receiver satisfies **n-1.R**;

n.R: the receiver satisfies **n-1.R** and strongly believes that, whaterver her type, the sender satisfies **n-1.S**.

∞.S-R: For every n, the sender satisfies **n.S** whatever her type and the receiver satisfies **n.R**; that is, **1.S** and **1.R** hold, and there is common strong belief thereof.

Of course, common strong belief of **1.S** and **1.R** may be incompatible with the belief restrictions imposed at the outset: this will happen when the belief that all types choose the expected messages and that they can rationally do so under the same first-order belief is at odds with strategic reasoning, that is, with some order of mutual strong belief in **1.S** and **1.R**.

Path rationalizability with second-order independence Our goal is now to define a rationalizability procedure that either rejects the epistemic conditions (that is, rejects the expected outcome distribution as inconsistent with the other hypotheses), or calculates their behavioral implications (which may be weaker than what is expected). We will construct a version of Strong-Δ-Rationalizability (Battigalli 2003, Battigalli and Siniscalchi 2003) that accommodates not only the restrictions on first-order beliefs given by the prior and by the expected on-path behavior, but also the restriction on the second-order beliefs of the receiver given by the independence hypothesis. We call it "Path-rationalizability with second-order independence". The baseline definition of Strong-Δ-Rationalizability has been given an epistemic justification of the kind outlined above by Battigalli and Siniscalchi (2007) for the case of first-order belief restrictions;[18] the extension to second-order belief restrictions is relatively straightforward.

For each $(\theta, m, \mu_1) \in \Theta \times M \times \Delta(A^M)$, slightly abusing notation, let $u_1(\theta, m, \mu_1)$ denote the expected payoff of type θ given message m and the belief induced by μ_1 over the receiver's actions after m. Similarly, for each $(p', m, a) \in \Delta(\Theta) \times M \times A$, let

$$u_2\left(p', m, a\right) = \sum_{\theta \in \Theta} p'(\theta) u_2(\theta, m, a).$$

be the receiver's expected payoff after message m and action a, given the belief about types p'.

Definition 1 *Consider the following elimination procedure.*

Step 0 *For every $\theta \in \Theta$, let $\Sigma^0_{1,\theta} = M \times \Delta_1$. Let $\Sigma^0_1 = \Theta \times M^\Theta \times \Delta_1$, $\Sigma^0_2 = A^M$.*

Step n > 0 *For every $\theta \in \Theta$ and $(m, \mu_1) \in M \times \Delta_1$, let $(m, \mu_1) \in \Sigma^n_{1,\theta}$ if:*

> ***S1*** $\mu_1\left(\Sigma^{n-1}_2\right) = 1$*;*
>
> ***S2*** *for every $m' \in M$,*
> $$u_1(\theta, m, \mu_1) \geq u_1(\theta, m', \mu_1).$$

Let

$$\Sigma^n_1 = \Theta \times \{(s_1, \mu_1) \in M^\Theta \times \Delta_1 : \forall \theta \in \Theta, (s_1(\theta), \mu_1) \in \Sigma^n_{1,\theta}\}.$$

For every $s_2 \in A^M$, let $s_2 \in \Sigma^n_2$ if there exists $\mu_2 \in \Delta_2$ such that:

[18] Battigalli and Prestipino (2013) provide an alternate epistemic justification where the first-order belief restrictions are transparent, i.e., there is common belief at every node of the game that the restrictions hold.

R1 *for every $k = 1, ..., n-1$ and $C \in 2^{\Theta \times M^{\Theta}}$,*

$$\left(C \times \Delta(A^M)\right) \cap \Sigma_1^k \neq \emptyset \quad \Rightarrow \quad \mu_2\left(\Sigma_1^k | C\right) = 1;$$

R2 *for every $m \in M$ and $a \in A$,*

$$u_2\left(\mathrm{marg}_{\Theta}\mu_2(\cdot|m), m, s_2(m)\right) \geq u_2\left(\mathrm{marg}_{\Theta}\mu_2(\cdot|m), m, a\right).$$

Finally, let $\Sigma_{1,\theta}^{\infty} = \cap_{n \geq 0}\Sigma_{1,\theta}^n$ for each $\theta \in \Theta$, and $\Sigma_2^{\infty} = \cap_{n \geq 0}\Sigma_2^n$. The elements of each $\Sigma_{1,\theta}^{\infty}$ and of Σ_2^{∞} are called path-rationalizable (with second-order independence).

Path-rationalizability with second-order independence works as follows. At every step n, since $\mu_1 \in \Delta_1$, the sender's belief is consistent with ν^m for each on-path message $m \in M^*$. After every other message $m \in M \backslash M^*$, by S1, the sender believes that the receiver will play actions that are consistent with step $n-1$. Then, by S2, the sender chooses a message that is optimal given her type and belief. The receiver reasons as follows. At the beginning of the game, since $\mu_2 \in \Delta_2$, his belief is consistent with the prior on types, with the independence hypothesis, and with each ν^{θ}. Moreover, by R1, the receiver also believes that every type of the sender would choose a message that is consistent with step $n-1$. The two requirements can be mutually inconsistent: for an arbitrarily given outcome distribution, it may be the case that there is no $\mu_2 \in \Delta_2$ such that $\mu_2\left(\Sigma_1^{n-1}|\varnothing\right) = 1$. This is the only way Path-rationalizability with second-order independence can yield the empty set; we will expand on this later. After receiving an on-path message $m \in M^*$, the receiver simply updates his initial belief (recall that the actions in the support of ν^m are optimal by assumption). After an off-path message $m \in M \backslash M^*$, by R1, the receiver follows a version of the best rationalization principle:[19] If message m is consistent with step of reasoning $k \leq n-1$ for at least one type, the receiver revises his initial belief by conditioning on m an alternative theory that is consistent with k steps of reasoning of the sender. The restrictions to off-path beliefs imposed by Δ_2 cannot induce the empty set (second-order independence and R1 are always mutually compatible), but they can refine the set of actions that, by R2, the receiver could choose after m. The following example of Path-rationalizability illustrates this refinement of the receiver's off-path beliefs. The first entry in each box is the payoff of the sender.

Example 1

m_1	a_1	a_2	a_3
θ	$1,1$	$0,0$	$0,0$
θ'	$1,1$	$0,0$	$0,0$

m_2	a_1	a_2	a_3
θ	$3,0$	$0,3$	$0,2$
θ'	$2,3$	$0,0$	$0,2$

We will write a choice function $s_1 \in M^{\Theta}$ as $s_1(\theta).s_1(\theta')$, and a strategy $s_2 \in A^M$ as $s_2(m_1).s_2(m_2)$.

Let $p(\theta) = p(\theta') = 1/2$, $\nu^{\theta}(m_1) = \nu^{\theta'}(m_1) = 1$, $\nu^{m_1}(a_1) = 1$. So, Δ_1 is the set of beliefs that assign probability 1 to strategies $s_2 \in A^M$ with $s_2(m_1) = a_1$. For the receiver, every

[19] See Battigalli (2003) and the relevant references therein.

initial belief $\nu \in P_2^*$ is the product of the prior (by B1), a Dirac on $m_1.m_1$ (by B3), and a (finite-support) probability measure over $\Delta(A^M)$ (by B2). So, for every $\mu_2 \in \Delta_2$, we have

$$\mu_2(\{(\theta, m_1.m_1)\} \times \Delta(A^M)|\varnothing) = \mu_2(\{(\theta', m_1.m_1)\} \times \Delta(A^M))|\varnothing) = 1/2.$$

Given the belief that m_1 would be followed by a_1, the sender has the incentive to deviate to m_2 only if she assigns sufficiently high probability to a_1 after m_2: at least $1/3$ for type θ and $1/2$ for θ'. So we have

$$\begin{aligned}\Sigma^1_{1,\theta} &= \{m_1\} \times \{\mu_1 \in \Delta_1 : \mu_1(a_1.a_1) \leq 1/3\} \cup \{m_2\} \times \{\mu_1 \in \Delta_1 : \mu_1(a_1.a_1) \geq 1/3\},\\ \Sigma^1_{1,\theta'} &= \{m_1\} \times \{\mu_1 \in \Delta_1 : \mu_1(a_1.a_1) \leq 1/2\} \cup \{m_2\} \times \{\mu_1 \in \Delta_1 : \mu_1(a_1.a_1) \geq 1/2\}.\end{aligned}$$

From this, in preparation for step 2, observe that

$$\Theta \times \{m_1.m_1\} \times \{\mu_1 \in \Delta_1 : \mu_1(a_1.a_1) \leq 1/3\} \quad \subset \quad \Sigma^1_1, \tag{1}$$

$$\Theta \times \{m_2.m_1\} \times \left\{\mu_1 \in \Delta_1 \middle| \mu_1(a_1.a_1) \in \left[\frac{1}{3}, \frac{1}{2}\right]\right\} \quad \subset \quad \Sigma^1_1 \subset \Theta \times (M^\Theta \setminus \{m_1.m_2\}) \times \Delta_1, \tag{2}$$

where the last inclusion follows from the fact that for every $\mu_1 \in \Delta_1$ such that $(m_2, \mu_1) \in \Sigma^1_{1,\theta'}$, $(m_1, \mu_1) \notin \Sigma^1_{1,\theta}$. For the receiver, the first step eliminates the strategies that prescribe the dominated actions a_2 and a_3 after m_1.

The elimination of a_2 and a_3 after m_1 does not refine the sender's beliefs at the second step. For the receiver, by (1), there exists $\nu \in P_2^*$ such that $\nu(\Sigma^1_1) = 1$, thus there are beliefs $\mu_2 \in \Delta_2$ that satisfy R1 at the second step. Now recall how we break down the receiver's belief revision in two parts: first he revises his initial theory by conditioning on $\Theta \times M^\Theta(m_2)$ (the event that the choice function of the sender allows for m_2) and then he updates the revised theory taking into account the interaction between type and choice function, i.e., conditioning on $\{\theta\} \times \{m_2.m_1, m_2.m_2\} \cup \{\theta'\} \times \{m_1.m_2, m_2.m_2\}$, which yields the choice-relevant belief $\mu_2(\cdot|m_2)$. With this, the off-path beliefs that satisfy R1 do not justify a_1: by the first inclusion in (2), R1 imposes $\mu_2(\Sigma^1_1|\Theta \times M^\Theta(m_2)) = 1$, but then by the second inclusion $\mu_2(\cdot|\Theta \times M^\Theta(m_2))$ must assign marginal probability 0 to $m_1.m_2$. Therefore, $\mu_2(\cdot|m_2)$ cannot assign to θ' higher probability than the prior, which implies

$$\Sigma^2_2 = \{a_1.a_2, a_1.a_3\}.$$

At the third step, both types of the sender eliminate m_2, because every belief over Σ^2_2 justifies only m_1. Therefore, we have

$$\Sigma^3_{1,\theta} = \Sigma^3_{1,\theta'} = \{m_1\} \times \{\mu_1 \in \Delta_1 : \mu_1(\Sigma^2_2) = 1\}.$$

At the fourth step, the receiver refines his initial beliefs by assigning probability 1 to the beliefs of the sender that are compatible with step 3, but cannot refine the beliefs after m_2 compared to step 2, because m_2 is incompatible with step 3 for both types of the sender. Therefore, the path-rationalizable strategies of the receiver are $\{a_1.a_2, a_1.a_3\}$, and for each type of the sender the only path-rationalizable message is m_1. So, ν^θ, $\nu^{\theta'}$, and ν^{m_1} are compatible with strategic reasoning. Note also that every type and the receiver have only one path-rationalizable move, so no path-rationalizable move is unexpected. This is far from true in general: In Example 3, the expected moves are compatible with strategic reasoning, but also different moves are. $\triangle$

Given the assumption that every action in the support of ν^m is optimal under the belief p^m induced by $(\nu^\theta)_{\theta\in\Theta}$, as long as Σ_2^{n-1} is non-empty, it contains strategies that make S1 compatible with Δ_1, thus each $\Sigma_{1,\theta}^n$ is non-empty as well. Instead, we obtain an empty Σ_2^n when $M^*(\theta) \nsubseteq \mathrm{Proj}_M \Sigma_{1,\theta}^{n-1}$ for some $\theta \in \Theta$. In this case, R1 cannot be satisfied by any $\mu_2 \in \Delta_2$, because it implies disagreement with the given distribution on terminal nodes. The interpretation is that sending some message $m \in M^*(\theta)$ is incompatible with the (belief-restricted) strategic reasoning for type θ. But Σ_2^n can be empty even if $M^*(\theta) \subseteq \mathrm{Proj}_M \Sigma_{1,\theta}^{n-1}$ for every $\theta \in \Theta$. This happens when different types of the sender find the messages in M^* optimal only for different beliefs. This means that no choice function in $\mathrm{Proj}_{M^\Theta}\Sigma_1^{n-1}$ prescribes a message in $M^*(\theta)$ to every type θ,[20] and then, by second-order independence, every $\nu \in P_2$ with $\nu(\Sigma_1^{n-1}) = 1$ assigns positive probability to a triple (θ, s_1, μ_1) with $s_1(\theta) \notin M^*(\theta)$, which implies disagreement with the given outcome distribution. The interpretation is the following: the belief that every type θ would send a message in $M^*(\theta)$, even when her beliefs are independent of her type, is not compatible with the (belief-restricted) strategic reasoning for the receiver. The following example illustrates this second kind of inconsistency.

Example 2

m_1	a_1	a_2
θ	$1,0$	$0,0$
θ'	$1,0$	$0,0$

m_2	a_1	a_2
θ	$3,0$	$0,0$
θ'	$0,0$	$3,0$

Let $p(\theta) = p(\theta') = 1/2$, $\nu^\theta(m_1) = \nu^{\theta'}(m_1) = 1$, $\nu^{m_1}(a_1) = 1$. *So,* Δ_1 *is the set of beliefs that give probability one to strategies* $s_2 \in A^M$ *with* $s_2(m_1) = a_1$, *and for every* $\mu_2 \in \Delta_2$, *we have*

$$\mu_2(\{(\theta, m_1.m_1)\} \times \Delta(A^M)|\varnothing) = \mu_2(\{(\theta', m_1.m_1)\} \times \Delta(A^M)|\varnothing) = 1/2.$$

Given the belief in a_1 *after* m_1, *types* θ *and* θ' *have the incentive to deviate to* m_2 *if they assign at least probability* $1/3$ *to, respectively,* a_1 *and* a_2. *So we have*

$$\begin{aligned}\Sigma_{1,\theta}^1 &= \{m_1\} \times \{\mu_1 \in \Delta_1 : \mu_1(a_1.a_1) \le 1/3\} \cup \{m_2\} \times \{\mu_1 \in \Delta_1 : \mu_1(a_1.a_1) \ge 1/3\},\\ \Sigma_{1,\theta'}^1 &= \{m_1\} \times \{\mu_1 \in \Delta_1 : \mu_1(a_1.a_1) \ge 2/3\} \cup \{m_2\} \times \{\mu_1 \in \Delta_1 : \mu_1(a_1.a_1) \le 2/3\}.\end{aligned}$$

Note that there is no $\mu_1 \in \Delta_1$ *such that both* $(m_1, \mu_1) \in \Sigma_{1,\theta}^1$ *and* $(m_1, \mu_1) \in \Sigma_{1,\theta'}^1$. *Hence,* $m_1.m_1 \notin \mathrm{Proj}_{M^\Theta}\Sigma_1^1$. *Therefore, there is no* $\mu_2 \in \Delta_2$ *that satisfies R1 at step 2.* △

How can we check that Σ_2^n ($n > 1$) is not empty? First, we need every type to be indifferent among all her on-path messages: Every $\mu_1 \in \Delta_1$ induces the same beliefs $(\nu^m)_{m\in M^*}$ after the messages in M^*, so if θ had a strict ranking over $M^*(\theta)$ under some $\mu_1 \in \Delta_1$, this ranking would be the same under all $\mu_1 \in \Delta_1$, and there would be some message $m \in M^*(\theta)$

[20] Even if only one possible assignment of on-path messages to types was missing from $\mathrm{Proj}_{M^\Theta}\Sigma_1^{n-1}$, R1 and B2 would still be incompatible with B3. In any case, if one assignment is missing, then all the assignments are missing, because all on-path messages of a type must be justified by the same beliefs, as we will argue later.

such that $m \notin \text{Proj}_M \Sigma^1_{1,\theta}$. Second, we need that *one* belief of the sender compatible with step $n-1$ justifies an on-path message for every type. Provided that every type of the sender is indifferent among all her on-path messages, this belief justifies *all* on-path messages of all types.

Lemma 1 *Fix $n > 1$. We have $\Sigma^n_2 \neq \emptyset$ if and only if:*

1. *every $\theta \in \Theta$ is indifferent among all messages in $M^*(\theta)$ under beliefs $(\nu^m)_{m\in M^*}$;*
2. *there exists $\overline{\mu}^1 \in \Delta_1$ such that, for every $\theta \in \Theta$, $\left(m, \overline{\mu}^1\right) \in \Sigma^{n-1}_{1,\theta}$ for any $m \in M^*(\theta)$.*

Lemma 1 guarantees that, if $(\nu^\theta)_{\theta\in\Theta}$ and $(\nu^m)_{m\in M^*}$ are compatible with strategic reasoning, there is a belief $\overline{\mu}^1 \in \Delta_1$ over the strategically sophisticated strategies of the receiver so that every type of the sender has the incentive to stay on path. For each off-path message m, the probability measure induced by $\overline{\mu}^1$ after m can assign positive probability to actions that are optimal for the receiver only under different beliefs about the sender's type. Then, $(\nu^\theta)_{\theta\in\Theta}$ and $\overline{\mu}^1$, which induces ν^m after each $m \in M^*$, define the behavioral strategies of a Perfect Bayesian Equilibrium with (possibly) heterogenous off-path beliefs (PBH; Fudenberg and He, 2018).

Proposition 1 *Suppose that $\Sigma^3_2 \neq \emptyset$. Then, there exists a PBH $(\beta_1, \beta_2) \in (\Delta(M))^\Theta \times (\Delta(A))^M$ such that $\beta_1(\cdot|\theta) = \nu^\theta$ for each $\theta \in \Theta$ and $\beta_2(\cdot|m) = \nu^m$ for each $m \in M^*$.*

We obtain a PBH, and not just a self-confirming equilibrium (Fudenberg and Levine 1993), because of our restriction on the receiver's second-order beliefs.[21] Battigalli and Siniscalchi (2003) have shown that when the first-order beliefs are restricted by a given outcome distribution (as we assume), non-emptiness of Strong-Δ-Rationalizability guarantees that the distribution is induced by a self-confirming equilibrium.[22] Moreover, they show that in a signaling game non-emptiness of Strong-Δ-Rationalizability is equivalent to passing the Iterated Intuitive Criterion (Cho and Kreps 1987). Since our restrictions on first-order beliefs are of the same kind, also Path-rationalizability with second-order independence, when non-empty, guarantees that the corresponding PBH satisfies the Iterated Intuitive Criterion.[23] The examples above illustrate how the independence restriction on the receiver's second-order beliefs further refines his first-order beliefs: in Example 1, at step 2, the Intuitive Criterion would allow the receiver to assign high probability to θ' after m_2 and thus to play a_1, because in his mind θ and θ' could have different beliefs where θ' has the incentive to play m_2 and θ has the incentive to play m_1; in Example 2, it would be allowed to justify m_1 with different beliefs for different types, so we would not get the empty set.

[21] On top of this, we obtain a PBH and not just a Bayes-Nash equilibrium because the first step of reasoning guarantees that the receiver best replies to some off-path beliefs.

[22] Also in absence of the independence hypothesis, for a given type $\theta \in \Theta$, the messages in the support of ν^θ would still have to be justified by the same belief, because of the indifference among them given $(\nu^m)_{m\in M^*}$. Fudenberg and Kamada (2015) call this property of a self-confirming equilibrium "unitary beliefs".

[23] Given the non-monotonicity of strong belief, this observation requires proof. In the Appendix, we prove that path rationalizability is stronger than FERO, and Sobel et al. (1990) show that FERO is stronger than the iterated intuitive criterion. A direct proof can be provided by observing that there are no restrictions on the first-order beliefs of the sender about the off-path reactions of the receiver, and then adapting the techniques of Catonini (2020) to a signaling game.

A simpler algorithm The systems of second-order beliefs of the receiver allow for a transparent representation of his process of belief formation. However, they are redundant in two dimensions for the calculation of the behavioral implications of the epistemic conditions. First, using choice functions in place of messages in the space of uncertainty of the receiver allows for an indirect representation of the independence hypothesis through his first-order belief: the receiver's first-order belief shall (be a product measure and) assign probability 1 to the strategies of the sender that prescribe to each type a message that is optimal under the same belief. Second, just a few conditional beliefs are sufficient to derive the beliefs of the receiver at the moment of choosing an action — those described in Remark 1. Note also that the corresponding conditioning events have disjoint supports. Then, adding the initial theory as first measure, they can be organized in a *Lexicographic Conditional Probability System* (Blume et al., 1991; henceforth, LCPS), a finite sequence of probability measures with disjoint supports. Conversely, from an LCPS, coupled with a full-support probability measure ν, one can derive a CCPS as follows: for each conditional event C, derive the belief from the first measure in the LCPS that assigns positive probability to C, if any, otherwise from ν. With this, we can rewrite path rationalizability with second-order independence as a simpler procedure that uses only first-order LCPSs. So, let Δ_2^ℓ denote the set of LCPSs $\bar{\mu}_2 = (\mu_2^1, ..., \mu_2^l)$ over $\Theta \times M^\Theta$, such that:

D1$^\ell$ $\mu_2^1 = p \times \eta^*$, where $\eta^*(s_1) = \prod_{\theta\in\Theta} \nu^\theta(s_1(\theta))$ for each $s_1 \in M^\Theta$;

D2$^\ell$ for each $j = 2, ..., l$, $\mu_2^j = p \times \eta$ for some $\eta \in \Delta(M^\Theta)$.

Conditions D1$^\ell$ and D2$^\ell$ mirror conditions D1 and D2: the theories about the sender must be product measures between the prior and, for the primary theory, a probability measure over strategies that is consistent with $(\nu^\theta)_{\theta\in\Theta}$. In the following definition, conditioning on m will stay for conditioning on $\cup_{\theta\in\Theta} \{\theta\} \times M^\Theta(\theta, m)$. The usual abuses of notation for expected payoffs apply.

Definition 2 *Consider the following reduction procedure.*

Step 0 *Let* $\Sigma_1^{\ell,0} = M^\Theta$ *and* $\Sigma_2^{\ell,0} = A^M$.

Step n > 0 *For each* $s_1 \in \Sigma_1^{\ell,n-1}$, *let* $s_1 \in \Sigma_1^{\ell,n}$ *if there exists* $\mu_1 \in \Delta_1$ *such that:*

S1$^\ell$ $\mu_1(\Sigma_2^{\ell,n-1}) = 1$;

S2$^\ell$ *for every* $\theta \in \Theta$, *for every* $m \in M$,

$$u_1(\theta, s_1(\theta), \mu_1) \geq u_1(\theta, m, \mu_1).$$

For each $s_2 \in \Sigma_2^{\ell,n-1}$, *let* $s_2 \in \Sigma_2^{\ell,n}$ *if there exists* $\bar{\mu}_2 = (\mu_2^1, ..., \mu_2^l) \in \Delta_2^\ell$ *such that:*

R1$^\ell$ $\mu_2^j(\Theta \times \Sigma_1^{\ell,n-1}) = 1$ *for each* $j = 1, ..., l$;

R2$^\ell$ *for each* $m \in M$ *with* $M^\Theta(m) \cap \Sigma_1^{\ell,n-1} \neq \emptyset$, *there exists* $j \in \{1, ..., l\}$ *such that* $\mu_2^j(\Theta \times M^\Theta(m)) > 0$, *and calling* k *the smallest of such* j*'s,*

$$u_2\left(\mathrm{marg}_\Theta(\mu_2^k|m), m, s_2(m)\right) \geq u_2\left(\mathrm{marg}_\Theta(\mu_2^k|m), m, a\right)$$

for every $a \in A$.

Finally, let $\Sigma_1^{\ell,\infty} = \cap_{n\geq 0}\Sigma_1^{\ell,n}$ *and* $\Sigma_2^{\ell,\infty} = \cap_{n\geq 0}\Sigma_2^{\ell,n}$.

Requirements S1$^\ell$ and S2$^\ell$ coincide with S1 and S2, except that here we directly construct the choice functions of the sender where every type best replies to the same belief. Requirement R1$^\ell$ is simpler than R1 in that it only deals with beliefs over the choice functions of the sender that survived the previous step of elimination, not the earlier ones. After the messages that cannot be rationalized based on such choice functions, the optimality of the reaction is guaranteed by the fact that only the strategies of the receiver that survived the previous step are considered (that is, we defined a *reduction* procedure). On the other hand, by R2$^\ell$ the beliefs in the LCPS must be able to explain all the messages that survived the previous step, and the first belief that can explain a message m must justify the reaction of the receiver (as in R2). We now formalize the equivalence between the two procedures.

Proposition 2 *For every* $n \geq 0$*, we have* $\Sigma_2^{\ell,n} = \Sigma_2^n$*, and for every* $(\theta, m) \in \Theta \times M$*, we have* $m \in \mathrm{Proj}_M \Sigma_{1,\theta}^n$ *if and only if* $M^\Theta(\theta, m) \cap \Sigma_1^{\ell,n} \neq \emptyset$*.*

It is also worth noting that considering theories with overlapping supports would not expand the set of justifiable strategies of the receiver, because the recevier relies on theory μ^k only conditional on the event that no type-function pair that is assigned positive probability by the theories $\mu^1, ..., \mu^{k-1}$ is consistent with the observed message.

4 Comparison with divinity

Banks and Sobel (1987) call a sequential equilibrium "divine" when it survives an iterated procedure of refinement of off-path beliefs inspired by the idea that the beliefs of the sender do not differ across types. To appreciate similarities and differences between our analysis and Divine Equilibrium, it is enough to focus on the first two steps of reasoning. To facilitate this comparison, we report here the first two steps of the iterative procedure that defines Divine Equilibrium, and we jointly call them "divinity criterion." Banks and Sobel focus directly on sequential equilibrium, which in signaling games coincides with Perfect Bayesian Equilibrium (with common off-path beliefs). However, we start from a Bayes-Nash equilibrium $(\beta_1^*, \beta_2^*) \in (\Delta(M))^\Theta \times (\Delta(A))^M$ to show that PBE emerges endogenously from their conditions. For each $\theta \in \Theta$, let $M^*(\theta) = \mathrm{supp}\beta_1^*(\cdot|\theta)$, and let $M^* := \cup_{\theta\in\Theta} M^*(\theta)$. For any $m \in M\backslash M^*$ and any map $\sigma \in [0,1]^\Theta$ that assigns to each $\theta \in \Theta$ a probability of playing m (positive for some θ), let $p(\cdot|m;\sigma)$ denote the probability measure over types derived from the prior with Bayes rule. Let σ^0 denote the constant map that assign 0 to every type. Finally. let $\mathrm{Conv}(Y)$ denote the convex hull of a set Y.

Definition 3 *Fix a Bayes-Nash equilibrium* (β_1^*, β_2^*)*. For each* $m \in M\backslash M^*$*, let*

$$\begin{aligned}
\Sigma_1^d(m) &:= \left\{\sigma \in [0,1]^\Theta : \exists\alpha \in \Delta(A), \forall\theta \in \Theta, \sigma(\theta) \in \arg\max_{\pi\in[0,1]} \pi u_1(\theta, m, \alpha) + (1-\pi)u_1(\theta, \beta_1^*, \beta_2^*)\right\};\\
\Gamma(m) &:= \left\{p' \in \Delta(\Theta) : \exists\sigma \in \Sigma_1^d(m)\backslash\left\{\sigma^0\right\}, p' = p(\cdot|m;\sigma)\right\};\\
\Sigma_2^d(m) &:= \left\{\alpha \in \Delta(A) : \exists p' \in \mathrm{Conv}(\Gamma(m)), \mathrm{supp}\alpha \subseteq \arg\max_{a\in A} u_2(p', m, a)\right\}.
\end{aligned}$$

We say that (β_1^*, β_2^*) satisfies the ***divinity criterion*** *if for each $m \in M \backslash M^*$*

$$\begin{aligned}\Gamma(m) &\neq \emptyset \Rightarrow \beta_2^*(\cdot|m) \in \Sigma_2^d(m),\\ \Gamma(m) &= \emptyset \Rightarrow \exists p' \in \Delta(\Theta), \text{supp}\beta_2^*(\cdot|m) \subseteq \arg\max_{a\in A} u_2(p', m, a).\end{aligned}$$

Note that all the actions in the support of each $\beta_2^*(\cdot|m)$ must best reply to the same belief over the sender's types. Then, we have the following.

Remark 2 *If (β_1^*, β_2^*) satisfies the divinity criterion, then it is a Perfect Bayesian Equilibrium.*

The divinity criterion guarantees that the equilibrium distributions over messages and actions are compatible with the first two steps of reasoning under our hypotheses.

Theorem 1 *Fix a Bayes-Nash Equilibrium (β_1^*, β_2^*) that satisfies the divinity criterion and let $\nu^\theta := \beta_1^*(\cdot|\theta)$ for each $\theta \in \Theta$, $\nu^m := \beta_2^*(\cdot|m)$ for each $m \in M^*$. Then*

$$\times_{m\in M} \text{supp}\beta_2^*(\cdot|m) \subseteq \Sigma_2^2.$$

Proof of Theorem 1. Fix $m \in M \backslash M^*$. Suppose first that $\Gamma(m) = \emptyset$. Then, there is $p' \in \Delta(\Theta)$ such that

$$\text{supp}\beta_2^*(\cdot|m) \subseteq \arg\max_{a\in A} u_2(p', m, a).$$

Fix $\nu_{2,m} \in P_2$ such that $\text{marg}_\Theta(\nu_{2,m}|m) = p'$; it exists because the prior has full support.

Now suppose that $\Gamma(m) \neq \emptyset$. Then, there is $p' \in \text{Conv}(\Gamma(m))$ such that $\text{supp}\beta_2^*(\cdot|m) \subseteq \arg\max_{a\in A} u_2(p', m, a)$. Write $p' = \gamma^1 p^1 + ... + \gamma^n p^n$ — a convex combination of points in $\Gamma(m)$. Fix $j = 1, ..., n$. Then, there is $\sigma^j \in \Sigma_1^d(m)$ such that $p^j = p(\cdot|m; \sigma^j)$. Hence, there is $\alpha \in \Delta(A)$ such that

$$\sigma^j(\theta) \in \arg\max_{\pi\in[0,1]} \pi u_1(\theta, m, \alpha) + (1-\pi) u_1(\theta, \beta_1^*, \beta_2^*) \tag{3}$$

for each $\theta \in \Theta$. Construct $\mu_1 \in \Delta(A^M)$ that induces α after m and $\beta_2^*(\cdot|m')$ after each $m' \neq m$. Since (β_1^*, β_2^*) is an equilibrium, for each $\theta \in \Theta$ we have

$$u_1(\theta, m', \mu_1) \geq u_1(\theta, m'', \mu_1)$$

for each $m' \in M^*$ and $m'' \in M \backslash M^*$ with $m'' \neq m$. But then, for each $\theta \in \Theta$, by (3) we get

$$m \in \arg\max_{m'} u_1(\theta, m', \mu_1) \quad \text{if } \sigma^j(\theta) > 0, \tag{4}$$

$$M^*(\theta) \subseteq \arg\max_{m'} u_1(\theta, m', \mu_1) \quad \text{if } \sigma^j(\theta) < 1. \tag{5}$$

Define $\widetilde{\nu}^\theta \in \Delta(M)$ as $\widetilde{\nu}^\theta(m) = \sigma^j(\theta)$ and $\widetilde{\nu}^\theta(m') = \nu^\theta(m')\left(1 - \sigma^j(\theta)\right)$ for each $m' \neq m$. Define $\eta^j \in \Delta(M^\Theta)$ as

$$\forall s_1 \in M^\Theta, \quad \eta^j(s_1) = \prod_{\theta\in\Theta} \widetilde{\nu}^\theta(s_1(\theta)).$$

Let $\widehat{\eta}^j = \eta^j \times \delta_{\mu_1}$. Note that, for each $\theta \in \Theta$,

$$\widehat{\eta}^1(M^\Theta(\theta, m) \times \Delta(A^M)) = \eta^1(M^\Theta(\theta, m)) = \sigma^j(\theta) \tag{6}$$

For every $s_1 \in M^\Theta$ with $\eta^j(s_1) > 0$, for each $\theta \in \Theta$, we have $\sigma^j(\theta) > 0$ if $s_1(\theta) = m$, and $\sigma^j(\theta) < 1$ if $s_1(\theta) \in M^*(\theta)$ (there is no third possibility). Then, by (4) and (5), we have $(s_1(\theta), \mu_1) \in \Sigma^1_{1,\theta}$. Thus, $\Theta \times \{(s_1, \mu_1)\} \subseteq \Sigma^1_1$. Hence, $\widehat{\eta}^j(\mathrm{Proj}_{M^\Theta \times \Delta(A^M)} \Sigma^1_1) = 1$. Finally, let

$$\begin{aligned}
\widetilde{\delta}^j &= \frac{\gamma^j}{\sum\limits_{\theta \in \Theta} p(\theta)\sigma^j(\theta)}, \\
\delta^j &= \frac{\widetilde{\delta}^j}{\sum\limits_{k=1,\dots n} \widetilde{\delta}^k},
\end{aligned}$$

and for future reference, observe that

$$\frac{\delta^j}{\sum\limits_{k=1,\dots n} \delta^k \sum\limits_{\theta \in \Theta} p(\theta)\sigma^k(\theta)} = \frac{\gamma^j}{\sum\limits_{\theta \in \Theta} p(\theta)\sigma^j(\theta) \cdot \sum\limits_{k=1,\dots n} \widetilde{\delta}^k} \cdot \frac{\sum\limits_{k=1,\dots n} \widetilde{\delta}^k}{\sum\limits_{k=1,\dots n} \gamma^k} = \frac{\gamma^j}{\sum\limits_{\theta \in \Theta} p(\theta)\sigma^j(\theta)}. \tag{7}$$

Now let $\widehat{\eta} = \delta^1 \widehat{\eta}^1 + \dots + \delta^n \widehat{\eta}^n$. Let $\nu_{2,m} = p \times \widehat{\eta}$. Clearly, $\nu_{2,m} \in P_2$ and $\nu_{2,m}(\Sigma^1_1) = 1$. Let $C_m = \left(\cup_{\theta \in \Theta} \{\theta\} \times M^\Theta(\theta, m)\right) \times \Delta(A^M)$. For each $\overline{\theta} \in \Theta$, we have

$$\begin{aligned}
& (\nu_{2,m}|C_m)\left(\{\overline{\theta}\} \times M^\Theta(\overline{\theta}, m) \times \Delta(A^M)\right) \\
= \ & \frac{\nu_{2,m}(\{\overline{\theta}\} \times M^\Theta(\overline{\theta}, m) \times \Delta(A^M))}{\nu_{2,m}(C_m)} \\
= \ & \frac{\delta^1 p(\overline{\theta})\widehat{\eta}^1(M^\Theta(\overline{\theta}, m) \times \Delta(A^M)) + \dots + \delta^n p(\overline{\theta})\widehat{\eta}^n(M^\Theta(\overline{\theta}, m) \times \Delta(A^M))}{\delta^1 \sum\limits_{\theta \in \Theta} p(\theta)\widehat{\eta}^1(M^\Theta(\theta, m) \times \Delta(A^M)) + \dots + \delta^n \sum\limits_{\theta \in \Theta} p(\theta)\widehat{\eta}^n(M^\Theta(\theta, m) \times \Delta(A^M)} \\
= \ & \frac{\delta^1 p(\overline{\theta})\sigma^1(\overline{\theta})}{\sum\limits_{k=1,\dots n} \delta^k \sum\limits_{\theta \in \Theta} p(\theta)\sigma^k(\theta)} + \dots + \frac{\delta^n p(\overline{\theta})\sigma^n(\overline{\theta})}{\sum\limits_{k=1,\dots n} \delta^k \sum\limits_{\theta \in \Theta} p(\theta)\sigma^k(\theta)} \\
= \ & \gamma^1 \frac{p(\overline{\theta})\sigma^1(\overline{\theta})}{\sum\limits_{\theta \in \Theta} p(\theta)\sigma^1(\theta)} + \dots + \gamma^n \frac{p(\overline{\theta})\sigma^n(\overline{\theta})}{\sum\limits_{\theta \in \Theta} p(\theta)\sigma^n(\theta)} \\
= \ & \gamma^1 p^1(\overline{\theta}) + \dots + \gamma^n p^n(\overline{\theta}) = p'(\overline{\theta}),
\end{aligned}$$

where the third equality follows from (6) and the fourth from (7). So,

$$\mathrm{supp}\beta^*_2(\cdot|m) \subseteq \arg\max_{a \in A} u_2(\mathrm{marg}_\Theta(\nu_{2,m}|C_m), m, a).$$

Note that, for all $m, m' \in M \backslash M^*$ with $m \neq m'$, $\nu_{2,m}(\Theta \times M^\Theta(m') \times \Delta(A^M)) = 0$. Hence, there exists $\mu_2 \in \Delta_2$ such that $\mu_2(\cdot|\varnothing) = \nu^*$ for some $\nu^* \in P^*_2$, and $\mu_2(\cdot|m) = \nu_{2,m}|C_m$

for each $m \in M \backslash M^*$. Fix $s_2 \in A^M$ such that $s_2(m) \in \text{supp}\beta_2^*(\cdot|m)$ for each $m \in M$. By construction, μ_2 satisfies R1 at step 2 and R2 with s_2. So $s_2 \in \Sigma_2^2$. ■

The converse of Theorem 1 is not true: even if $\times_{m\in M}\text{supp}\beta_2^*(\cdot|m) \subseteq \Sigma_2^\infty$, (β_1^*, β_2^*) might not satisfy the divinity criterion. To see this, we now formalize the solution to the example of Section 2.

Example 3

m_1	a_1	a_2	a_3
θ^h	$0,3$	$4,2$	$9,0$
θ^ℓ	$0,3$	$4,2$	$9,0$

m_2	a_1	a_2	a_3
θ^h	$-2,3$	$2,5$	$7,3$
θ^ℓ	$-3,3$	$1,2$	$6,0$

m_3	a_1	a_2	a_3
θ^h	$-5,3$	$-1,5$	$4,6$
θ^ℓ	$-8,3$	$-4,2$	$1,0$

The prior is $p(\theta^h) = p(\theta^\ell) = 1/2$. *Consider the equilibrium* (β_1^*, β_2^*) *with* $\beta_1^*(m_1|\theta^h) = \beta_1^*(m_1|\theta^\ell) = 1$, $\beta_2^*(a_1|m_1) = \beta_2^*(a_1|m_2) = \beta_2^*(a_2|m_3) = 1$. *We have*

$$\Sigma_1^d(m_2) = \Sigma_1^d(m_3) = \{[0,1] \times \{0\}\} \cup \{\{1\} \times [0,1]\}.$$

For each $k = 2, 3$, *the first component in the union is justified by beliefs of the sender that make* θ^h *indifferent between* m_1 *and* m_k, *thus* θ^ℓ *strictly prefer* m_1, *and the second component by beliefs that make* θ^ℓ *indifferent between* m_1 *and* m_k, *thus* θ^h *strictly prefer* m_k. *With this, we get*

$$\Gamma(m_2) = \Gamma(m_3) = \left\{p' \in \Delta(\Theta) \middle| p'(\theta^h) \geq 1/2\right\}.$$

But then (abusing notation),

$$\Sigma_2^d(m_2) = \{a_2\}, \quad \Sigma_2^d(m_3) = \Delta(\{a_2, a_3\}).$$

so $\beta_2^*(\cdot|m_2) \notin \Sigma_2^d(m_2)$: (β_1^*, β_2^*) *does not satisfy the divinity criterion (and no equilibrium with* $\beta_1^*(m_1|\theta^h) = \beta_1^*(m_1|\theta^\ell) = 1$ *does).*

Now we turn to Path-rationalizability with second-order independence. We have

$$\{(m_1, \delta_{a_1.a_1.a_2}), (m_3, \delta_{a_1.a_2.a_3})\} \subset \Sigma^1_{1,\theta^h}, \quad (8)$$

$$\{(m_1, \delta_{a_1.a_1.a_2})\} \cup (\{m_2, m_3\} \times \{\delta_{a_1.a_2.a_3}\}) \subset \Sigma^1_{1,\theta^\ell}. \quad (9)$$

We check whether $a_1.a_1.a_2 \in \Sigma_2^2$. *Consider any* $\mu_2 \in \Delta_2$ *such that*

$$\mu_2(\cdot|\Theta \times \left(M^\Theta \backslash \{m_1.m_1\}\right)) = p \times \eta \times \delta_{\mu_1},$$

where $\bar{\mu}_1 = \delta_{a_1.a_2.a_3}$ *and* η *assigns probability* $1/2$ *to* $m_3.m_2$ *and* $m_3.m_3$. *By (8) and (9),* μ_2 *satisfies R1. We get*

$$\begin{aligned} \text{marg}_\Theta(\nu|m_2)(\theta^\ell) &= 1, \\ \text{marg}_\Theta(\nu|m_3)(\theta^\ell) &= 1/3; \end{aligned}$$

thus $a_1.a_1.a_2$ *satisfies R2 with* μ_2. *Hence,* $a_1.a_1.a_2 \in \Sigma_2^2$.

Since also $a_1.a_2.a_3$ *survives the second step, an easy inductive argument shows that the equilibrium survives all steps of path rationalizability.* △

To conclude, note that in the example, after m_3, a_1 and a_3 best reply only to disjoint sets of beliefs about the sender's type. Path-rationalizability with second-order independence allows the sender to assign positive probability both to a_1 and a_3 at all steps. For Divine Equilibrium, this is not allowed, because a mix of a_1 and a_3 is not a best response to any belief.

5 Appendix

5.1 Complete information game

In this appendix, we analyze the sender-receiver game as a complete-information game with asymmetric observation of an initial chance move. Then, we show the equivalence of the analysis with the incomplete-information approach of the main body, and the relationship with the notion of fixed-equilibrium rationalizable outcome (henceforth, FERO) of Sobel et al. (1990).

The timing of the game is as follows.

1. The pseudo-player chance chooses the value of θ from Θ, according to the commonly known distribution p.
2. The sender observes θ and chooses a message m from M.
3. The receiver observes m but not θ and chooses an action a from A.

At the beginning of the game, the sender and the receiver have a belief over the strategies[24] of the other player and chance. We assume that the *sender believes that there is no correlation between the move of chance and the strategy of the receiver*. Therefore, after observing the move of chance, her marginal belief about the strategy of the receiver will stay the same. For this reason, instead of writing the sender's entire conditional probability system, we simply write one probability measure $\mu_1 \in \Delta(A^M)$ over the receiver's choice functions, with the understanding that the initial belief of the sender is $p \times \mu_1 \in \Delta\left(\Theta \times A^M\right)$.[25] As in the main body, the sender's belief about the actions of the receiver after the messages in M^* is given by $(\nu^m)_{m \in M^*}$. So, μ_1 must belong to this restricted set:

$$\hat{\Delta}_1 = \left\{\mu_1 \in \Delta(A^M) \left| \forall (m,a) \in M^* \times A, \mu_1\left(\left\{s_2 \in A^M \left| s_2(m) = a\right.\right\}\right) = \nu^m(a)\right.\right\},$$

which coincides with Δ_1 defined in the main body.

Given the independence restriction on the sender's first-order beliefs, we do not need to restrict the second-order beliefs of the receiver exogenously. Therefore, we endow the receiver with a CCPS over $\Theta \times M^\Theta$, thus considering the collection of conditional events: $2^{\Theta \times M^\Theta} \backslash \{\emptyset\}$. Let $\hat{\Delta}_2$ be the set of CCPSs $\hat{\mu}_2$ that satisfy the following two conditions:

[24] Here we talk of strategies of the sender because the sender can formulate a plan at the ex ante stage, before observing the chance move.

[25] More generally, the sender's prior might be any product measure $q \times \mu_1$ with $q \in \Delta^\circ(\Theta)$. What is important is that it is commonly believed that the receiver's prior on Θ is p.

D1' $\hat{\mu}_2(\cdot|\varnothing) = p \times \eta^*$, where $\eta^*(s_1) = \prod_{\theta\in\Theta} \nu^\theta(s_1(\theta))$ for each $s_1 \in M^\Theta$;

D2' for each $E \in 2^{M^\Theta} \setminus \{\emptyset\}$, $\hat{\mu}_2(\cdot|\Theta \times E) = p \times \eta$ for some $\eta \in \Delta(M^\Theta)$.

Condition D1' requires that the initial belief of the receiver is a product measure between the objective distribution of chance moves p and a distribution over sender's choice functions that is consistent with $(\nu^\theta)_{\theta\in\Theta}$. Condition D2' requires that, conditional on every subset of choice functions, the receiver's belief is a product measure between p and some distribution over choice functions. While imposing independence between the chance move and the strategy of the sender is natural (for the reasons we mentioned in the main body), one may wonder why we need to do it, given the independence condition we have already imposed on the sender's beliefs. The reason is the following: assigning probability one to choice functions of the sender that are justifiable under independence does not automatically imply that the receiver believes in independence, because his belief could still match different chance moves with different sender's choice functions.

It is easy to see that a CCPS $\hat{\mu}_2$ is in $\hat{\Delta}_2$ if and only if it can be obtained by marginalization from some $\mu_2 \in \Delta_2$.

Remark 3 *Fix a CCPS $\hat{\mu}_2$ on $\Theta \times M^\Theta$ with conditional events $2^{\Theta\times M^\Theta}$. We have $\hat{\mu}_2 \in \hat{\Delta}_2$ if and only if there exists $\mu_2 \in \Delta_2$ such that, for each $C \in 2^{\Theta\times M^\Theta}$, $\hat{\mu}_2(\cdot|C) = \mathrm{marg}_{\Theta\times M^\Theta}\mu_2(\cdot|C)$.*

Proof. If: Let $\mathcal{D}$ denote the collection of all $D \in 2^{\Theta\times M^\Theta}$ such that $\mu_2(D\times\Delta(A^M)|C) = 0$ for every $C \supset D$. Thus, $\mathcal{D}$ is also the collection of all $D \in 2^{\Theta\times M^\Theta}$ such that $\hat{\mu}_2(D|C) = 0$ for every $C \supset D$. By D2, for each $D \in \mathcal{D}$, $\mu_2(\cdot|D) = p \times \eta$ for some $\eta \in \Delta(M^\Theta \times \Delta(A^M))$. Thus, $\hat{\mu}_2(\cdot|D) = p \times \mathrm{marg}_{M^\Theta}\eta$. Hence, $\hat{\mu}_2$ satisfies D2'. Moreover, by D1, there exists $\eta \in \Delta(M^\Theta \times \Delta(A^M))$ such that $\mu_2(\cdot|\varnothing) = p \times \eta$ and $\eta(M^\Theta(\theta, m) \times \Delta(A^M)) = \nu^\theta(m)$ for each $(\theta, m) \in \Theta \times M$. Hence, $\hat{\mu}_2$ satisfies D1'.

Only if: Fix any $\mu_1 \in \Delta(A^M)$ and define μ_2 as $\mu_2(\cdot|C) = \hat{\mu}_2(\cdot|C) \times \delta_{\mu_1}$ for each $C \in 2^{\Theta\times M^\Theta}$. The proof that $\mu_2 \in \Delta_2$ mirrors the one above and is therefore omitted. ■

Given our first-order belief restrictions, we now define Strong-Δ-Rationalizability for the game with complete information, which for future reference we call "Path-rationalizability with first-order independence".

Definition 4 *Consider the following elimination procedure.*

Step 0 *Let $\overline{\Sigma}_1^0 = M^\Theta$, $\overline{\Sigma}_{-2}^0 = \Theta \times M^\Theta$, and $\overline{\Sigma}_2^0 = A^M$.*

Step n > 0 *For each $s_1 \in M^\Theta$, let $s_1 \in \overline{\Sigma}_1^n$ if there exists $\mu_1 \in \hat{\Delta}_1$ such that:*

> ***S1'*** $\mu_1(\overline{\Sigma}_2^{n-1}) = 1$;
>
> ***S2'*** *for every $m \in M$,*
> $$u_1(\theta, s_1(\theta), \mu_1) \geq u_1(\theta, m, \mu_1).$$

Let $\overline{\Sigma}_{-2}^n = \Theta \times \overline{\Sigma}_1^n$.

For each $s_2 \in A^M$, let $s_2 \in \overline{\Sigma}_2^n$ if there exists $\hat{\mu}_2 \in \hat{\Delta}_2$ such that

R1' *for every $k = 1, ..., n-1$ and $C \in 2^{\Theta \times M^\Theta}$, if $C \cap \overline{\Sigma}_{-2}^k \neq \emptyset$, then $\hat{\mu}_2(\overline{\Sigma}_{-2}^k | C) = 1$;*

R2' *for every $m \in M$ and $a \in A$,*

$$u_2\left(\text{marg}_\Theta \hat{\mu}_2(\cdot|m), m, s_2(m)\right) \geq u_2\left(\text{marg}_\Theta \hat{\mu}_2(\cdot|m), m, a\right).$$

Finally, let $\overline{\Sigma}_1^\infty = \cap_{n \geq 0} \overline{\Sigma}_1^n$ and $\overline{\Sigma}_2^\infty = \cap_{n \geq 0} \overline{\Sigma}_2^n$. The elements in $\overline{\Sigma}_1^\infty$ and $\overline{\Sigma}_2^\infty$ are called path-rationalizable with first-order independence.

The difference between Path-rationalizability with first-order and second-order independence lies in the substitution between the restriction that the sender's first-order belief is independent of the chance move/type, and the restriction that the receiver believes this is the case. While the first restriction truly constrains the strategy of the sender, the second only establishes connections between the moves of dfferent types in the mind of the receiver. Nonetheless, the two scenarios are equivalent for the receiver's choices, and then also for the choices of the sender for each chance move/type after strategic reasoning.

Proposition 3 *A strategy of the receiver is path-rationalizable with first-order independence if and only if it is path-rationalizable with second-order independence.*

Every strategy of the sender that is path-rationalizable with first-order independence prescribes after each move of chance a message that is path-rationalizable with second-order independence for the corresponding type.

Every message that is path-rationalizable with second-order independence for a type is prescribed after the corresponding move of chance by a strategy that is path-rationalizable with first-order independence.

Proof.

Assume by way of induction that, for each $k = 0, ..., n$,

(i) $\overline{\Sigma}_2^k = \Sigma_2^k$,

(ii) for every $s_1 \in \overline{\Sigma}_1^k$, for every $\theta \in \Theta$, $s_1(\theta) \in \text{Proj}_M \Sigma_{1,\theta}^k$,

(iii) for every $\mu_1 \in \Delta_1$ with $\mu_1(\Sigma_2^{k-1}) = 1$, for every $s_1 \in M^\Theta$ such that $(s_1(\theta), \mu_1) \in \Sigma_{1,\theta}^k$ for each $\theta \in \Theta$, $s_1 \in \overline{\Sigma}_1^k$.

All is trivally true for $n = 0$.

Recall that $\Delta_1 = \hat{\Delta}_1$. Moreover, by the induction hypothesis (i), S1 and S1' at step $n+1$ coincide. Then, (ii) and (iii) for $n+1$ follow.

To prove (i) for $n+1$, we will show in the next paragraph that, for each $k = 1, ..., n$, $\overline{\Sigma}_{-2}^k = \text{Proj}_{\Theta \times M^\Theta} \Sigma_1^k$. But then, $\mu_2 \in \Delta_2$ satisfies R1 if and only if its marginal over $\Theta \times M^\Theta$ satisfies R1'. At the same time, by Remark 3, Δ_2 and $\hat{\Delta}_2$ have the same marginals over $\Theta \times M^\Theta$. Therefore, the subset of Δ_2 that satisfies R1 and the subset of $\hat{\Delta}_2$ that satisfies R1' have the same marginals over Θ. Given that R2 and R2' are identical for the same marginal over Θ, we obtain $\overline{\Sigma}_2^n = \Sigma_2^n$.

Now we prove $\overline{\Sigma}_{-2}^{k} = \text{Proj}_{\Theta\times M^{\Theta}}\Sigma_1^k$. First the inclusion $\overline{\Sigma}_{-2}^{k} \subseteq \text{Proj}_{\Theta\times M^{\Theta}}\Sigma_1^k$. Fix $(\theta, s_1) \in \overline{\Sigma}_{-2}^{k}$. Thus, there is $\bar{\mu}_1 \in \hat{\Delta}_1$ with $\bar{\mu}_1(\overline{\Sigma}_2^{k-1}) = 1$ such that s_1 and $\bar{\mu}_1$ satisfy S2'. By the induction hypothesis (i), $\bar{\mu}_1(\Sigma_2^{k-1}) = 1$, and by $\Delta_1 = \hat{\Delta}_1$, $\bar{\mu}_1 \in \Delta_1$, so by S2', $s_1(\theta')$ is optimal for every $\theta' \in \Theta$ under $\bar{\mu}_1$, thus $(s_1(\theta'), \bar{\mu}_1) \in \Sigma_{1,\theta'}^k$. Then, $(\theta, s_1, \bar{\mu}_1) \in \Sigma_1^k$. Now the opposite inclusion. Fix $(\theta, s_1, \mu_1) \in \Sigma_1^k$. Thus, $\mu_1 \in \Delta_1$, $\mu_1(\Sigma_2^{k-1}) = 1$, and $(s_1(\theta'), \mu_1) \in \Sigma_{1,\theta'}^k$ for each $\theta' \in \Theta$. Then, by the induction hypothesis (iii), $s_1 \in \overline{\Sigma}_1^k$, thus $(\theta, s_1) \in \overline{\Sigma}_{-2}^{k}$. ■

Comparison with Fixed-Equilibrium Rationalizable Outcomes (Sobel et al. 1990)
Now we present Fixed Equilibrium Rationalizable Outcomes of Sobel et al. (1990). To facilitate the comparison with Path rationalizability, we will slightly modify their definition in terms of language and we consider both players at each iteration instead of alternating between them.[26] Fix an equilibrium $(\beta_1^*, \beta_2^*) \in (\Delta(M))^{\Theta} \times (\Delta(A))^{M}$. For each $\theta \in \Theta$, let $M^*(\theta) = \text{supp}\beta_1^*(\cdot|\theta)$, and let $M^* := \cup_{\theta\in\Theta} M^*(\theta)$. Modify the game by substituting the on-path messages M^* with a unique message m^* that terminates the game. Let $\widetilde{M} = M\backslash M^*$ and $\widehat{M} = \widetilde{M} \cup \{m^*\}$. Let $\hat{s}_1^*$ be the choice function $\hat{s}_1 \in \widehat{M}^{\Theta}$ such that $\hat{s}_1(\theta) = m^*$ for each $\theta \in \Theta$. Finally, for each $\theta \in \Theta$, let $u_1(\theta, m^*, \cdot) = u_1(\theta, m, \beta_2^*)$ for any $m \in M^*(\theta)$ (by equilibrium, every $m \in M^*(\theta)$ gives the same expected payoff). The usual abuse of notation for conditioning on m applies.

Definition 5 *Consider the following reduction procedure.*

Step 0 *Let* $\widehat{\Sigma}_1^0 = \widehat{M}^{\Theta}$ *and* $\widehat{\Sigma}_2^0 = A^{\widetilde{M}}$.

Step n *For each* $\hat{s}_1 \in \widehat{\Sigma}_1^{n-1}$, *let* $\hat{s}_1 \in \widehat{\Sigma}_1^n$ *if there exists* $\mu_1 \in \Delta(A^{\widetilde{M}})$ *such that:*

S1^f $\mu_1(\widehat{\Sigma}_2^{n-1}) = 1$;

S2^f *for each* $\theta \in \Theta$ *and* $m \in \widehat{M}$,

$$u_1(\theta, \hat{s}_1(\theta), \mu_1) \geq u_1(\theta, m, \mu_1).$$

For each $\hat{s}_2 \in \widehat{\Sigma}_2^{n-1}$, *let* $\hat{s}_2 \in \widehat{\Sigma}_2^n$ *if, for each* $m \in \widetilde{M}$ *with* $\widehat{\Sigma}_1^{n-1} \cap \widehat{M}^{\Theta}(m) \neq \emptyset$, *there exists* $\eta \in \Delta(\widehat{M}^{\Theta})$ *such that:*

R1^f $\eta\left(\widehat{\Sigma}_1^{n-1}\right) = 1$;

R2^f $\eta\left(\widehat{M}^{\Theta}(m)\right) > 0$;

R3^f *for every* $a \in A$,

$$u_2(\text{marg}_{\Theta}((p\times\eta)|m), m, \hat{s}_2(m)) \geq u_2(\text{marg}_{\Theta}((p\times\eta)|m), m, a).$$

[26] For simplicity, we also maintain the assumption that the sender has the same available messages after every chance move, although this is not assumed by Sobel et al (1990).

Say that (β_1^*, β_2^*) determines a fixed-equilibrium rationalizable outcome if $\hat{s}_1^* \in \cap_{n>0}\widehat{\Sigma}_1^n$.

FERO is weaker than Path-rationalizability with first-order independence because it only requires the belief of the receiver after each message to be derived from a product measure over chance moves and choice functions — an assumption of *structural consistency* (Kreps and Wilson, 1982) — without any coherency rule between the beliefs after different messages. In other words, there is no analog of Remark 1 for FERO. The following example is used by Sobel et al (1990) to show that FERO is weaker than co-divinity,.a weakening of divinity that allows for mixes of best replies of the receiver to the different beliefs instead of just mixed best replies to a unique belief. While this example shows that FERO, like us, considers the possibility that some "types" would have preferred a different deviation than the observed one, it also highlights how FERO does not require coherence of beliefs after different deviations.

Example 4 (Sobel et al. (1990, page 321)) *Consider the following game,[27] with prior $p(\theta) = 1/4$.*

m_1	a_1	a_2	a_3
θ	$3, -1$	$2, 2$	$-2, 1$
θ'	$6, 2$	$3, -1$	$-1, 1$

m_2	a_1	a_2	a_3
θ	$4, -1$	$1, 2$	$-2, 1$
θ'	$5, 2$	$4, -1$	$-1, 1$

m_3	a_1	a_2	a_3
θ	$0, 0$	$0, 0$	$0, 0$
θ'	$0, 0$	$0, 0$	$0, 0$

Consider a Bayes-Nash equilibrium where the sender always chooses m_3 and the receiver responds to a deviation with a_3.

FERO goes as follows. Let $m^* = m_3$. All choice functions of the receiver survive the first step. For the sender, we have

$$\widehat{\Sigma}_1^1 = \widehat{M}^{\Theta} \setminus \{m_1.m^*, m_2.m^*\}.$$

Strategies $m_1.m^*, m_2.m^*$ cannot be justified because for every belief $\mu_1 \in \Delta(A^{\widetilde{M}})$, if m^* is optimal after θ', it is strictly optimal after θ. Strategy $m_1.m_2$ best replies to $a_2.a_2$ and strategy $m_2.m_1$ best replies to $a_1.a_1$. It is easy to see that all other strategies are optimal under some belief. All strategies of the receiver survive the second step. Fix $m, m' \in \{m_1, m_2\}$ with $m \neq m'$. After m, the receiver may derive his belief about the chance move from a Dirac on m, m', from a Dirac on $m'.m$, or from $\eta \in \Delta(\widehat{M}^{\Theta})$ with $\eta(m.m') = \eta(m.m) = 1/2$. The first belief justifies a_2, the second justifies a_1, the third justifies a_3 (the induced belief is $p'(\theta) = 2/5$). Note that strategy $a_3.a_3$ can only be justified by two different theories about the sender's strategy after m_1 and m_2, whereby after m_1 the sender is supposed to choose m_2 after θ' more frequently than after θ, and after m_2 to choose m_1 more frequently after θ' than after θ. Given that no strategy is eliminated at the second step, we get $m^*.m^* \in \widehat{\Sigma}_1^\infty$: the equilibrium outcome is FERO.

Now we move to Path-rationalizability with first-order independence. The first step is identical to FERO:

$$\overline{\Sigma}_1^1 = M^{\Theta} \setminus \{m_1.m_3, m_2.m_3\}.$$

[27] Sobel et al. use different labels for the receiver's actions after different messages; we adopt the same labels for coherence with our notation.

We are going to show that no $s_2 \in A^M$ with $s_2(m_1) = s_2(m_2) = a_3$ survives the second step of the receiver; then, for every μ_1 with $\mu_1(\overline{\Sigma}_2^2) = 1$, m_3 is suboptimal after θ', and thus $\overline{\Sigma}_2^4 = \emptyset$. Fix any $\hat{\mu}_2 \in \hat{\Delta}_2$ that satisfies R1' at step 2. Let $C = \Theta \times \left(M^\Theta(m_1) \cup M^\Theta(m_2)\right)$. By D2' and R1', $\hat{\mu}_2(\cdot|C) = p \times \eta$ for some η with $\eta(m_1.m_3) = \eta(m_2.m_3) = 0$. Suppose first that $\eta(m_1.m_2) + \eta(m_2.m_1) > 0$. Then, the receiver must update $\mu_2(\cdot|C)$ both after m_1 and m_2. Then, if $\eta(m_1.m_2) \geq \eta(m_2.m_1)$, a_3 is suboptimal after m_2, and if $\eta(m_1.m_2) \leq \eta(m_2.m_1)$, a_3 is suboptimal after m_1 — in both cases, after the corresponding message, the receiver does not raise the probability of θ with respect to $p(\theta)$. Suppose now that $\eta(m_1.m_2) = \eta(m_2.m_1) = 0$. Still, either after m_1 or after m_2 (or both) the receiver's must update $\mu_2(\cdot|C)$, but then he cannot assign to θ probability higher than the prior, therefore a_3 is suboptimal. $\triangle$

To conclude, we formalize the observation that our solution concept is indeed stronger than FERO.

Proposition 4 *Fix an equilibrium (β_1^*, β_2^*) and let $\left(\nu^\theta\right)_{\theta\in\Theta} = (\beta_1^*(\cdot|\theta))_{\theta\in\Theta}$, $\left(\nu^m\right)_{m\in M} = \left(\beta_2^*(\cdot|m)\right)_{m\in M}$. If $\overline{\Sigma}_1^\infty \neq \emptyset$, then (β_1^*, β_2^*) determines a fixed-equilibrium rationalizable outcome.*

Proof. Let φ be the map that associates each $m \in \widetilde{M}$ with itself and each $m \in M^*$ with m^*. Let ς be the map that associates each $s_1 = (s_1(\theta))_{\theta\in\Theta}$ with $\varsigma(s_1) = (\varphi(s_1(\theta)))_{\theta\in\Theta} \in \widehat{M}^\Theta$. Let ϱ be the map that associates each $s_2 = (s_2(m))_{m\in M} \in M^\Theta$ with $\varrho(s_2) = \left(s_2(m)\right)_{m\in\widetilde{M}} \in A^{\widetilde{M}}$. Suppose by contraposition that $\hat{s}_1^* \notin \widehat{\Sigma}_1^\infty$. Then, by finiteness of the game there exists $n \in \mathbb{N}$ such that $\hat{s}_1^* \notin \widehat{\Sigma}_1^n$, and let k be the smallest of such n's. If $\widehat{\Sigma}_1^k \supseteq \varsigma(\overline{\Sigma}_1^k)$, $\hat{s}_1^* \notin \widehat{\Sigma}_1^k$ implies that for every $s_1 \in \overline{\Sigma}_1^k$, there exists $\theta \in \Theta$ such that $s_1(\theta) \notin M^*$. But then, there is no $\hat{\mu}_2 \in \hat{\Delta}_2$ with $\hat{\mu}_2(\overline{\Sigma}_{-2}^k|\varnothing) = 1$, thus $\overline{\Sigma}_2^{k+1} = \emptyset$, and then $\overline{\Sigma}_1^{k+2} = \emptyset$, completing the proof. So there only remains to show that $\widehat{\Sigma}_1^k \supseteq \varsigma(\overline{\Sigma}_1^k)$.

Fix $n = 0, ..., k-1$ and assume by way of induction that $\widehat{\Sigma}_1^n \supseteq \varsigma(\overline{\Sigma}_1^n)$ and $\widehat{\Sigma}_2^n \supseteq \varrho(\overline{\Sigma}_2^n)$. (The basis step is given by $\varsigma(M^\Theta) = \widehat{M}^\Theta$ and $\varrho(A^M) = A^{\widetilde{M}}$.)

Fix $s_2 \in \overline{\Sigma}_2^{n+1}$. Then, by R1' and R2', there exists $\hat{\mu}_2 \in \hat{\Delta}_2$ such that $\hat{\mu}_2(\overline{\Sigma}_{-2}^n|\Theta \times M^\Theta(m)) = 1$ for every $m \in M$ with $\overline{\Sigma}_1^n \cap M^\Theta(m) \neq \emptyset$, and

$$u_2(\mathrm{marg}_\Theta \hat{\mu}_2(\cdot|m), m, s_2(m)) \geq u_2(\mathrm{marg}_\Theta \hat{\mu}_2(\cdot|m), m, a) \tag{10}$$

for each $m \in M$ and every $a \in A$. Fix $m \in \widetilde{M}$. By D2', we have $\mu_2(\cdot|\Theta \times M^\Theta(m)) = p \times \eta^m$ for some $\eta^m \in \Delta(M^\Theta)$. Define $\widehat{\eta}^m \in \Delta(\widehat{M}^\Theta)$ as follows: for each $\hat{s}_1 \in \widehat{M}^\Theta$, let $\widehat{\eta}^m(\hat{s}_1) = \eta^m(\varsigma^{-1}(\hat{s}_1))$. Thus, by $\eta^m(\overline{\Sigma}_1^n) = 1$ and the induction hypothesis, $\widehat{\eta}^m(\widehat{\Sigma}_1^n) = 1$ (R1^f). For each $s_1 \in M^\Theta$ and $\theta \in \Theta$, we have $\varsigma(s_1)(\theta) = m$ if and only if $s_1(\theta) = m$. Therefore, $\widehat{\eta}^m(\widehat{M}^\Theta(m)) > 0$ (R2^f) and

$$\mathrm{marg}_\Theta((p \times \eta^m)\,|m)) = \mathrm{marg}_\Theta((p \times \widehat{\eta}^m)\,|m)). \tag{11}$$

Therefore, by (10) and $\varrho(s_2)(m) = s_2(m)$, $\varrho(s_2)$ and $\widehat{\eta}^m$ satisfy R3^f. By the induction hypothesis, $\varrho(s_2) \in \widehat{\Sigma}_2^n$. Thus, $\varrho(s_2) \in \widehat{\Sigma}_2^{n+1}$. Hence, $\widehat{\Sigma}_2^{n+1} \supseteq \varrho(\overline{\Sigma}_2^{n+1})$.

Now fix $s_1 \in \overline{\Sigma}_1^{n+1}$. By S1' and S2', there exists $\mu_1 \in \hat{\Delta}_1$ with $\mu_1(\overline{\Sigma}_2^n) = 1$ that justifies s_1. Define $\hat{\mu}_1 \in \Delta(A^{\widetilde{M}})$ as $\hat{\mu}_1(\hat{s}_2) = \hat{\mu}_1(\varrho^{-1}(\hat{s}_2))$ for each $\hat{s}_2 \in A^{\widetilde{M}}$. By the induction hypothesis, $\hat{\mu}_1(\widehat{\Sigma}_2^n) = 1$, thus $\hat{\mu}_1$ satisfies S1^f. The sender expects the same payoff under $\hat{\mu}_1$ and μ_1 after each $m \in \widetilde{M}$, and the equilibrium payoff after m^* or each $m \in M^*$. So, $\hat{\mu}_1$ justifies every $\hat{s}_1 \in \varsigma(s_1)$ (S2^f). By the induction hypothesis, $\varsigma(s_1) \subseteq \widehat{\Sigma}_1^n$. Thus, $\varsigma(s_1) \subseteq \widehat{\Sigma}_1^{n+1}$. Hence, $\widehat{\Sigma}_1^{n+1} \supseteq \varsigma(\overline{\Sigma}_1^{n+1})$. ■

5.2 Omitted proofs

Proof of Lemma 1. Necessity. Since $\Sigma_2^n \neq \emptyset$, there exists $\mu_2 \in \Delta_2$ such that $\mu_2(\Sigma_1^{n-1}|\varnothing) = 1$. For any $(s_1, \bar{\mu}_1) \in M^\Theta \times \Delta(A^M)$ such that $\mu_2(\Theta \times \{(s_1, \bar{\mu}_1)\}\,|\varnothing) > 0$, by B2 we have $\mu_2((\theta, s_1, \bar{\mu}_1)|\varnothing) > 0$ for every $\theta \in \Theta$. So, to satisfy B3, we need $\nu^\theta(s_1(\theta)) > 0$. This yields 2. We prove 1 by contraposition: if some $m \in M^*(\theta)$ were to give θ a strictly lower expected payoff than some other $m' \in M^*(\theta)$ under $(\nu^m)_{m\in M^*}$, we would have $m \notin \mathrm{Proj}_M \Sigma_{1,\theta}^1$, and therefore we would get $\Sigma_2^2 = \emptyset$.

Sufficiency. Let $S_1^* = \times_{\theta\in\Theta} M^*(\theta)$. Define $\eta \in \Delta(M^\Theta)$ as follows:

$$\begin{array}{rcll} \forall s_1 & \in & S_1^*, & \eta(s_1) = \prod_{\theta\in\Theta} \nu^\theta(s_1(\theta)); \\ \forall s_1 & \notin & S_1^*, & \eta(s_1) = 0. \end{array}$$

Let $\nu^* = p \times \eta \times \delta_{\mu_1}$, where δ_{μ_1} denotes the Dirac measure supported by $\bar{\mu}_1$. We have $\nu^* \in P_2^*$ because $\eta(M^\Theta(\theta, m)) = \nu^\theta(m)$ for each $(\theta, m) \in \Theta \times M$. By 1 and 2, we have $(\theta, s_1, \bar{\mu}_1) \in \Sigma_1^{n-1}$ for each $\theta \in \Theta$ and $s_1 \in S_1^*$. Hence, $\nu^*(\Sigma_1^{n-1}) = 1$. Hence, letting $\mu_2(\cdot|\varnothing) = \nu^*$, we can construct $\mu_2 \in \Delta_2$ that satisfies R1. Thus, $\Sigma_2^n \neq \emptyset$. ■

Proof of Proposition 1. By Lemma 1, there exists $\bar{\mu}_1 \in \Delta_1$ such that $\bar{\mu}_1\left(\Sigma_2^1\right) = 1$, and for each $(\theta, m) \in \Theta \times M$ with $m \in M^*(\theta)$,

$$m \in \arg\max_{m'\in M} u_1(\theta, m', \bar{\mu}_1).$$

Consider the profile of behavioral strategies $(\beta_1, \beta_2) \in (\Delta(M))^\Theta \times (\Delta(A))^M$ defined by:

1. $\beta_1(\cdot|\theta) = \nu^\theta$ for each $\theta \in \Theta$;
2. $\beta_2(\cdot|m) = \nu^m$ for each $m \in M^*$;
3. for each $m \in M\backslash M^*$ and $a \in A$,

$$\beta_2(a|m) = \bar{\mu}_1\left(\left\{s_2 \in A^M | s_2(m) = a\right\}\right).$$

For each $(\theta, m) \in \Theta \times M^*$ with $\beta_1(m|\theta) > 0$, i.e., $\nu^\theta(m) > 0$, m is optimal against β_2 because β_2 and $\bar{\mu}_1$ induce the same belief after every $m' \in M$. Moreover, β_1 induces belief p^m after every $m \in M^*$, and by assumption the actions in the support of $\beta_2(\cdot|m) = \nu^m$ are optimal under p^m. Finally, for each $m \in M\backslash M^*$ and $a \in A$ with $\beta_2(a|m) > 0$, we have $\bar{\mu}_1\left(\left\{s_2 \in A^M | s_2(m) = a\right\}\right) > 0$, thus there is $s_2 \in \Sigma_2^1$ such that $s_2(m) = a$. Hence, by R2,

$$\forall a' \in A, \quad u_2\left(\mathrm{marg}_\Theta \mu_2(\cdot|m), m, a\right) \geq u_2\left(\mathrm{marg}_\Theta \mu_2(\cdot|m), m, a'\right).$$

Thus, a is optimal given $\text{marg}_{\Theta}\mu_2(\cdot|m)$. ■

Proof of Proposition 2.

We show the equivalence between the reduction procedure for LCPS and path rationalizability with *first*-order independence; then Proposition 2 follows from Proposition 3.

Assume by way of induction that the two procedures are equivalent at step k. For step $k+1$ of the sender, the equivalence between S1$^{\ell}$,S2$^{\ell}$ and S1',S2' can be seen by inspection of the definitions. Consider step $k+1$ of the receiver.

For every CPS $\hat{\mu}_2 \in \hat{\Delta}_2$ that satisfies R1' one can derive an LCPS $\bar{\mu}_2 \in \Delta_2^{\ell}$ that satisfies R1$^{\ell}$ and the first part of R2$^{\ell}$ with the procedure for Remark 1, stopping at the conditional events that are consistent with step k. Then, every s_2 that satisfies R2' with $\hat{\mu}_2$ satisfies R2$^{\ell}$ with $\bar{\mu}_2$ because the first measure of $\bar{\mu}_2$ consistent with a message m, after conditioning on $\Theta \times M^{\Theta}(m)$, coincides with $\hat{\mu}_2(\cdot|m)$.

For every s_2 that survive step $k+1$, for each $j = 1, ..., k+1$, one can find a LCPS $\bar{\mu}^{2,j} \in \Delta_2^{\ell}$ that satisfies R1$^{\ell}$ and R2$^{\ell}$ at step j with s_2. Given the concatenation $(\bar{\mu}^{2,k+1}, ..., \bar{\mu}^{2,1})$, derive $\hat{\mu}_2 \in \hat{\Delta}_2$, for each conditional event C, from the first measure that gives positive probability to C. From R1$^{\ell}$, it is easy to check that $\hat{\mu}_2$ satisfies R1'. Moreover, for each $m \in M$, $\hat{\mu}_2(\cdot|m)$ coincides with the first measure of the concatenation consistent with m, after conditioning on $\Theta \times M^{\Theta}(m)$. Then, since $s_2(m)$ satisfies R2$^{\ell}$ with it, it also satisfies R2' with $\hat{\mu}_2(\cdot|m)$. ■

References

[1] Banks, J.S. and J. Sobel (1987): "Equilibrium selection in signaling games," *Econometrica,* 55(3), 647-661.

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

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

[4] Battigalli, P. and M. Siniscalchi (2007): "Interactive epistemology in games with payoff uncertainty," *Research in Economics*, 61, 165-184.

[5] Battigalli, P. and A. Prestipino (2013): "Transparent Restrictions on Beliefs and Forward Induction Reasoning in Games with Asymmetric Information," *The B.E. Journal of Theoretical Economics*, 13(1), 79-130.

[6] Battigalli, P., Catonini, E., Manili, J. (2021): "Belief change, rationality, and strategic reasoning in sequential games," Igier Working Paper n. 679, Universita' Bocconi.

[7] Blume, L., Brandenburger, A., Dekel, E. (1991): "Lexicographic probabilities and choice under uncertainty," *Econometrica*, 59, 61-79.

[8] Catonini, E. (2021): "Self-enforcing agreements and forward induction reasoning," *Review of Economic Studies*, **2**, 610-642.

[9] Catonini, E. (2022): "A Dutch book argument for belief consistency," working paper.

[10] Cho, I.K. (1987): "A refinement of sequential equilibrium," *Econometrica*, 55(6), 1367-1389.

[11] Cho, I.K. and D. Kreps (1987): "Signaling Games and Stable Equilibria," *The Quarterly Journal of Economics*, 102(2), 179-221.

[12] Fudenberg, D. and K. He (2018): "Learning and type compatibility in signaling games," *Econometrica*, 86, 1215-1255.

[13] Fudenberg, D. and Y. Kamada (2015): "Rationalizable Partition-Confirmed Equilibrium,". *Theoretical Economics*, 10, 775-806.

[14] Kohlberg, E. and J.-F. Mertens (1986): "On the strategic stability of equilibria," *Econometrica*, 54, 1003-1037.

[15] Kreps, D. and R. Wilson (1982): "Sequential Equilibria," *Econometrica*, 50, 863-894.

[16] Myerson, R. (1986): "Multistage Games with Communication," *Econometrica*, 54, 323-358.

[17] Sobel, J., L. Stole, I. Zapater (1990): "Fixed-Equilibrium Rationalizability in Signaling Games," *Journal of Economic Theory*, 52, 304-331.