

Structural Rationality in Dynamic Games

Marciano Siniscalchi*

May 4, 2017

Abstract

The analysis of dynamic games hinges on assumptions about players' actions and beliefs at information sets that are not actually reached during game play, and that players themselves do not expect to reach. However, it is not obvious how to elicit intended actions and conditional beliefs at such information sets. Hence, key concepts such as sequential rationality, backward induction, and forward induction do not readily translate to testable behavioral assumptions. This paper addresses this concern by introducing a novel optimality criterion, *structural rationality*. In any dynamic game, structural rationality implies sequential rationality. In addition, if players are structurally rational, their intended actions and conditional beliefs can be elicited via the *strategy method* (Selten, 1967). Finally, structural rationality is consistent with experimental evidence indicating that subjects behave differently in the strategic and extensive form, but take the extensive form into account even if they are asked to commit to strategies ahead of time.

Keywords: conditional probability systems, sequential rationality, strategy method.

*Economics Department, Northwestern University, Evanston, IL 60208; marciano@northwestern.edu. Earlier drafts were circulated with the titles 'Behavioral counterfactuals,' 'A revealed-preference theory of strategic counterfactuals,' 'A revealed-preference theory of sequential rationality,' and 'Sequential preferences and sequential rationality.' I thank Amanda Friedenberg and participants at RUD 2011, D-TEA 2013, and many seminar presentations for helpful comments on earlier drafts.

1 Introduction

The analysis of simultaneous-move games is grounded in single-person choice theory. Players are assumed to maximize their expected utility (EU)—a criterion characterized by well-known, testable properties of observed choices (Savage, 1954; Anscombe and Aumann, 1963). Furthermore, beliefs can be elicited via incentive-compatible “side bets” whose outcomes depend upon the strategies of coplayers (Luce and Raiffa, 1957, §13.6).¹ Hence, assumptions about players’ beliefs can be translated into testable restrictions on behavior.

In dynamic games, the central notion of sequential rationality (Kreps and Wilson, 1982) requires that a player’s strategy prescribe an optimal continuation at every information set. A common interpretation of this requirement is that the player is endowed with EU preferences at every information set; a sequentially rational strategy simultaneously maximizes every such preference.² However, this interpretation effectively precludes the elicitation of players’ beliefs at information sets they do not expect to reach, and that indeed are not reached during observed play. This is concerning because key concepts such as backward or forward induction entail specific assumptions on beliefs off the predicted path of play. If such beliefs cannot be elicited, these assumptions cannot be tested.

To illustrate, consider the centipede game in Fig. 1 (Rosenthal, 1981). Suppose that an experimenter wishes to verify whether, per the standard backward-induction argument, Ann believes at the third node (denoted I) that Bob will choose d_2 at the last node. If I were reached during game play, the experimenter could in principle do so by offering Ann side bets on a_2 vs. d_2 . However, if the backward-induction analysis of the game in Fig. 1 is correct, I will

¹The experimental literature illustrates how to implement side-bets in practice: see e.g. Van Huyck, Battalio, and Beil (1990); Nyarko and Schotter (2002); Costa-Gomes and Weizsäcker (2008); Rey-Biel (2009). See also Aumann and Dreze, 2009 and Gilboa and Schmeidler, 2003.

²Kreps and Wilson (1982, p. 863) explicitly suggest this interpretation. Myerson (1979) analyzes “conditional EU maximization” axiomatically.

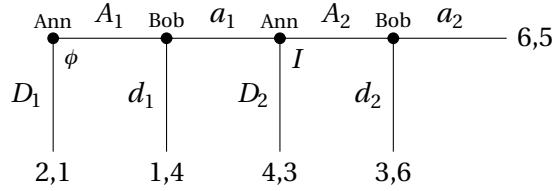


Figure 1: A centipede game.

not be reached—Ann will choose D_1 at the first node. Hence, the experimenter can only elicit Ann’s *prior* beliefs (the beliefs she holds at the initial node ϕ). If Ann assigned positive prior probability to I being reached, then her beliefs at I could be derived from her prior beliefs by updating. However, if—again, consistently with backward induction—Ann assigns probability zero to I being reached, updating is not possible. Hence, if Ann has EU preferences at nodes ϕ and I , her beliefs at I cannot be elicited.³ As a result, it is not actually possible to test experimentally whether Ann’s beliefs at I are consistent with backward induction.

The present paper proposes an alternative perspective on sequential rationality that resolves this difficulty, and thus places the analysis of dynamic games on firm choice-theoretic grounds. Theorem 1 shows that sequential rationality follows from the maximization of a *single* preference relation, which characterizes a player’s ex-ante perspective, but takes into account the information about opponents’ moves she may receive throughout the game. I call this preference relation *structural* because this information is represented by the extensive-form structure of the game. Theorem 2 then shows that, if players are endowed with structural preferences, their prior and conditional beliefs can be elicited in an incentive-compatible way.

In addition, structural rationality (the ex-ante maximization of structural preferences) is consistent with experimental evidence that poses a challenge for the view that players maximize EU preferences at each information set—and, indeed, for sequential rationality itself.

³A further difficulty is that Ann’s own choice of D_1 makes it impossible to observe *any* action by Bob. As discussed in Section 5.2, this implies that even the elicitation of prior beliefs is potentially problematic.

On one hand, experimental findings indicate that subjects behave differently in a dynamic game and in the associated strategic form (Cooper, DeJong, Forsythe, and Ross, 1993; Schotter, Weigelt, and Wilson, 1994; Cooper and Van Huyck, 2003; Huck and Müller, 2005); this is predicted by both sequential and structural rationality. On the other hand, evidence also suggests that qualitatively similar behavior is observed in experiments when subjects play the actual game in extensive form (“direct response”) and when the *strategy method* (Selten, 1967) is adopted: see Brandts and Charness (2011) (a broad meta-analysis) and Fischbacher, Gächter, and Quercia (2012). In the strategy-method implementation of a dynamic game, players simultaneously commit to an extensive-form strategy, which the experimenter then implements. When the game is so implemented, sequential rationality reduces to ex-ante optimality, and hence does not explain why behavior under direct response and in the strategy-method implementation should be consistently similar. Section 5.2 instead shows that, if subjects observe the realized path of play as the experimenter implements their choices, their structural preferences are unchanged, and hence imply the same observed behavior.⁴ Thus, unlike received theories, structural rationality provides a positive rationale for the use of the strategy method in experiments.

This paper is organized as follows. Section 2 introduces the basic notation and defines structural preferences for extensive-form games in which information sets satisfy a regularity condition. While restrictive, this class includes several games of interest in applications and experiments, and permits a more straightforward definition of structural preferences. Section 4 generalizes this definition to arbitrary extensive forms with perfect recall. Section 5.1 contains the main result on structural and sequential rationality. Section 5.2 describes the elicitation procedure in detail, and contains the main result on the incentive-compatible elicitation

⁴ To the best of my knowledge, no other theory of play can accommodate both the evidence on behavior in the strategic-form vs. the extensive form, and the evidence on direct-response vs. strategy-method treatments. In particular, the invariance hypothesis (Kohlberg and Mertens, 1986) predicts that behavior should be the same in all presentations of the game.

of structural preferences. Section 6 discusses alternative notions of preferences (including lexicographic maximization) and other related issues.

Finally, the companion paper Siniscalchi (2016a) provides an axiomatic characterization of structural preferences. Siniscalchi (2016b) instead incorporates structural rationality into the analysis of forward-induction reasoning.

2 Basic notation and definitions

This paper considers dynamic games with imperfect information. The analysis only requires that certain familiar reduced-form objects be defined. The Online Appendix describes how these objects are derived from a complete description of the underlying game, as e.g. in Osborne and Rubinstein (1994, Def. 200.1, pp. 200-201; OR henceforth) Section 6 indicates how to extend the notation to allow for incomplete information as well.

A dynamic game will be represented by a tuple $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$, where:

- N is the set of **players**.
- S_i is the set of **strategies** of player i .
- \mathcal{I}_i is the collection of **information sets** player i ; it is convenient to assume that the **root**, ϕ , is an information set for all players.
- $U_i : \prod_{j \in N} S_j \rightarrow \mathbb{R}$ is the reduced-form **payoff function** for player i .
- For every $i \in N$ and $I \in \mathcal{I}_i$, $S(I)$ is the set of strategy profiles $(s_j)_{j \in N} \in \prod_j S_j$ that **reach** I .

I adopt the usual conventions for product sets: thus, $S_{-i} = \prod_{j \neq i} S_j$ and $S = S_i \times S_{-i}$. I assume that the game has **perfect recall**, as per Def. 203.3 in OR. In particular, this implies that, for every $i \in N$ and $I \in \mathcal{I}_i$, $S(I) = S_i(I) \times S_{-i}(I)$, where $S_i(I) = \text{proj}_{S_i} S(I)$ and $S_{-i}(I) = \text{proj}_{S_{-i}} S(I)$. If

$s_{-i} \in S_{-i}(I)$, I say that s_{-i} **allows** I .⁵ It is also convenient to use the notation

$$S_{-i}(\mathcal{I}_i) = \{S_{-i}(I) : I \in \mathcal{I}_i\}.^6 \quad (1)$$

Finally, for every player $i \in N$ and information set $I \in \mathcal{I}_i$, the set $S(I)$ is required to satisfy *strategic independence* (Mailath, Samuelson, and Swinkels, 1990, Def. 2): for every $s_i, t_i \in S_i(I)$ there is $r_i \in S_i(I)$ such that $U_i(r_i, s_{-i}) = U_i(r_i, t_{-i})$ for all $s_{-i} \in S_{-i}(I)$, and $U_i(r_i, s_{-i}) = U_i(s_i, s_{-i})$ for all $s_{-i} \in S_{-i} \setminus S_{-i}(I)$. Intuitively, r_i is the strategy that coincides with s_i everywhere except at I and all subsequent information sets, where it coincides with t_i : see the Online Appendix, or Theorem 1 in Mailath et al. (1990).

At any information set $I \in \mathcal{I}_i$, player i 's beliefs about the past and future moves of her coplayers are represented by a probability distribution over S_{-i} . These beliefs are conditional upon the (possibly partial) information she has at I about coplayers' previous moves; this information is represented by the event $S_{-i}(I)$. Collectively, player i 's beliefs are required to satisfy the chain rule of conditioning: starting with the prior, player i updates her beliefs in the usual way "whenever possible"—that is, until a zero-probability event occurs. Then, player i formulates a new belief, but from that point on, she again applies the updating formula, until a new, zero-probability event is observed; and so on.

Definition 1 below takes as given an arbitrary collection \mathcal{C}_i of conditioning events. The preceding paragraph suggests the specification $\mathcal{C}_i = S_{-i}(\mathcal{I}_i)$; this is sufficient to define sequential rationality, and is also enough to define structural rationality in the class of games considered in Section 3. The general definition of structural rationality requires a richer set of conditioning events; see Section 4. Myerson (1986) considers the case $\mathcal{C}_i = 2^{S_{-i}} \setminus \{\emptyset\}$.

Definition 1 (Rényi, 1955; Myerson, 1986; Ben-Porath, 1997; Battigalli and Siniscalchi, 1999,

⁵That is: if i 's coplayers follow the profile s_{-i} , I can be reached; whether it is reached depends upon whether or not i plays a strategy in $S_i(I)$.

⁶If one views $S_{-i}(\cdot)$ as a function from \mathcal{I}_i to $2^{S_{-i}}$, then $S_{-i}(\mathcal{I}_i)$ is its range.

2002b) Fix a dynamic game $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$, a player $i \in N$, and a non-empty collection \mathcal{C}_i of non-empty subsets of S_{-i} . A **conditional probability system (CPS)** on (S_{-i}, \mathcal{C}_i) is a collection $\mu_i \equiv (\mu_i(\cdot|F))_{F \in \mathcal{C}_i}$ such that:

(1) for every $F \in \mathcal{C}_i$, $\mu_i(\cdot|F) \in \Delta(S_{-i})$ and $\mu_i(F|F) = 1$;

(2) for every $E \subseteq S_{-i}$ and $F, G \in \mathcal{C}_i$ such that $E \subseteq F \subseteq G$,

$$\mu_i(E|G) = \mu_i(E|F) \cdot \mu_i(F|G). \quad (2)$$

The set of CPS on (S_{-i}, \mathcal{C}_i) is denoted by $\Delta(S_{-i}, \mathcal{C}_i)$. For any probability distribution $\pi \in \Delta(S_{-i})$ and function $a : S_{-i} \rightarrow \mathbb{R}$, let $E_\pi[a] = \sum_{s_{-i} \in S_{-i}} a(s_{-i})\pi(s_{-i})$; when no confusion can arise, I omit the square brackets.

Sequential rationality requires that a strategy be optimal at every information set that it does not preclude: for the rationale behind this definition, see [Reny \(1992\)](#), [Rubinstein \(1991\)](#), and [Battigalli and Siniscalchi \(2002a\)](#).

Definition 2 (Sequential rationality) Fix a dynamic game $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$, a player $i \in N$, and a CPS $\mu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i))$ for player i . Strategy $s_i \in S_i$ is **sequentially rational (given μ)** if, for every $I \in \mathcal{I}_i$ such that $s_i \in S_i(I)$, and all $t_i \in S_i(I)$, $E_{\mu(\cdot|S_{-i}(I))} U_i(s_i, \cdot) \geq E_{\mu(\cdot|S_{-i}(I))} U_i(t_i, \cdot)$.

Finally, two definitions related to CPSs play a key role in the analysis of structural preferences. First, fix an event $F \in \mathcal{C}_i$; by Equation (2), if there is $G \in \mathcal{C}_i$ such that $G \supset F$ and $\mu(F|G) > 0$, then $\mu(\cdot|F)$ is derived from $\mu(\cdot|G)$ by updating. If instead $\mu(F|G) = 0$ for every event $G \in \mathcal{C}_i$ such that $G \supset F$, then $\mu(\cdot|G)$ cannot be derived from any other element of the CPS μ by updating. I call such a belief *basic* to emphasize this fact:

Definition 3 Consider a dynamic game the tuple $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$, a player $i \in N$, and a CPS $\mu \in \Delta(S_{-i}, \mathcal{C}_i)$. An event $F \in \mathcal{C}_i$ is **basic** if, for all $G \in \mathcal{C}_i$, $G \supset F$ implies $\mu(F|G) = 0$. If $F \in \mathcal{C}_i$ is basic, then $\mu(\cdot|F)$ is a **basic belief** of player i .

The prior belief $\mu(\cdot|S_{-i})$ is always basic. The collection of basic beliefs is a “sufficient statistic” for the entire CPS μ : any other conditional belief can be derived by updating some basic belief.

Second, a CPS also conveys information about the relative “infinitesimal” likelihood of events. Elaborating upon the classic example by [Blume, Brandenburger, and Dekel \(1991a\)](#), suppose a die is thrown, and allow for the possibility that it lands either on a face or on an edge. Let E (resp. O) be the event that the die lands on an edge, one of whose adjoining faces is even (resp. odd). Suppose that the individual assigns positive probability to O given E . This suggests that she does deem E *not* infinitely less likely than O . (For instance, she may believe that the die is just as (un)likely to land on any one edge.) The following definition augments this intuition by requiring that likelihood comparisons be transitive.

Definition 4 Consider a dynamic game the tuple $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$, a player $i \in N$, a CPS $\mu \in \Delta(S_{-i}, \mathcal{C}_i)$, and two events $F, G \in \mathcal{C}_i$. Then F is **not infinitely less likely than G given μ** ($F \geq^\mu G$) if there is an ordered list $F_1, \dots, F_L \in \mathcal{C}_i$ such that $F_1 = G$, $F_L = F$, and $\mu(F_{\ell+1}|F_\ell) > 0$ for all $\ell = 1, \dots, L$.

The relation \geq^μ is a preorder (i.e., reflexive and transitive), but in general not complete: see [Example 3](#) in [Section 3](#). Its strict (i.e., asymmetric) part $>^\mu$ is defined as usual by letting $F >^\mu G$ iff $F \geq^\mu G$ and not $G \geq^\mu F$.

3 Structural preferences under nested strategic information

This section restricts attention to games that satisfy a convenient regularity condition.

Definition 5 A dynamic game $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$ has **nested strategic information** if

$$\forall i \in N, I, J \in \mathcal{I}_i: \quad \text{either } S_{-i}(I) \cap S_{-i}(J) = \emptyset \text{ or } S_{-i}(I) \subseteq S_{-i}(J) \text{ or } S_{-i}(J) \subseteq S_{-i}(I). \quad (3)$$

In a game with nested strategic information, either *every* strategy profile that allows I (resp. J) set also allows J (resp. I), or *no* strategy profile allows both I and J . All signalling games, and, more broadly, all games in which each player moves only once on each path of play, have nested strategic information. So do centipede game forms, and ascending-clock auctions.

While Definition 5 rules out many games of interest (see Section 4), it does allow for a simple definition of structural preferences and structural rationality.

Definition 6 (Structural preferences with nested strategic information) Fix a dynamic game with nested strategic information $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N})$, a player $i \in N$, a CPS $\mu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i))$, and strategies $s_i, t_i \in S_i$. Then s_i is (weakly) **structurally preferred** to t_i ($s_i \succsim^\mu t_i$) iff, for any basic event $F \in S_{-i}(\mathcal{I}_i)$ such that $E_{\mu(\cdot|F)} U_i(s_i, \cdot) < E_{\mu(\cdot|F)} U_i(t_i, \cdot)$, there is a basic event $G \in S_{-i}(\mathcal{I}_i)$, with $G \succ^\mu F$, such that $E_{\mu(\cdot|G)} U_i(s_i, \cdot) > E_{\mu(\cdot|G)} U_i(t_i, \cdot)$.

In words, $s_i \succsim^\mu t_i$ means that s_i is infinitely more likely to be better than t_i than to be worse than t_i : if t_i yields a strictly higher payoff given some basic belief $\mu(\cdot|F)$, then s_i must yield a strictly higher payoff given some basic belief $\mu(\cdot|G)$ associated with an infinitely more likely conditioning event.

Strict preference, denoted “ $s_i \succ^\mu t_i$,” is defined as “ $s_i \succsim^\mu t_i$ and not $t_i \succsim^\mu s_i$,” indifference, denoted “ $s_i \sim^\mu t_i$,” is defined as “both $s_i \succsim^\mu t_i$ and $t_i \succsim^\mu s_i$.” A strategy $s_i \in S_i$ is **structurally rational** if there is no $t_i \in S_i$ with $t_i \succ^\mu s_i$.⁷ Structural preferences are transitive (for a proof, see Siniscalchi, 2016a, Appendix B); hence, every finite game admits a structurally rational strategy for every CPS.

In games with nested strategic information, the likelihood ranking of basic events reduces to set inclusion (see Appendix C for a proof):

Remark 1 For any game with nested strategic information $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N})$, player $i \in N$, CPS $\mu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i))$, and basic events $F, G \in S_{-i}(\mathcal{I}_i)$: $G \succ^\mu F$ if and only if $G \supset F$.

This result simplifies the analysis, as the examples in this Section demonstrate.

Example 1 Consider the centipede game in the Introduction, and continue to assume that Ann’s beliefs are consistent with backward induction: $\mu(\{d_1 d_2\} | S_b(\phi)) = 1$ and $\mu(\{a_1 d_2\} | S_b(I)) =$

⁷ That is s_i is *maximal* with respect to \succsim^μ . With complete preferences, maximality coincides with optimality (s_i is at least as good as any other strategy). However, as Example 3 shows, structural preferences may be incomplete.

1. By Definition 3, $S_b = S_b(\phi)$ and $S_b(I)$ are both basic events. Figure 2 reproduces the game for convenience, and indicates the expected payoff of each of Ann's strategies with respect to her beliefs at ϕ and I . Since this is the first application of Definition 6, I analyze this example

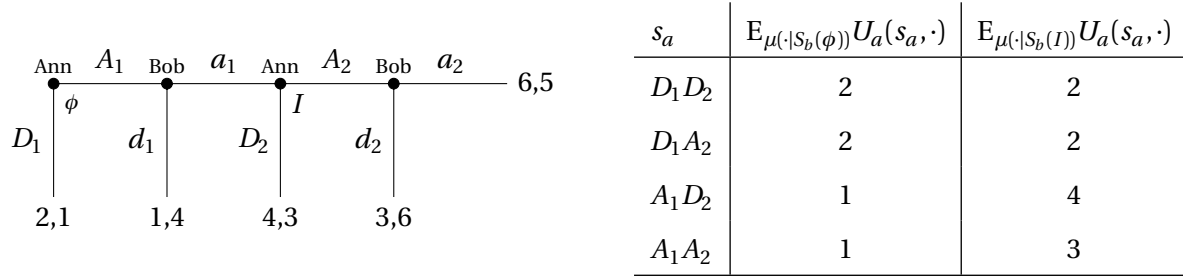


Figure 2: A four-legged centipede game. Ann's CPS: $\mu(\{d_1 d_2\}|S_b(\phi)) = \mu(\{a_2 d_2\}|S_b(I)) = 1$

in detail. Note first that, since the strategies D_1D_2 and D_1A_2 are realization-equivalent, they yield the same expected payoffs given Ann's beliefs at ϕ and I . Hence, they are equivalent from the perspective of structural preferences. This is a general property of Definition 6. For short, denote either one of D_1D_2, D_1A_2 simply as D_1 .

Second, while A_1D_2 and A_1A_2 yield a strictly higher expected payoff given $\mu(\cdot|S_b(I))$ than D_1 , they yield a strictly lower expected payoff given the prior belief $\mu(\cdot|S_b(\phi))$. Since $S_b(\phi) = S_b \supset \{a_1 d_2, a_1 a_2\} = S_b(I)$, Definition 6 and Remark 1 imply that $D_1 \succ^\mu A_1D_2$ and $D_1 \succ^\mu A_1A_2$. Furthermore, since there is no $J \in \mathcal{J}_a$ such that $S_b(J) \supset S_b = S_b(\phi)$, the fact that $E_{\mu(\cdot|S_b(\phi))}U_a(D_1, \cdot) > E_{\mu(\cdot|S_b(\phi))}U_a(A_1D_2, \cdot)$ implies that it is *not* the case that $A_1D_2 \succ^\mu D_1$; thus, $D_1 \succ^\mu A_1D_2$. Similarly, $D_1 \succ^\mu A_1A_2$. This is also a general property of Definition 6: if strategy s_i has strictly greater ex-ante expected payoff than strategy t_i , it is structurally strictly preferred to it.

Definition 6 also implies that $A_1D_2 \succ^\mu A_1A_2$: there is no basic event $S_b(J) \in S_b(\mathcal{J}_a)$ such that $E_{\mu(\cdot|S_b(J))}U_a(A_1D_2, \cdot) < E_{\mu(\cdot|S_b(J))}U_a(A_1A_2, \cdot)$, so $A_1D_2 \succ^\mu A_1A_2$; furthermore, since $E_{\mu(\cdot|S_b(I))}U_a(A_1D_2, \cdot) > E_{\mu(\cdot|S_b(I))}U_a(A_1A_2, \cdot)$, it is not the case that $A_1A_2 \succ^\mu A_1D_2$.

Thus, D_1 is maximal (indeed, optimal) for \succ^μ , and it is also sequentially rational per Definition 2. This is consistent with Theorem 1 in Section 5.1. \square

Example 1 illustrates one simple manifestation of the fact that Definition 6 reflects an *ex-ante perspective*. Ann's prior $\mu(\cdot|S_b(\phi))$ is treated as her primary hypothesis about Bob's play: if s_a yields strictly higher payoff than t_b ex-ante, it is strictly structurally preferred, regardless of the expected payoff given $\mu(\cdot|S_b(I))$. The latter is only used to break ex-ante ties, and is thus regarded as an alternative, secondary hypothesis about Bob's play. (The discussion following Example 2 below highlights another, subtler aspect of the ex-ante nature of Definition 6.)

At the same time, since it does also take Ann's belief given $S_b(I)$ into account, Definition 6 is *forward-looking*: it allows for the possibility that observed play *might* contradict Ann's prior. This forward-looking nature leads to the conclusion that Ann prefers A_1D_2 to A_1D_1 ex-ante.

Finally, structural preferences reflect the *extensive form* of the game. First of all, a player's CPS is explicitly defined with reference to the collection \mathcal{S}_i of her information sets. Second, Definition 6 makes explicit reference to the basic events in $S_{-i}(\mathcal{S}_i)$. Third, the order in which the associated basic beliefs are used to break ties is also determined by her CPS μ . Indeed, under nested strategic information, this ordering coincides with set inclusion of the corresponding basic events, which reflects the extensive-form structure of the game.

This dependence upon the extensive form of the game differentiates structural preferences from lexicographic expected-utility preferences (Blume et al., 1991a; Blume, Brandenburger, and Dekel, 1991b). Fix a player i and an ordered list $\sigma = (p_1, \dots, p_L)$ of probabilities over S_{-i} (henceforth, a lexicographic probability system, or LPS). Strategy s_i is lexicographically (weakly) preferred to t_i if, for every $k \in \{1, \dots, L\}$ such that $E_{p_k} U_i(s_i, \cdot) < E_{p_k} U_i(t_i, \cdot)$ there is $\ell < k$ such that $E_{p_\ell} U_i(s_i, \cdot) > E_{p_\ell} U_i(t_i, \cdot)$. While there are analogies with Definition 6, the key difference is that the LPS σ , and the ordering of its component probabilities p_1, \dots, p_L , are entirely unrelated to the extensive form.⁸ In addition to these substantive differences, a formal

⁸Lexicographic EU maximization was introduced into game theory in order to study *strategic-form* refinements (Blume et al., 1991b). In addition, lexicographic EU maximization with respect to a full-support LPS implies a strong form of invariance (see footnote 4): in single-person decision problems, it is equivalent to sequential rationality in every tree with the same matrix form (Brandenburger, 2007). This further indicates that

difference is the fact that structural preferences may be incomplete: see Example 3. Section 6 provides a more detailed comparison of lexicographic and structural rationality.

Structural preferences reduce to standard expected-utility maximization in two cases: when the game has simultaneous moves, and when every information set of player i has positive prior probability (because in this case only the prior $\mu(\cdot|S_{-i})$ is a basic belief). Outside of these special cases, structural preferences refine ex-ante EU maximization. This immediately delivers the first prediction anticipated in the Introduction: if a player has structural preferences, her behavior will, in general, differ in a dynamic game and in the associated strategic form.

The following examples illustrate other features of structural preferences.

Example 2 (Refining sequential rationality) The ex-ante perspective that characterizes structural preferences can sometimes lead to a refinement of the predictions of sequential rationality. Indeed, the analysis of elicitation in Section 5.2 builds upon this. Figure 3 provides a simpler illustration.

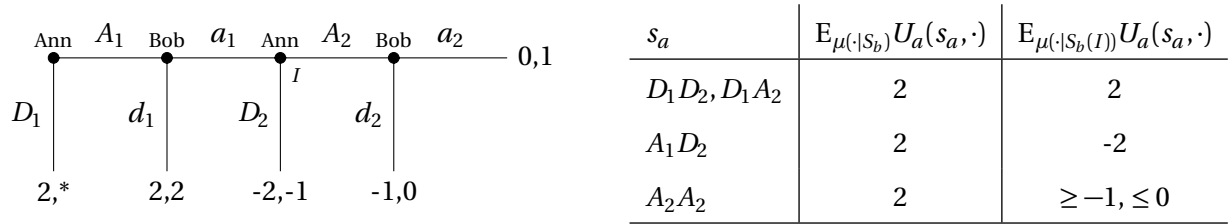


Figure 3: Structural and sequential rationality. Ann's CPS: $\mu(S_b(I)|S_b(\phi)) = 0$

Again, take the perspective of Ann, and assume that her CPS μ satisfies $\mu(S_b(I)|S_b(\phi)) = 0$ and (as required by Definition 1) $\mu(S_b(I)|S_b(I)) = 1$. Both events of Ann are basic. The game has nested strategic information: $S_b(\phi) = S_b \supset \{a_1 d_2, a_1 a_2\} = S_b(I)$. As in Example 1, I identify realization-equivalent strategies: thus, D_1 stands for either $D_1 D_2$ or $D_2 A_2$.

lexicographic maximization disregards the given extensive form, whereas structural rationality is built upon it.

Given the CPS μ , both strategies D_1 and A_1A_2 are sequentially rational. However, $D_1 \succ^\mu A_1D_2$: $E_{\mu(\cdot|S_b(I))}U_a(D_1, \cdot) > E_{\mu(\cdot|S_b(I))}U_a(A_1A_2, \cdot)$, and there is no basic event $S_b(J) \in S_b(\mathcal{I}_i)$ such that $E_{\mu(\cdot|S_b(J))}U_a(D_1, \cdot) < E_{\mu(\cdot|S_b(J))}U_a(A_1A_2, \cdot)$. As a result, D_1 is the only structurally rational strategy for Ann, given the CPS μ .

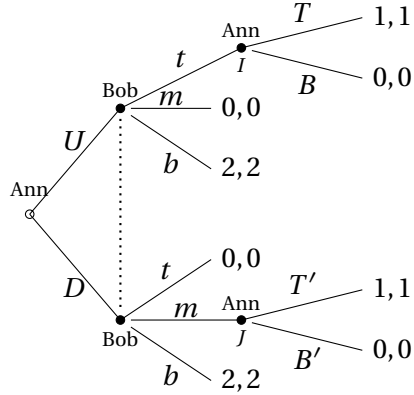
In addition, D_1 is the only structurally rational strategy for *any* CPS ν for Ann. Since the only assumption about μ was that $\mu(S_b(I)|S_b(\phi)) = 0$, it is enough to consider the case of a CPS ν with $\nu(S_b(I)|S_b(\phi)) > 0$. In this case, only $S_b = S_{-a}(\phi)$ is basic, so structural rationality coincides with ex-ante payoff maximization. Furthermore, since $\nu(\cdot|S_b(\phi))$ assigns positive probability to at least one of a_1d_2, a_1a_2 , the ex-ante expected payoff from A_1A_2 (or A_1D_2) is strictly less than 2. Hence, D_1 is also the unique structurally rational strategy for such a CPS ν .

Refinements of sequential equilibrium, such as trembling-hand perfection (Selten, 1975) or properness (Myerson, 1978) also rule out A_1A_2 . Structural rationality delivers this conclusion without explicitly perturbing beliefs. \square

Example 2 also emphasizes the interpretation of Definition 6 as an *ex-ante* ranking. Strategy D_1 does not allow information set I to be reached; yet, Definition 6 requires comparing the expected payoffs of D_1 and A_1A_2 given $\mu(\cdot|S_b(I))$. This makes sense from an ex-ante perspective. At the initial node, Ann still must choose between the actions D_1 and A_1 ; to do so, she takes into account the basic belief $\mu(\cdot|S_b(I))$, albeit only to break ex-ante ties. By way of contrast, from an *interim* perspective—that is, once I is actually reached—it is not meaningful to compare the payoff of A_1A_2 with the payoff that D_1 would have achieved instead. Sequential rationality does not require this comparison because it reflects the *interim* perspective.

Example 3 (Incompleteness) Consider the game in Figure 4, which satisfies nested strategic information. Take Ann's perspective, and assume that her CPS μ satisfies $\mu(\{b\}|S_b(\phi)) = 1$ (the other conditional probabilities are pinned down by Definition 1).

All conditional beliefs of Bob are basic. Moreover, $S_b(I) = \{t\}$ and $S_b(J) = \{m\}$ are disjoint, hence not ordered by relative likelihood (Remark 1). Definition 6 implies that UTT' and



s_a	$E_{\mu(\cdot S_b)}U_a(s_a, \cdot)$	$E_{\mu(\cdot S_b(I))}U_a(s_a, \cdot)$	$E_{\mu(\cdot S_b(J))}U_a(s_a, \cdot)$
UTT', UTB'	2	1	0
UBT', UBB'	2	0	0
DTT', DBT'	2	0	1
DTB', DBB'	2	0	0

Figure 4: Incomplete structural preferences. Ann's CPS: $\mu(\{b\}|S_b(\phi)) = 1$

UTB' are strictly preferred to UBT', UBB' ; similarly, DTT', DBT' are strictly preferred to DTB', DBB' . However, there are no further strict rankings of strategies: UTT', UTB', DTT' and DBT' are all structurally rational.⁹ For instance, UTT' and DTT' are incomparable: UTT' yields a strictly higher expected payoff given $\mu(\cdot|S_b(I))$ than DTT' , but DTT' does strictly better given $\mu(\cdot|S_b(J))$. Notice that UTT', UTB', DTT', DBT' are also the sequentially rational strategies given μ . \square

In Example 3, preferences are incomplete because the ranking of basic events is itself incomplete: Ann's CPS does not indicate whether or not one of the events $S_a(I)$ and $S_a(J)$ is infinitely more likely than the other. However, this incompleteness is inconsequential as far as the connection between structural and sequential rationality is concerned.

4 Structural preferences for general dynamic games

While convenient, the assumption of nested strategic information rules out several games of economic interest; notable examples include English (rather than ascending-clock) auctions,

⁹Note that each of these strategies is *maximal* (i.e., undominated for structural preferences), but not *optimal* (i.e., structurally preferred to all other strategies): cf. footnote 7. Thus, this game has no optimal strategies.

alternating-offer bargaining, and the chain-store game.¹⁰ Unfortunately, without nested strategic information, a strategy may be maximal in the order formalized by Definition 6, and yet fail to be sequentially rational.

Example 4 Consider the “signal-choice” game in Figure 5. Ann and Bob choose an action simultaneously. If Bob chooses o , the game ends. Otherwise, Ann’s action determines what she learns about Bob’s action. This game does not have nested strategic information: $S_b(I) = \{t, m\}$ and $S_b(J) = \{m, b\}$, so $S_b(I) \cap S_b(J) \neq \emptyset$ but $S_b(I)$ and $S_b(J)$ are not nested. Bob’s payoffs are omitted in Fig. 5 as they are not relevant to the discussion.

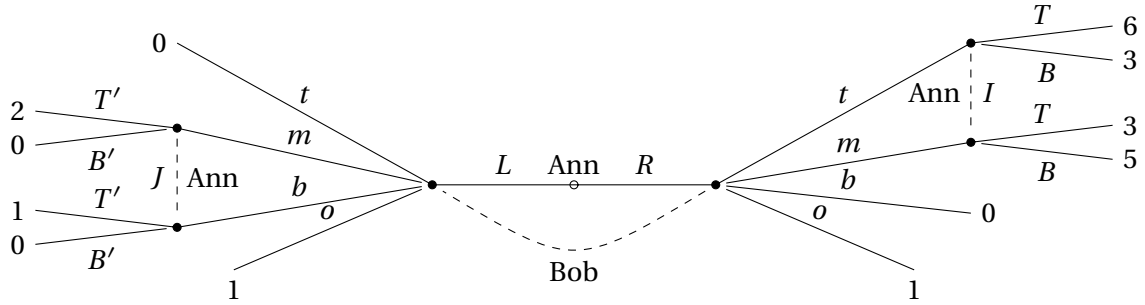


Figure 5: A signal-choice game.

Define Ann’s CPS μ by $\mu(\{o\}|S_b(\phi)) = 1$, $\mu(\{t\}|S_b(I)) = \mu(\{m\}|S_b(I)) = \frac{1}{2}$, and $\mu(\{m\}|S_b(J)) = \mu(\{b\}|S_b(J)) = \frac{1}{2}$. All three elements of Ann’s CPS are basic beliefs; expected payoffs are displayed in Table I.

According to Definition 6, RB is structurally rational given μ . In particular, RT is *not*

¹⁰For example, consider a (discretized, finite-horizon) bargaining game between Ann and Bob. Let I (resp. J) denote the history in which Ann offers to leave the entire pie to Bob (resp. keep the entire pie for herself), and Bob rejects the offer. Any strategy of Bob that always rejects Ann’s initial offer, no matter what it is, allows both I and J : thus, $S_a(I) \cap S_a(J) \neq \emptyset$. However, a strategy of Bob that accepts the offer of the whole pie and rejects a zero share allows J but not I ; and a strategy of Bob that rejects the whole pie but accepts a zero share allows I but not J . Thus, $S_a(I)$ and $S_a(J)$ are not nested.

s_a	$E_{\mu(\cdot S_b(\phi))}U_a(s_a, \cdot)$	$E_{\mu(\cdot S_b(I))}U_a(s_a, \cdot)$	$E_{\mu(\cdot S_b(J))}U_a(s_a, \cdot)$
RT	1	4.5	1.5
RB	1	4	2.5
LT'	1	1	1.5
LB'	1	0	0

Table I: $\mu(\{o\}|S_b(\phi)) = 1$, $\mu(\{t\}|S_b(I)) = \mu(\{m\}|S_b(I)) = \mu(\{m\}|S_b(J)) = \mu(\{b\}|S_b(J)) = \frac{1}{2}$

strictly preferred to RB . While $E_{\mu(\cdot|S_b(I))}U_a(RT, \cdot) > E_{\mu(\cdot|S_b(I))}U_a(RB, \cdot)$, it is also the case that $E_{\mu(\cdot|S_b(J))}U_a(RB, \cdot) > E_{\mu(\cdot|S_b(J))}U_a(RT, \cdot)$. Moreover, since $\mu(S_b(I)|S_b(J)) > 0$ and $\mu(S_b(J)|S_b(I)) > 0$, one has $S_b(I) \stackrel{\mu}{=} S_b(J)$. Thus, it is *not* the case that RT is infinitely more likely to be better than RB (or vice versa). Yet RB is not sequentially rational given μ . \square

One reason the conclusion in Example 4 might seem counterintuitive is that, while RB does yield a higher expected payoff than RT conditional on $S_b(J) = \{m, b\}$, it delivers a lower payoff than RT in case Bob plays t . From an ex-ante perspective, one might expect Ann to take this into account when comparing RT and RB . However, this requires considering the expected payoffs of RT and RB conditional on $\{t, m, b\}$, which is not a conditioning event for Ann: there is no information set $K \in \mathcal{I}_a$ such that $S_b(K) = \{t, m, b\}$.

The key insight of this section is that, despite the fact that $\{t, m, b\} \notin S_b(\mathcal{I}_a)$, there is a *unique* belief conditional on $\{t, m, b\}$ that is consistent with Ann's CPS μ . Formally, there is a unique CPS ν with conditioning events $S_b(\mathcal{I}_a) \cup \{\{t, m, b\}\}$ that coincides with μ on $S_b(\mathcal{I}_a)$. Note first that, for any such CPS ν , the fact that $\{t, m, b\} = S_b(I) \cup S_b(J)$ and $S_b(I) \stackrel{\mu}{=} S_b(J)$ imply that $\nu(\cdot|\{t, m, b\})$ must assign positive probability to both $S_b(I)$ and $S_b(J)$.¹¹ But then, the chain rule, applied to the conditioning events $\{t, m, b\}$ and $S_b(K)$ for $K = I, J$, implies that

¹¹ $S_b(I) \stackrel{\mu}{=} S_b(J)$ implies that $\mu(S_b(I)|S_b(J)) > 0$ and $\mu(S_b(J)|S_b(I)) > 0$; since $\mu(S_b(K)|S_b(K)) = 1$ for $K = I, J$, $\mu(\{m\}|S_b(K)) = \mu(S_b(I) \cap S_b(J)|S_b(K)) > 0$, and so $\nu(\{m\}|S_b(K)) > 0$, for $K = I, J$. Also, either $\nu(S_b(I)|\{t, m, b\}) > 0$ or $\nu(S_b(J)|\{t, m, b\}) > 0$. In either case, by the chain rule $\nu(\{m\}|\{t, m, b\}) > 0$; the claim follows.

$\nu(\cdot|\{t, m, b\})$ must be the uniform distribution on $\{t, m, b\}$.

Notice that RT has a strictly higher expected payoff conditional on $\{t, m, b\}$ than RB . Furthermore, with respect to the CPS ν , only the events S_b and $\{t, m, b\}$ are basic, because $\nu(\{t, m\}|\{t, m, b\}) = \nu(\{m, b\}|\{t, m, b\}) = \frac{2}{3} > 0$. Therefore, if one modifies Definition 6 so as to employ the extended CPS ν rather than the original CPS μ , strategy RB is no longer structurally rational. This resolves the issue highlighted in Example 4.

To generalize this argument, fix a game $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N})$, a player $i \in N$, and a CPS $\mu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i))$. Let

$$S_{-i}(\mathcal{I}_i; \mu) = \left\{ \bigcup_{\ell=1}^L S_{-i}(I_\ell) : L \geq 1, I_1, \dots, I_L \in \mathcal{I}_i, S_{-i}(I_\ell) =^\mu S_{-i}(I_{\ell+1}) \forall \ell = 1, \dots, L-1 \right\} \quad (4)$$

That is, $S_{-i}(\mathcal{I}_i; \mu)$ consists of unions of events in $S_{-i}(\mathcal{I}_i)$ that belong to the same equivalence class for $=^\mu$. In particular, $S_{-i}(\mathcal{I}_i) \subseteq S_{-i}(\mathcal{I}_i; \mu)$. Finally, I say that $\nu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i; \mu))$ is an **extension** of $\mu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i))$ if $\nu(\cdot|F) = \mu(\cdot|F)$ for all $F \in S_{-i}(\mathcal{I}_i)$. If μ admits such an extension, say that it is **extensible**.

Appendix A shows that, if a CPS is extensible, its extension is unique. Furthermore, while not every CPS in a general dynamic game is extensible, those that do not are inconsistent in the way they encode the relative likelihood of coplayers' strategies. Appendix A then shows that a CPS is extensible if and only if it satisfies a strengthening of the chain rule, or, equivalently, if it can be obtained as the limit of strictly positive probabilities. The latter result thus connects the approach of this paper with "belief trembles."

Definition 7 (Structural preferences for general games) Fix a dynamic game $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N})$, a player $i \in N$, a CPS $\mu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i))$ that admits an extension $\nu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i; \mu))$, and strategies $s_i, t_i \in S_i$. Then s_i is (weakly) **structurally preferred** to t_i ($s_i \succsim^\mu t_i$) iff, for every basic event $F \in S_{-i}(\mathcal{I}_i; \mu)$ such that $E_{\nu(\cdot|F)} U_i(s_i, \cdot) < E_{\nu(\cdot|F)} U_i(t_i, \cdot)$, there is a basic event $G \in S_{-i}(\mathcal{I}_i; \mu)$, with $G >^\nu F$, such that $E_{\nu(\cdot|G)} U_i(s_i, \cdot) > E_{\nu(\cdot|G)} U_i(t_i, \cdot)$.

As anticipated, Definition 7 only differs from Definition 6 because it employs an extension of the given CPS μ . Thus, the intuition given in Section 3 for structural preference applies here as well: s_i is infinitely more likely to be better than t_i than to be worse.

I continue to employ the notation $s_i \succ^\mu t_i$ (rather than $s_i \succ^\nu t_i$) because the extended CPS ν is uniquely determined by μ ; thus, ultimately, preferences are determined by μ as well.

Appendix C proves formally that Definition 6 is indeed a special case of Definition 7. Also, the above formalization emphasizes the similarities between the two definitions. Corollary 2 in Appendix B provides an alternative formulation that simplifies Definition 7 in two ways: first, it explicitly identifies the basic events for the extension of a CPS μ ; second, it employs the original likelihood ordering induced by μ , instead of the ordering induced by its extension.

5 Main Results

5.1 Structural Rationality implies Sequential Rationality

The first main result of this paper can now be stated. For readers who skipped Section 4: in a game with nested strategic information, every CPS is extensible.

Theorem 1 *Fix a dynamic game $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$, a player $i \in N$, and an extensible CPS $\mu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i))$ for player i . if strategy $s_i \in S_i$ is structurally rational for μ , then it is sequentially rational for μ .*

Example 2 shows that the converse does not hold: there may be strategies that are sequentially rational for a given CPS of player i , but not structurally rational for *any* CPS of i . As noted above, this is because structural rationality takes into account all of the player's basic beliefs, not just the ones she holds at information sets allowed by the strategy under consideration. Yet, structural rationality employs the *minimal* amount of information about a player's conditional beliefs that is necessary to ensure sequential rationality:

Example 5 Consider the game in Fig. 6, which differs from Fig. 4 in the payoffs assigned following U, m and D, t . Assume that $\mu(\{b\}|S_b) = 1$.

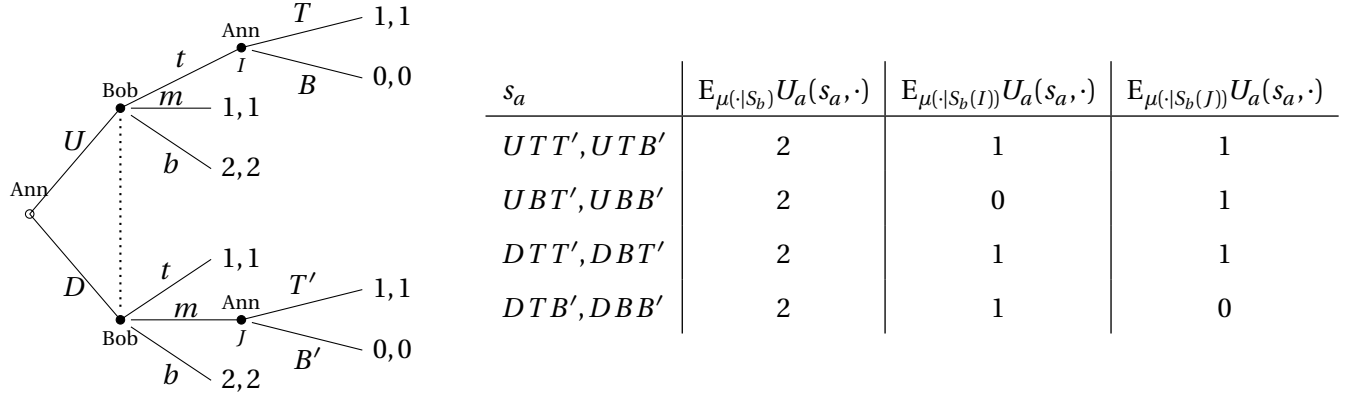


Figure 6: Minimality of basic beliefs. Ann's CPS: $\mu(\{b\}|S_b(\phi)) = 1$

Suppose that, in applying Definition 6, one considers Ann's basic beliefs $\mu(\cdot|S_b)$ and $\mu(\cdot|S_b(I))$, but not $\mu(\cdot|S_b(J))$. Then strategies DTB', DBB' are undominated in the resulting ordering, despite the fact that they are *not* sequentially rational. Definition 6 instead rules out these strategies precisely because it takes the basic belief $\mu(\cdot|S_b(J))$ into account as well. \square

The intuition for Theorem 1 is easiest to convey in a game with nested strategic information. The basic argument is reminiscent of the proof that, with EU preferences, an ex-ante optimal strategy must be optimal conditional upon a positive-probability information set. The additional power of structural preferences allows one to extend the argument to information sets that have zero prior probability.

Fix a player i , a CPS μ for i , and an information set I . The case of greatest interest occurs when $S_{-i}(I)$ is basic, as per Definition 3 (the other cases can be reduced to this). Suppose that strategy s_i is structurally rational given μ , and assume by contradiction that there is another strategy t_i such that $E_{\mu(\cdot|S_{-i}(I))}U_i(t_i, \cdot) > E_{\mu(\cdot|S_{-i}(I))}U_i(s_i, \cdot)$. Define a strategy s_i^* that agrees everywhere with s_i except at I and subsequent information sets, where it agrees with t_i . I now argue that $s_i^* \succ^\mu s_i$, which contradicts the assumption that s_i is structurally rational.

Since s_i^* agrees with t_i at I and all subsequent information sets, $E_{\mu(\cdot|S_{-i}(I))}U_i(s_i^*, \cdot) > E_{\mu(\cdot|S_{-i}(I))}U_i(s_i, \cdot)$. Furthermore, let $S_{-i}(J) \in \mathcal{S}_{-i}$ be a basic event for i such that $S_{-i}(J) \supset S_{-i}(I)$. Since $S_{-i}(I)$ is basic, $\mu(S_{-i}(I)|S_{-i}(J)) = 0$. Therefore, the only strategy profiles to which $\mu(\cdot|S_{-i}(J))$ gives positive probability are those that, starting from J , do *not* reach I . But at all such information sets, s_i and s_i^* coincide. Thus, $E_{\mu(\cdot|S_{-i}(J))}U_i(s_i^*, \cdot) = E_{\mu(\cdot|S_{-i}(J))}U_i(s_i, \cdot)$. Since this holds for all basic $S_{-i}(J)$ with $S_{-i}(J) \supset S_{-i}(I)$, it follows that it is *not* the case that $s_i \succsim^\mu s_i^*$.

Finally, consider an arbitrary, basic event $S_{-i}(J)$ for i and assume that $E_{\mu(\cdot|S_{-i}(J))}U_i(s_i^*, \cdot) < E_{\mu(\cdot|S_{-i}(J))}U_i(s_i, \cdot)$. Since s_i and s_i^* agree at information sets that do not (weakly) follow I , it must be the case that $\mu(\cdot|S_{-i}(J))$ assigns positive probability to some s_{-i} that reaches I , so that $\mu(S_{-i}(I)|S_{-i}(J)) > 0$. This means that $S_{-i}(I) \cap S_{-i}(J) \neq \emptyset$, and since the game has nested strategic information, either $S_{-i}(I) \supseteq S_{-i}(J)$ or $S_{-i}(J) \supseteq S_{-i}(I)$. But since $S_{-i}(I)$ is basic, $\mu(S_{-i}(I)|S_{-i}(J)) > 0$ implies that we must have $S_{-i}(I) \supset S_{-i}(J)$. By construction, $E_{\mu(\cdot|S_{-i}(I))}U_i(s_i^*, \cdot) > E_{\mu(\cdot|S_{-i}(I))}U_i(s_i, \cdot)$. Since this holds for all basic $S_{-i}(J)$ with $S_{-i}(I) \supset S_{-i}(J)$, conclude that $s_i^* \succsim^\mu s_i$. Hence, $s_i^* \succ^\mu s_i$, as claimed: contradiction.

Besides handling cases in which $S_{-i}(I)$ is not basic, the proof of Theorem 1 exploits the properties of extensible CPSs to adapt the analysis to general dynamic games.

5.2 Eliciting Conditional Beliefs

To elicit players' beliefs, I build upon the *side-bets* approach suggested by Luce and Raiffa (1957, §13.6) for games with simultaneous moves. To illustrate, Figure 7 depicts the Matching Pennies game, and an augmented game that elicits a bound on the probability that Ann assigns to Bob's choice of H . In the augmented game, both players choose a strategy from the original game (H or T); in addition, Ann must either *bet* on Bob's choice of H , or take a constant payoff of p "utils." The bet b pays 1 "util" if Bob chooses H , and 0 otherwise. Chance chooses the *left* or *right* matrix (Chance's payoff is constant and hence omitted in Figure 7). Bob's payoffs are as in Matching Pennies, regardless of Chance's move and Ann's choice of

b or p ; Ann receives the Matching Pennies payoff if Chance chooses ℓ , and the payoff from her choice of b or p otherwise. Assume that Ann (i) expects Bob to play H with probability $\pi > \frac{1}{2}$ ¹² in *both* the original Matching Pennies game and in the elicitation game, and (ii) deems each Chance move equally likely and independent of Bob's move. Then her best replies in the elicitation game are (H, b) if $\pi > p$, (H, p) if $\pi < p$, and both (H, b) and (H, p) if $\pi = p$. Thus, as in Matching Pennies, H is optimal for Ann (recall that $\pi > \frac{1}{2}$); in addition, her choice of b or p reveals whether $\pi \geq p$ or $\pi \leq p$.

Ann \ Bob	H	T
H	1,-1	-1,1
T	-1,1	1,-1

Ann \ Bob	H	T
H, b	1,-1	-1,1
H, p	1,-1	-1,1
T, b	-1,1	1,-1
T, p	-1,1	1,-1

Ann \ Bob	H	T
H, b	1,-1	0,1
H, p	$p,-1$	$p,1$
T, b	1,1	0,-1
T, p	$p,1$	$p,-1$

ℓ
 r

Chance

Figure 7: Matching Pennies (left) and the Luce-Raiffa approach to eliciting Ann's beliefs (right)

In actual experiments, this approach is modified to derive tighter bounds on Ann's beliefs, while preserving incentive compatibility. For instance, when Chance chooses r , a random-lottery incentive scheme (Grether and Plott, 1979) can be used to elicit Ann's ranking of the bet b against different utility values p_1, \dots, p_n , rather than a single value p . Alternatively, the Becker, DeGroot, and Marschak (1964) mechanism can be played when Chance chooses r .

Adapting Luce and Raiffa's construction to dynamic games requires addressing novel incentive issues. Consider the two-legged centipede game in Fig. 8. Suppose one is interested in the probability that Ann assigns at the root of the tree to Bob's choice of d .

As in the Luce-Raiffa procedure, one might offer Ann a side bet b that yields 1 util in case

¹²In this example, I assume a non-Nash belief so that Ann has a unique best reply in Matching Pennies.

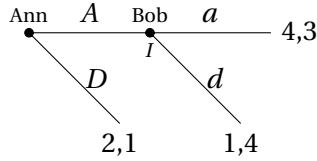


Figure 8: A two-legged centipede game.

Bob chooses d and 0 otherwise, and a constant bet that yields $p \in [0, 1]$ utils for sure. However, an issue arises. If Ann—as predicted by backward induction—expects Bob to choose d , then any ex-ante optimal strategy in both the original game and the Luce-Raiffa game must prescribe that she choose D . But since Bob only moves if Ann chooses A , *Bob's choice cannot be observed, and therefore bet b cannot actually be “paid out.”* As a consequence, whether Ann chooses b or p is *not* in response to real incentives.

The solution I propose is to add side bets a la Luce-Raiffa to the *strategy-method implementation* of the game. Recall that, in the strategy method, players commit ex-ante and simultaneously to extensive-form strategies. Therefore, the experimenter can always observe their choices, and actually pay out bets on coplayers' strategies. Thus, real incentives *can* be provided, just as in simultaneous-move games.

The required construction for the two-legged centipede game is displayed in Fig. 9. For simplicity, the unobserved chance move is *not shown*. Instead, Ann's payoff at each terminal node is specified as a vector; Chance's move determines which of its two components accrues to Ann. As in Figure 7, Bob's payoffs instead are unaffected by Chance's move. The root and Bob's first information set correspond to the first stage of the strategy method. Ann commits to a strategy in the original game (i.e., an element of $\{A, D\}$) as well as one of b or p . Simultaneously, Bob commits to a strategy in the original game (d or a); the symbol \emptyset indicates that Bob is not asked to engage in any side bet. Then, in the second stage of the strategy method, the players' choices are implemented; to formalize this, at every second-stage information set, the only action available is the one specified by the strategy the player has committed to (see

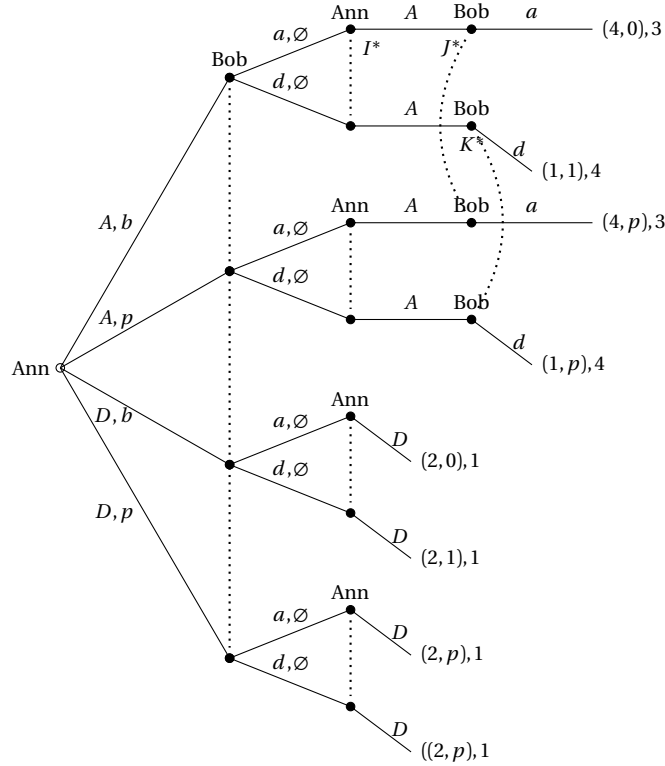


Figure 9: Strategy-method elicitation in the two-legged centipede (simplified).

e.g. I^* , J^* and K^* in Figure 9). Crucially, *players observe the resulting play as it unfolds*, just as they would in the original game. For instance, at information set I^* , Ann is not informed of Bob's first-stage choice. Similarly, at J^* and K^* , Bob learns that Ann chose A ; note however that he does *not* learn Ann's choice of side bet (whether it is b or p). Thus, the information Bob has at both J^* and K^* about Ann's choices corresponds to the information he has at I in the original game; J^* and K^* are distinct only because, consistently with perfect recall, they also encode Bob's own first-stage choice.

Notice that, even if Ann chooses D, b or D, p at the initial node, the experimenter can still observe Bob's (simultaneous) choice of a, \emptyset or b, \emptyset and pay out Ann's chosen side bet accordingly, if Chance's move is t . Thus, Ann's choice of b vs. p reflects real incentives.

Of course, this construction modifies the game in a significant way; players' strategic rea-

soning may change as a result. For instance, in the game of Figure 9, sequential rationality reduces to ex-ante optimality for Bob. However, Theorem 2 below shows that, *if instead players are structurally rational, then strategic incentives are preserved in the elicitation game*. This provides a rationale for the elicitation of beliefs using the strategy method.

The construction in Figure 9 can be adapted to elicit *conditional* beliefs. Consider the four-legged centipede game in Figure 1 in the Introduction, and suppose one is interested in Ann's beliefs at the third node I about Bob's choice of d_2 —that is, $\mu(\{a_1 d_2\} | S_b(I))$, where μ denotes Ann's CPS. This can be achieved by offering Ann a *conditional side bet* at the beginning of the game. If she chooses b , Ann will receive 1 util if Bob indeed chooses $a_1 d_2$, and zero otherwise. If she chooses p , she will receive p utils if Bob chooses $a_1 d_2$ or $a_1 a_2$, and zero if he chooses d_1 . In other words, the bet is “called off” if Bob does not play a strategy in $S_b(I)$. If Ann has ex-ante EU preferences and the prior probability of $S_b(I)$ is positive, then Ann strictly prefers b if and only if $\mu(\{a_1 d_2\} | S_b(I)) \geq p$. The proof of Theorem 2 establishes that, if Ann has structural preferences, this is true *even if she initially assigns probability zero to $S_b(I)$* . As in the case of the Luce-Raiffa procedure, this construction can be extended to obtain tighter bounds on $\mu(\{a_1 d_2\} | S_b(I))$, and indeed elicit Ann's entire CPS μ .

I now formalize the definition of the elicitation game. It is convenient to allow for the simultaneous elicitation of beliefs from zero, one or more players: see Corollary 1. Given a dynamic game $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$, a **questionnaire** is a collection $Q = (Q_i)_{i \in N}$ such that, for every $i \in N$, either $Q_i = \emptyset$ or $Q_i = (I, E, p)$, with $I \in \mathcal{I}_i$, $E \subseteq S_{-i}(I)$, and $p \in [0, 1]$.

Definition 8 Fix a dynamic game $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$ and a questionnaire $Q = (Q_i)_{i \in N}$. The **elicitation game for Q** is the tuple $(N \cup \{c\}, (S_i^*, \mathcal{I}_i^*, U_i^*)_{i \in N^*}, S^*(\cdot))$, where $S_c^* = \{\emptyset\} \cup \{i \in N : Q_i \neq \emptyset\}$, $\mathcal{I}_c^* = \{\phi^*\}$, $S^*(\phi^*) = S^*$, $U_c^* \equiv 0$, and the following properties hold for all $i \in N$:

1. *Strategies:* $S_i^* = S_i \times W_i$, where $W_i = \{\emptyset\}$ if $Q_i = \emptyset$ and $W_i = \{b, p\}$ if $Q_i = (I, E, p)$;
2. *Information:* $\mathcal{I}_i^* = \{I_i^1\} \cup \{(s_i, w_i, I) : (s_i, w_i) \in S_i^*, I \in \mathcal{I}_i, s_i \in S_i(I)\}$;
3. *First stage:* $S^*(I_i^1) = S^*$

4. *Second stage: for all $(s_i, w_i, I) \in \mathcal{I}_i^*$, $S^*((s_i, w_i, I)) = \{(s_i, w_i)\} \times S_{-i}(I) \times W_{-i} \times S_c^*$,*¹³
5. *Payoffs: for all $((s_i, w_i), (s_{-i}, w_{-i}), s_c^*) \in S^*$ with $s_c^* \neq i$, $U_i^*((s_i, w_i), (s_{-i}, w_{-i}), s_c^*) = U_i(s_i, s_{-i})$; and for all $(s_{-i}, w_{-i}) \in S_{N \setminus \{i\}}^*$, if $Q_i = (I, E, p)$, then*

$$U_i((s_i, b), (s_{-i}, w_{-i}), i) = \begin{cases} 1 & s_{-i} \in E \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad U_i((s_i, p), (s_{-i}, w_{-i}), i) = \begin{cases} p & s_{-i} \in S_{-i}(I) \\ 0 & \text{otherwise.} \end{cases}$$

Thus, chance can select either \emptyset , in which case the payoffs are as in the original game, or else one of the players whose beliefs one wishes to elicit. Each (non-chance) player i has a first-stage information set I_i^1 . By Part 4 of the definition, at every second-stage information set (s_i, w_i, I) , player i recalls her first-stage choice (s_i, w_i) ; furthermore, what i learns about (real) coplayers at (s_i, w_i, I) is precisely what she learns about them at I in the original game.¹⁴ In the game of Figure 9, for instance, $J^* = (a, \emptyset, I)$ and $S^*(J^*) = \{(a, \emptyset)\} \times \{(A, b), (A, p)\} \times \{\text{Ann}, \emptyset\}$; similarly, $K^* = (d, \emptyset, I)$ and $S^*(K^*) = \{(d, \emptyset)\} \times \{(A, b), (A, p)\} \times \{\text{Ann}, \emptyset\}$. Finally, part 5 defines payoffs. In particular, if chance selects a player i whose beliefs we wish to elicit, the first-period choice of b vs. p represents a conditional side bet, as described above.

The next step is to formalize the assumption that a player holds the same conditional beliefs about coplayers in the original game and in the elicitation game. In addition, as in the Luce-Raiffa procedure, I assume that players view chance moves as stochastically independent of coplayers' strategies.

¹³ $W_{-i} = (W_j)_{j \neq i}$, where each W_j is as defined in part 1. Also, here and in part 5, it is convenient to decompose $S^* = (S_i \times W_i) \times (S_{N \setminus \{i\}} \times W_{N \setminus \{i\}}) \times S_c^*$.

¹⁴This part of the definition also implies that i has a single action available at (s_i, w_i, I) . To prove this formally, one has to start from a full specification of the extensive form; this is done in the Online Appendix. Yet, the reduced-form description in part 4 of Definition 8 captures the basic intuition: (s_i, w_i) is the *only* strategy of i that allows (s_i, w_i, I) to be reached; if there were two or more actions available at (s_i, w_i, I) , then for each of them there would be a different strategy of i that also allows (s_i, w_i, I) , and only differs from (s_i, w_i) in the choice made at (s_i, w_i, I) .

Definition 9 Fix a dynamic game $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$, and a questionnaire $(Q_i)_{i \in N}$. Let $(N^*, (S_i^*, \mathcal{I}_i^*, U_i^*)_{i \in N^*}, S^*(\cdot))$ be the corresponding elicitation game. For any $i \in N$ and CPS $\mu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i))$, the CPS $\mu^* \in \Delta(S_{-i}^*, S_{-i}^*(\mathcal{I}_i^*))$ **agrees with** μ if, for every $I^* \in \mathcal{I}_i^*$,

$$\text{marg}_{S_{-i} \times S_c^*} \mu^*(\cdot | S_{-i}^*(I^*)) = \frac{1}{|S_c^*|} \mu(\cdot | \text{proj}_{S_{-i}} S_{-i}^*(I^*)).^{15} \quad (5)$$

More than one CPS for player i in the elicitation game may agree with her CPS in the original game. This is because i may assign different probabilities to her coplayers' choices of side bets in the elicitation game. However, these differences are irrelevant for her strategic reasoning, because her payoff does not depend on these choices. On the other hand, independence of Chance's moves is important: if i believes that her coplayers correlate their choices with Chance, this may impact her expected payoffs, and hence her strategic incentives.

The main result of this section can now be stated.

Theorem 2 Fix a dynamic game $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$, a questionnaire Q , and the corresponding elicitation game $(N^*, (S_i^*, \mathcal{I}_i^*, U_i^*)_{i \in N^*}, S^*(\cdot))$. For any player $i \in N$, fix an extensible CPS $\mu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i))$. Then there exists an extensible CPS $\mu^* \in \Delta(S_{-i}^*, S_{-i}^*(\mathcal{I}_i^*))$ that agrees with μ . For any such CPS,

- (1) for all $(s_i, w_i), (t_i, w_i) \in S_i^*$, $(s_i, w_i) \succ^{\mu^*} (t_i, w_i)$ if and only if $s_i \succ^\mu t_i$;
- (2) if $Q_i = (I, E, p)$, then for all $s_i \in S_i$, $p > \mu(E | S_{-i}(E))$ implies $(s_i, p) \succ^{\mu^*} (s_i, b)$ and $p < \mu(E | S_{-i}(I))$ implies $(s_i, b) \succ^{\mu^*} (s_i, p)$.

Hence, if $Q_i = (I, E, p)$ and (s_i, b) (resp. (s_i, p)) is structurally rational in the elicitation game, then s_i is structurally rational in the original game, and $\mu(E | S_{-i}(I)) \geq p$ (resp. $\mu(E | S_{-i}) \leq p$).

Statement (1) formalizes the assertion that every player's strategic incentives are preserved:

¹⁵By Definition 8, $\text{proj}_{S_i} S_{-i}^*(\phi^*) = \text{proj}_{S_{-i}} S_{-i}^*(I_i^1) = S_{-i}$ and, for all $(s_i, w_i, I) \in \mathcal{I}_i^*$, $\text{proj}_{S_{-i}} S_{-i}^*((s_i, w_i, I)) = S_{-i}(I)$.

if their beliefs are the same as in their original game, so is their ranking of strategies.¹⁶ Statement (2) and the last assertion of the theorem formalize the claim that, by observing first-stage betting choices, it is possible to elicit a bound on players' first-order beliefs.¹⁷

Statement (1) also provides a positive rationale for the use of the strategy method. Indeed, notice that if the questionnaire Q is such that $Q_i = \emptyset$ for all i , the elicitation game simply implements the strategy method (Chance's only move is dummy). One then has:

Corollary 1 *Under the assumptions of Theorem 2, suppose that $Q_i = \emptyset$ for all $i \in N$. Then, for all $i \in N$ and all $s_i, t_i \in S_i$, $s_i \succ^\mu t_i$ if and only if $(s_i, \emptyset) \succ^{\mu^*} (t_i, \emptyset)$. In particular, s_i is structurally rational in the original game if and only if (s_i, \emptyset) is structurally rational in the elicitation game.*

These results do not hold if players are only sequentially, rather than structurally rational. Even if players' conditional beliefs are the same as in the original game, the set of sequentially rational strategies in the elicitation game coincides with the set of ex-ante rational strategies. Furthermore, the ranking of utility acts is determined by ex-ante beliefs alone; hence, it is not informative about players' conditional beliefs at zero-probability information sets:

¹⁶In Figure 9, there is a unique structurally rational strategy for Bob in the elicitation game— (d, \emptyset) —which corresponds to the unique structurally rational strategy in the original centipede game. Thus, the assumptions that Ann believes that Bob (a) is rational, and (b) has the same beliefs as in the original game are enough to imply that Ann *must* hold the same beliefs about Bob as in the original game. This is not the case in general. Yet, I suggest that, if i 's preferences are preserved in the elicitation game, it is at least reasonable to assume that the coplayers $j \neq i$ continue to hold the same beliefs about i . By comparison, Remark 2 below shows that, with conditional EU maximization, i 's preferences are *not* preserved in the elicitation game, so the rationale for keeping j 's beliefs fixed is weaker.

¹⁷The weak inequality in the last part of the Theorem accounts for the fact that, if $p = \mu(E|S_{-i})$, the strategies (s_i, b) and (s_i, p) may be incomparable: howevern, if e.g. (s_i, b) is structurally rational, one can definitely rule out the case $p > \mu(E|S_{-i}(E))$, because by Statement (2) that would imply $(s_i, p) \succ^{\mu^*} (s_i, b)$, contradicting the structural rationality of (s_i, b) .

Remark 2 *Under the assumptions of Theorem 2, for every player $i \in N$, $(s_i, w_i) \in S_i^*$ is sequentially rational in the elicitation game if and only if (i) $s_i \in \arg \max_{t_i \in S_i} \mathbb{E}_{\mu(\cdot|S_{-i})} U_i(t_i, \cdot)$, and (ii) if $Q_i = (I, E, p)$ and $w_i = b$ (resp. $w_i = p$), then $\mu(E|S_{-i}) \geq p$ (resp. $\mu(E|S_{-i}) \leq p$).*

This is an immediate consequence of the fact that, for each player i , the only information set in the elicitation game where more than one action is available is I_i^1 .

6 Discussion

Incomplete-information games The analysis of this paper may be easily adapted to accommodate incomplete information. Fix a dynamic game with N players, strategy sets S_i and information sets \mathcal{I}_i for each $i \in N$, and a strategy profile correspondence $S(\cdot)$. Consider a set Θ_0 , and sets Θ_i for each $i \in N$; each $\theta_i \in \Theta_i$ represents a possible realization of player i 's private information, which may or may not be payoff-relevant; Θ_0 captures residual uncertainty that is not reflected in players' private information. Player i 's payoff function is a map $U_i : S \times \Theta \rightarrow \mathbb{R}$, where $\Theta = \Theta_0 \times \prod_{j \in N} \Theta_j$. The set of conditioning events for player i is $\mathcal{F}_i = \{S_{-i}(I) \times \Theta_{-i} : I \in \mathcal{I}_i\}$, where $\Theta_{-i} = \Theta_0 \times \prod_{j \in N \setminus \{i\}} \Theta_j$; player i 's conditional beliefs can be represented via a CPS μ on $S_{-i} \times \Theta_{-i}$, with conditioning events \mathcal{F}_i . Depending on the application, player i may hold a different CPS for each realization of her private information θ_i .

Definitions 2, 6 and 7 can be readily adapted to characterize notions of sequential and structural rationality at the interim stage (that is, for fixed private-information parameters $\theta_i \in \Theta_i$). Theorems 1, 2 and 3 admit straightforward, corresponding extensions.

Lexicographic expected utility. As noted in Section 3, the definition of structural preferences bears formal resemblance to that of lexicographic expected utility. The key difference is that lexicographic probability systems (LPSs) are defined purely in terms of the strategic form of the game; they do not take the extensive form into account. On the other hand, structural preferences are explicitly defined for a particular, given extensive form. This has an important

consequence: an LPS can generate a CPS by conditioning,¹⁸ but the same CPS may be generated by *multiple* LPSs. For instance, the CPS μ in Example 3 can be generated by the LPS $\lambda^1 = (\delta_O, \delta_U, \delta_D)$, but also by the LPS $\lambda^2 = (\delta_O, \delta_D, \delta_U)$, where δ_ω denotes the Dirac measure concentrated on $\{\omega\}$. Intuitively, λ^1 deems U to be infinitely more likely than D , whereas the opposite is true of λ^2 . However, *this likelihood assessment is not derived from Bob's CPS μ* . By way of contrast, by design, structural preferences are defined solely in terms of information that can be derived from the player's conditional beliefs. This reflects the fact that structural rationality is explicitly motivated by extensive-form analysis. Lexicographic expected-utility maximization is instead a strategic-form concept; it was introduced into game theory to analyze refinements for games with simultaneous moves (Blume et al., 1991b), and moreover, when coupled with a full-support assumption, it incorporates an *invariance* requirement; see Brandenburger (2007), §12. Finally, as noted in Section 3, structural preferences reduce to EU in simultaneous-move games; on the other hand, lexicographic preferences may of course differ from EU in such games.

Conditional expected-utility maximization Myerson (1986) axiomatizes conditional expected utility maximization with respect to a CPS. The analysis assumes that a family of conditional preferences is taken as given. Preferences conditional upon nested events are related by *subjective substitution*, which is shown to characterize the chain rule of conditioning for CPSs. Just like prior beliefs do not fully determine the player's CPS due to the presence of ex-ante zero-probability events, prior preferences do not fully determine the entire system of conditional preferences. Thus, in Myerson's analysis, it is necessary to assume that all conditional preferences are observable. As argued in the Introduction, this may be problematic in many dynamic games. By way of contrast, the present paper defines an *ex-ante* preference relation; Theorem 2 shows that it is elicitable by observing initial choices in suitably-designed experi-

¹⁸Fix an LPS $\lambda = (p_1, \dots, p_M)$. For every $I \in \mathcal{I}_i$, let $\mu(\cdot | S_{-i}(I)) = p_{m(I)}(\cdot | S_{-i}(I))$, where $m(I)$ is the *lowest* index m for which $p_m(S_{-i}(I)) > 0$ —assuming one such index can be found.

ments.¹⁹

Structural and lexicographic consistency The notion of basis incorporates a version of *structural consistency* (Kreps and Wilson, 1982; Kreps and Ramey, 1987): conditional beliefs are derived from a collection of alternative prior probabilistic hypotheses about the play of opponents. To elaborate, every CPS reflects a trivial notion of structural consistency: every conditional belief in a CPS can be interpreted as an alternative “prior” hypothesis that is adopted once the corresponding information set is reached. However, as Theorem 3 shows, an arbitrary CPS may assign relative likelihoods in inconsistent ways, and thus represent alternative hypotheses that are not just trivial, but contradictory. The existence of a basis avoids such inconsistencies. In addition, a basis incorporates a notion of parsimony: it identifies the minimal set of alternative hypotheses that generate the CPS.

Kreps and Wilson (1982) also consider a notion of *lexicographic consistency*. Their definition is stated in the setting of equilibrium, rather than individual maximization; furthermore, conditional beliefs are represented by consistent assessments. Translated to the present setting and notation, lexicographic consistency requires that the player’s CPS can be generated by an LPS, as described above. Hence, the above comparison with LPSs applies: in the present analysis, the basis and its ordering is entirely derived from the CPS. Thus, CPSs are the starting point of the analysis. Lexicographic consistency, on the other hand, gives priority to an LPS, which adds information not present in the player’s CPS.

Preferences for the timing of uncertainty resolution The fact that structural preferences depend upon the extensive form of the dynamic game can be seen as loosely analogous to the issue of sensitivity to the timing of uncertainty resolution: see e.g. Kreps and Porteus (1978); Epstein and Zin (1989), and in particular Dillenberger (2010). In the latter paper, preferences are allowed to depend upon whether information is revealed gradually rather than in a single

¹⁹The same observability issue applies to Asheim and Perea (2005), who generalize Myerson’s analysis.

period, even if no action can be taken upon the arrival of partial information. This is close in spirit to the observation that subjects behave differently in the strategic form of a dynamic game (where all uncertainty is resolved in one shot), and when the game is played with commitment as in the strategy method (where information arrives gradually). The key difference is that, for structural preference, this dependence only affects preferences when some piece of partial information has zero prior probability—that is, when there is *unexpected* partial information. If all conditioning events have positive probability, structural preference reduce to standard expected-utility preferences. (Of course, the same is true for sequential rationality, when all information set have positive prior probability.)

A Appendix: Characterizing extensible CPSs

Example 6 (A conditional Newcomb Paradox) Consider the game in Figure 10.

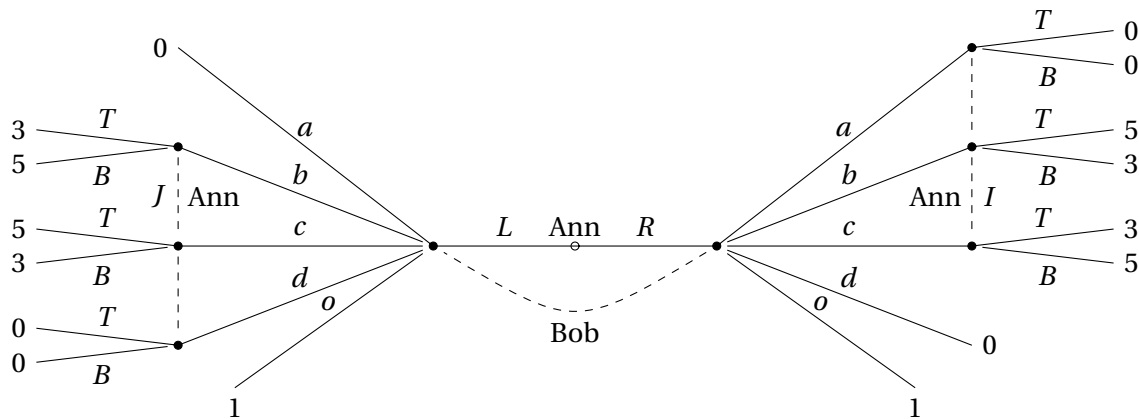


Figure 10: Newcomb's paradox?

Suppose that Ann's CPS μ satisfies $\mu(\{o\}|S_b(\phi)) = \mu(\{b\}|S_b(I)) = \mu(\{c\}|S_b(J)) = 1$. Observe that $S_b(\phi) \succ^\mu S_b(I) \simeq^\mu S_b(J)$. Hence, $\mathcal{S}_b(\mathcal{I}_a; \mu)$ contains the event $S_b(I) \cup S_b(J) = \{a, b, c, d\}$,

and any extension $\nu \in \Delta(S_b, \mathcal{S}_b(\mathcal{I}_a; \mu))$ must satisfy, in particular,

$$\begin{aligned} \nu(S_b(I)|S_b(I) \cup S_b(J)) &= \mu(\{b\}|S_b(I))\nu(S_b(I)|S_b(I) \cup S_b(J)) = \\ &= \nu(\{b\}|S_b(I) \cup S_b(J)) = \mu(\{b\}|S_b(J))\nu(S_b(J)|S_b(I) \cup S_b(J)) = 0 \text{ and} \\ \nu(S_b(J)|S_b(I) \cup S_b(J)) &= \mu(\{c\}|S_b(J))\nu(S_b(J)|S_b(I) \cup S_b(J)) = \nu(\{c\}|S_b(I) \cup S_b(J)) = \\ &= \mu(\{c\}|S_b(I))\nu(S_b(I)|S_b(I) \cup S_b(J)) = 0. \end{aligned}$$

But then $\nu(S_b(I) \cup S_b(J)|S_b(I) \cup S_b(J)) \leq \nu(S_b(I)|S_b(I) \cup S_b(J)) + \nu(S_b(J)|S_b(I) \cup S_b(J)) = 0$, contradiction. Hence μ is not extensible. \square

A peculiar feature of the CPS μ in Example 6 is that Ann's own initial choice of R vs. L determines her conditional beliefs on the relative likelihood of b and c , despite the fact that Bob does not observe Ann's initial choice. (In fact, Ann's first action and Bob's move may well be simultaneous.) This phenomenon is reminiscent of Newcomb's paradox (Weirich, 2016).

Observe that, if $S_b(I)$ and $S_b(J)$ both had positive prior probability, the definition of conditional probability would imply that the relative likelihood of b and c must be the same at both information sets. The same conclusion holds in any *consistent assessment* in the sense of Kreps and Wilson (1982), and in any *complete* CPS in the sense of Myerson (1986). The reason is that, in both cases, Ann's conditional beliefs at I and J are obtained by fixing a sequence (p^m) of strictly positive probability distributions on S_b , and taking the limit of the conditional probabilities $p^m(\cdot|S_b(I))$ and $p^m(\cdot|S_b(J))$.²⁰ For every index k , the relative likelihood of b vs. c is the same conditional on I and J , so this is true in the limit as well.

To sum up, CPSs that are not extensible fail consistency requirements that are both intuitive and follows from well-understood arguments based on "belief trembles." The following definition identifies an intrinsic property of CPSs that captures this condition.

²⁰Modulo notational differences, this is true by definition for consistent assessments; for complete CPSs, it follows from a result in Myerson (1986).

Definition 10 Fix a dynamic game $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$, a player $i \in N$, and a CPS $\mu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i))$. The CPS μ is **congruent** if, for every ordered list $I_1, \dots, I_L \in \mathcal{I}_i$ such that $\mu([I_{\ell+1}|S_{-i}(I_\ell)]) > 0$ for all $\ell = 1, \dots, L-1$, and all $E \subseteq S_{-i}(I_1) \cap S_{-i}(I_L)$,

$$\mu(E|S_{-i}(I_1)) \cdot \prod_{\ell=1}^{L-1} \frac{\mu(S_{-i}(I_\ell) \cap S_{-i}(I_{\ell+1})|S_{-i}(I_{\ell+1}))}{\mu([I_\ell] \cap S_{-i}(I_{\ell+1})|S_{-i}(I_\ell))} = \mu(E|S_{-i}(I_L))$$

Congruence is a strengthening of the chain rule of conditioning.²¹ Furthermore, it sheds light on the pathological nature of the beliefs in Example 6. Take $I_1 = I$ and $I_2 = J$ in Definition 10, and note that $S_{-i}(I) \cap S_{-i}(J) = \{b, c\}$, $\mu(S_{-i}(I) \cap S_{-i}(J)|S_{-i}(I)) > 0$, and $\mu(S_{-i}(I) \cap S_{-i}(J)|S_{-i}(J)) > 0$. Then the equation in Definition 10 implies that, in particular,

$$\frac{\mu(\{b\}|S_{-i}(I))}{\mu(\{b, c\}|S_{-i}(I))} = \frac{\mu(\{b\}|S_{-i}(J))}{\mu(\{b, c\}|S_{-i}(J))}.$$

Intuitively, the probability of b given $\{b, c\}$ should be the same, whether it is calculated from the perspective of I or J . This is a reasonable requirement, given that Ann's information about the relative likelihood of b vs. c is the same at I and J —neither has yet been ruled out. Yet this condition is violated in Example 6: $\mu(\{b\}|S_{-i}(I)) = 1$, but $\mu(\{b\}|S_{-i}(J)) = 0$.

The following theorem shows that a CPS is congruent if and only if it admits a (unique) extension. Furthermore, it formalizes the connection between congruence and belief trembles.

Theorem 3 Fix a dynamic game $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$, a player $i \in N$, and a CPS $\mu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i))$. The following are equivalent:

1. μ is congruent;
2. μ admits a unique extension $\nu \in \Delta(S_{-i}, \mathcal{I}_i(\mathcal{I}_i; \mu))$;
3. there is a sequence $(p^n) \in \Delta(S_{-i})$ such that $p^m(S_{-i}(I)) > 0$ for all m and $I \in \mathcal{I}_i$, and $p^m(E \cap S_{-i}(I))/p^m(S_{-i}(I)) \rightarrow \mu(E \cap S_{-i}(I)|S_{-i}(I))$ for all $I \in \mathcal{I}_i$ and $E \subseteq S_{-i}$.

²¹To see that it implies the chain rule, take $L = 2$ and consider the case $E \subseteq S_{-i}(I_1) \subseteq S_{-i}(I_2)$ in Definition 10.

B Appendix: Properties of extensible CPSs

Throughout, fix a dynamic game $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$.

For every player i , collection $\mathcal{C}_i \subseteq 2^{S_i} \setminus \{\emptyset\}$, and CPS $\mu \in \Delta(S_{-i}, \mathcal{C}_i)$, a μ -**sequence** is an ordered list $F_1, \dots, F_K \in \mathcal{C}_i$ such that $\mu(F_{k+1}|F_k) > 0$ for all $k = 1, \dots, K-1$. Thus, for all $F, G \in \mathcal{C}_i$, $F \geq^\mu G$ iff there is a μ -sequence F_1, \dots, F_K with $F_1 = G$ and $F_K = F$.

The following result states that every equivalence class of \geq^μ can be arranged in a μ -sequence. Observe that the elements of the μ -sequence constructed in the proof are not all distinct.

Lemma 1 *For every player i , collection $\mathcal{C}_i \subseteq 2^{S_i} \setminus \{\emptyset\}$, CPS $\mu \in \Delta(S_{-i}, \mathcal{C}_i)$, and event $F \in \mathcal{C}_i$, there is a μ -sequence $F_1, \dots, F_M \in \mathcal{C}_i$ such that $F_1 = F_M = F$ and, for all $G \in \mathcal{C}_i$, $G =^\mu F$ if and only if $G = F_m$ for some $m = 1, \dots, M$.*

Proof: let $\{F_1, \dots, F_L\}$ be an enumeration of the equivalence class of \geq^μ containing F ; in particular, assume without loss that $F_1 = F$. Then in particular $F_1 \geq^\mu F_2 \geq^\mu \dots \geq^\mu F_L$ and $F_L \geq^\mu F_1$. By definition, for every $\ell = 1, \dots, L-1$, there is a μ -sequence $F_1^\ell, \dots, F_{M(\ell)}^\ell$ such that $F_1^\ell = F_{\ell+1}$ and $F_{M(\ell)}^\ell = F_\ell$; furthermore, there is a μ -sequence $F_1^L, \dots, F_{M(L)}^L$ such that $F_1^L = F_1$ and $F_{M(L)}^L = F^L$. Then the ordered list

$$F_1^L, F_2^L, \dots, F_{M(L)}^L = F_1^{L-1}, \dots, F_{M(L-1)}^{L-1} = F_1^{L-2}, \dots, F_{M(1)}^1.$$

is a μ -sequence, with $F_1^L = F_1 = F$ and $F_{M(1)}^1 = F_1 = F$.

By construction, $F_\ell = F_{M(\ell)}^\ell$ for every $\ell = 1, \dots, L$, so this μ -sequence contains the equivalence class $\{F_1, \dots, F_L\}$ for F . Finally, notice that, for every $\ell = 1, \dots, L$ and $m = 1, \dots, M(\ell)$, the ordered sublist beginning with F_1^L and ending with F_m^ℓ , and the ordered sublist beginning with F_m^ℓ and ending with $F_{M(1)}^1$, are both μ -sequences, so $F_m^\ell \geq^\mu F_1^L$ and $F_{M(1)}^1 \geq^\mu F_m^\ell$. Furthermore, $F_1^L = F_{M(1)}^1 = F_1 = F$, so in fact $F_m^\ell \geq^\mu F$ and $F \geq^\mu F_m^\ell$, so $F_m^\ell = F_{\bar{\ell}}$ for some $\bar{\ell} = 1, \dots, L$. ■

The next result is useful to analyze extensions of CPSs. A collection $F_1, \dots, F_L \in \mathcal{C}_i$ μ -*supports* the probability measure $p \in \Delta(S_{-i})$ if $p(\cup_\ell F_\ell) = 1$ and, for all $\ell = 1, \dots, L$ and $E \subseteq F_\ell$,

$p(E) = \mu(E|F_\ell)p(F_\ell)$. In particular, if F_1, \dots, F_L is a subset of a \geq^μ -equivalence class and ν extends μ , then F_1, \dots, F_L μ -supports $\nu(\cdot|\cup_\ell F_\ell)$.

Lemma 2 Fix a player $i \in N$ and a CPS $\mu \in \Delta(S_{-i}, \mathcal{C}_i)$. Consider collections $F_1, \dots, F_L \in \mathcal{C}_i$ and $G_1, \dots, G_M \in \mathcal{C}_i$ such that $F_\ell =^\mu F_{\bar{\ell}}$ for all $\ell, \bar{\ell} \in \{1, \dots, L\}$, and $G_m =^\mu G_{\bar{m}}$ for all $m, \bar{m} \in \{1, \dots, M\}$. Suppose that G_1, \dots, G_M μ -supports the probability $p \in \Delta(S_{-i})$, and that $p(\cup_\ell F_\ell) > 0$. Then there are $\hat{\ell} \in \{1, \dots, L\}$ and $\hat{m} \in \{1, \dots, M\}$ such that

(i) $p(F_{\hat{\ell}} \cap G_{\hat{m}}) > 0$;

(ii) $p(F_{\hat{\ell}}) > 0$; and

(iii) $\mu(F_{\hat{\ell}}|G_{\hat{m}}) > 0$.

In particular, if G_1, \dots, G_M is an \geq^μ -equivalence class, then $p(G_m) > 0$ for all m .

Proof: We have

$$0 < p(\cup_\ell F_\ell) \leq \sum_{\bar{m}} p(G_{\bar{m}} \cap [\cup_\ell F_\ell]),$$

so there must be \hat{m} with $p(G_{\hat{m}} \cap [\cup_\ell F_\ell]) > 0$. Furthermore,

$$0 < p(G_{\hat{m}} \cap [\cup_\ell F_\ell]) \leq \sum_{\bar{\ell}} p(G_{\hat{m}} \cap F_{\bar{\ell}}),$$

so there is $\hat{\ell}$ with $p(G_{\hat{m}} \cap F_{\hat{\ell}}) > 0$, i.e., (i) holds. A fortiori, (ii) holds, and similarly $p(G_{\hat{m}}) > 0$. Since G_1, \dots, G_M μ -supports p , $p(F_{\hat{\ell}} \cap G_{\hat{m}}) = \mu(F_{\hat{\ell}} \cap G_{\hat{m}}|G_{\hat{m}}) \cdot p(G_{\hat{m}})$, so $\mu(F_{\hat{\ell}}|G_{\hat{m}}) = \mu(F_{\hat{\ell}} \cap G_{\hat{m}}|G_{\hat{m}}) > 0$, i.e., (iii) holds.

For the last statement, fix $m \in \{1, \dots, M\}$. By Lemma 1, there is a μ -sequence $F_1, \dots, F_L \in \mathcal{C}_i$ such that $F_1 = F_L = G_m$ and $\{F_1, \dots, F_L\}$ is the \geq^μ -equivalence class of G_m —hence, it coincides with $\{G_1, \dots, G_M\}$. Therefore $p(\cup_\ell F_\ell) = p(\cup_m G_m) = 1$. Part (ii) of this Lemma then implies that there is $\bar{\ell} \in \{1, \dots, L\}$ such that $p(F_{\bar{\ell}}) > 0$.

I claim that this implies $p(F_\ell) > 0$ for all $\ell = \bar{\ell} + 1, \dots, L$. The claim is trivially true if $\bar{\ell} = L$; otherwise, suppose that $p(F_\ell) > 0$ for some $\ell = \bar{\ell}, \dots, L-1$, and consider $\ell+1$. Since G_1, \dots, G_M μ -supports p and by construction $F_\ell \in \{G_1, \dots, G_M\}$, $p(F_\ell \cap F_{\ell+1}) = \mu(F_\ell \cap F_{\ell+1}|F_\ell)p(F_\ell)$. By the induction hypothesis, $p(F_\ell) > 0$; and since F_1, \dots, F_L is a μ -sequence, $\mu(F_\ell \cap F_{\ell+1}|F_\ell) = \mu(F_{\ell+1}|F_\ell) > 0$.

Thus, $p(F_{\ell+1}) \geq p(F_\ell \cap F_{\ell+1}) > 0$, as claimed. Since $F_L = G_m$, this completes the proof. ■

For any $I \in \mathcal{I}_i$, let $[S_{-i}(I)]_\mu = \bigcup_{J \in \mathcal{I}_i: S_{-i}(J) =^\mu S_{-i}(I)} S_{-i}(J)$: that is, $[S_{-i}(I)]_\mu$ is the union of all elements of the \geq^μ -equivalence class containing $S_{-i}(I)$. Also denote by $[S_{-i}(\mathcal{I}_i)]_\mu$ the range of the map from $S_{-i}(\mathcal{I}_i)$ to $2^{S_{-i}}$ defined by $F \mapsto [F]_\mu$ —that is, the set of all $[S_i(I)]_\mu$, for $I \in \mathcal{I}_i$. The next lemma shows that, if μ can be (uniquely) extended to $[S_{-i}(\mathcal{I}_i)]_\mu$, then it can be (uniquely) extended to all of $S_{-i}(\mathcal{I}_i; \mu)$.

Lemma 3 *Fix a player $i \in N$ and a CPS $\mu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i))$. If there exists $\mu^+ \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i) \cup [S_{-i}(\mathcal{I}_i)]_\mu)$ such that $\mu^+(\cdot|F) = \mu(\cdot|F)$ for all $F \in S_i(\mathcal{I}_i)$, then there exists a unique extension $\nu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i; \mu))$ of μ such that $\nu(\cdot|G) = \mu^+(\cdot|G)$ for all $G \in [S_{-i}(\mathcal{I}_i)]_\mu$.*

Proof: Note first that the last statement in Lemma 2 implies that, for every subset F_1, \dots, F_L of a \geq^μ -equivalence class, $\mu^+(\cup_\ell F_\ell|[F_1]_\mu) > 0$. I shall use this fact throughout this proof without further reference to that Lemma.

Consider two collections $F_1, \dots, F_L \in S_{-i}(\mathcal{I}_i)$ and $G_1, \dots, G_M \in S_{-i}(\mathcal{I}_i)$ such that (i) $F_\ell =^\mu F_1$ for all ℓ , and $G_m =^\mu G_1$ for all m ; and (ii) $\cup_\ell F_\ell = \cup_m G_m$. Since $\mu^+(\cup_\ell F_\ell|[F_1]_\mu) > 0$ and $\mu^+(\cup_m G_m|[G_1]_\mu) > 0$, assumption (ii) implies that also $\mu^+(\cup_m G_m|[F_1]_\mu) > 0$ and $\mu^+(\cup_\ell F_\ell|[G_1]_\mu) > 0$. Since $\{G : G =^\mu G_1\}$ μ -supports $\mu^+(\cdot|[G_1]_\mu)$, and as just noted $\mu^+(\cup_\ell F_\ell|[G_1]_\mu) > 0$, by Lemma 2 part (iii) there is $G \in S_{-i}(\mathcal{I}_i)$ with $G =^\mu G_1$ and $\ell \in \{1, \dots, L\}$ such that $\mu(F_\ell|G) > 0$; thus, $F_\ell \geq^\mu G$, and by transitivity $F_1 \geq^\mu G_1$. By a symmetric argument, $G_1 \geq^\mu F_1$. Therefore, $F_1 =^\mu G_1$, which implies that $[F_1]_\mu = [G_1]_\mu$.

Define an array $\nu \in \Delta(S_{-i}^{S_{-i}(\mathcal{I}_i; \mu)})$ by letting, for every collection $F_1, \dots, F_L \in S_{-i}(\mathcal{I}_i)$ with $F_\ell =^\mu F_m$ for all $\ell, m \in \{1, \dots, L\}$,

$$\nu(E|\cup_\ell F_\ell) = \frac{\mu^+(E \cap (\cup_\ell F_\ell)|[F_1]_\mu)}{\mu^+(\cup_\ell F_\ell|[F_1]_\mu)}. \quad (6)$$

The argument given above shows that, if $G_1, \dots, G_M \in S_{-i}(\mathcal{I}_i)$ satisfies $G_1 =^\mu G_m$ for all m and $\cup_m G_m = \cup_\ell F_\ell$, then $\frac{\mu^+(E \cap (\cup_\ell F_\ell)|[F_1]_\mu)}{\mu^+(\cup_\ell F_\ell|[F_1]_\mu)} = \frac{\mu^+(E \cap (\cup_m G_m)|[G_1]_\mu)}{\mu^+(\cup_m G_m|[G_1]_\mu)}$, so this definition is well-posed.

In addition, for each such collection F_1, \dots, F_M , $\nu(\cup_m F_m | \cup_m F_m) = 1$. Finally, consider $E \subseteq S_{-i}$, $F_1, \dots, F_L \in S_{-i}(\mathcal{I}_i)$ and $G_1, \dots, G_M \in S_{-i}(\mathcal{I}_i)$ such that (i) $F_\ell =^\mu F_1$ for all ℓ , and $G_m =^\mu G_1$ for all m ; and (ii) $E \subseteq \cup_\ell F_\ell \subseteq \cup_m G_m$. If $\nu(\cup_\ell F_\ell | \cup_m G_m) = 0$, then by monotonicity $\nu(E | \cup_m G_m) = 0$ and the chain rule holds. Otherwise, since $0 < \nu(\cup_\ell F_\ell | \cup_m G_m) = \frac{\mu^+(\cup_\ell F_\ell | [G_1]_\mu)}{\mu^+(\cup_m G_m | [G_1]_\mu)}$, so $\mu^+(\cup_\ell F_\ell | [G_1]_\mu) > 0$. Therefore, since the \geq^μ -equivalence class of G_1 μ -supports $\mu^+(\cdot | [G_1]_\mu)$, by Lemma 2 part (iii) there are $\ell \in \{1, \dots, L\}$ and $G \in S_{-i}(\mathcal{I}_i)$ with $G =^\mu G_1$ such that $\mu(F_\ell | G) > 0$, so $F_\ell \geq^\mu G$ and therefore by transitivity $F_1 \geq^\mu G_1$. On the other hand, $\mu^+(\cup_m G_m | [F_1]_\mu) \geq \mu^+(\cup_\ell F_\ell | [F_1]_\mu) > 0$, so a symmetric argument yields $G_1 \geq^\mu F_1$. Therefore $[F_1]_\mu = [G_1]_\mu$, so

$$\begin{aligned} \nu(E | \cup_m G_m) &= \frac{\mu^+(E | [G_1]_\mu)}{\mu^+(\cup_m G_m | [G_1]_\mu)} = \frac{\mu^+(E | [F_1]_\mu)}{\mu^+(\cup_m G_m | [G_1]_\mu)} = \\ &= \frac{\mu^+(E | [F_1]_\mu)}{\mu^+(\cup_\ell F_\ell | [F_1]_\mu)} \frac{\mu^+(\cup_\ell F_\ell | [F_1]_\mu)}{\mu^+(\cup_m G_m | [G_1]_\mu)} = \\ &= \frac{\mu^+(E | [F_1]_\mu)}{\mu^+(\cup_\ell F_\ell | [F_1]_\mu)} \frac{\mu^+(\cup_\ell F_\ell | [G_1]_\mu)}{\mu^+(\cup_m G_m | [G_1]_\mu)} = \nu(E | \cup_\ell F_\ell) \nu(\cup_\ell F_\ell | \cup_m G_m), \end{aligned}$$

so the chain rule holds. Hence ν is a CPS.

Finally, for every $F \in \mathcal{S}_{-i}(I_i)$, the definition of ν for the trivial μ -sequence F yields $\nu(E | F) = \frac{\mu^+(E | [F]_\mu)}{\mu^+(F | [F]_\mu)}$. But by the chain rule, the r.h.s. equals $\mu^+(E | F)$, which by assumption equals $\mu(E | F)$. Thus, ν extends μ .

The uniqueness assertion follows because, if $\nu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i, \mu))$ satisfies $\nu(\cdot | G) = \mu^+(\cdot | G)$ for all $G \in [S_{-i}(\mathcal{I}_i)]_\mu$, then by the chain rule it must also satisfy Eq. (6). ■

Lemma 4 Fix a player $i \in N$ and a CPS $\mu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i))$ with extension $\nu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i; \mu))$.

1. For every $I \in \mathcal{I}_i$, $\nu(S_{-i}(I) | [S_{-i}(I)]_\mu) > 0$.
2. $F \in S_{-i}(\mathcal{I}_i; \mu)$ is basic for ν if and only if $F = [S_{-i}(I)]_\mu$ for some $I \in \mathcal{I}_i$.
3. For every $I, J \in \mathcal{I}_i$, $[S_{-i}(I)]_\mu \geq^\nu [S_{-i}(J)]_\mu$ if and only if $S_{-i}(I) \geq^\mu S_{-i}(J)$.

Notice that parts 2 and 3 of the above Lemma allow one to simplify Definition 7:

Corollary 2 Fix strategies $s_i, t_i \in S_i$. Then $s_i \succ^\mu t_i$ iff and only if, for every $J \in \mathcal{J}_i$ such that $E_{\nu(\cdot|S_{-i}(J))_\mu} U_i(s_i, \cdot) < E_{\nu(\cdot|S_{-i}(J))_\mu} U_i(t_i, \cdot)$, there is $I \in \mathcal{J}_i$ such that $E_{\nu(\cdot|S_{-i}(I))_\mu} U_i(s_i, \cdot) > E_{\nu(\cdot|S_{-i}(I))_\mu} U_i(t_i, \cdot)$ and $S_{-i}(I) \succ^\mu S_{-i}(J)$.

Proof: 1: follows from the last statement of Lemma 2 and the fact that the \geq^μ -equivalence class of $S_{-i}(I)$ supports $\nu(\cdot|S_{-i}(I))_\mu$.

2: Fix $I \in \mathcal{J}_i$ and consider $G \in S_{-i}(\mathcal{J}_i; \mu)$ such that $G \supset [S_{-i}(I)]_\mu$. By construction, $G = S_{-i}(J_1) \cup \dots \cup S_{-i}(J_M)$, where $J_1, \dots, J_M \in \mathcal{J}_i$ and $S_{-i}(J_m) \equiv^\mu S_{-i}(J_n)$ for all $m, n = 1, \dots, M$. I claim that $\nu(S_{-i}(J_m)|\cup_n S_{-i}(J_n)) > 0$ for all m : to see this, observe that $S_{-i}(J_1), \dots, S_{-i}(J_M)$ belong to the same \geq^μ -equivalence class, which may contain additional elements. The union of all elements of the \geq^μ -equivalence class containing $S_{-i}(J_1), \dots, S_{-i}(J_M)$ is then $[S_{-i}(J_1)]_\mu$. By part 1, $\nu(S_{-i}(J_m)|[S_{-i}(J_1)]_\mu) > 0$ for all $m = 1, \dots, M$. Hence $\nu(G|[S_{-i}(J_1)]_\mu) > 0$, so by the chain rule, for every $m = 1, \dots, M$, $0 < \nu(S_{-i}(J_m)|[S_{-i}(J_1)]_\mu) = \nu(S_{-i}(J_m)|G)\nu(G|[S_{-i}(J_1)]_\mu)$, which implies that $\nu(S_{-i}(J_m)|G) > 0$ for all $m = 1, \dots, M$.

By contradiction, suppose that $\nu([S_{-i}(I)]_\mu|G) > 0$. Let $\{S_{-i}(I_1), \dots, S_{-i}(I_L)\}$ be the \geq^μ -equivalence class containing $S_{-i}(I)$, so $[S_{-i}(I)]_\mu = \cup_\ell S_{-i}(I_\ell)$ and $\nu(\cup_\ell S_{-i}(I_\ell)|G) > 0$. Then, by Lemma 2 part (iii), there are $\hat{\ell}$ and \hat{m} such that $\mu(S_{-i}(I_{\hat{\ell}})|S_{-i}(J_{\hat{m}})) > 0$. Therefore $S_{-i}(I_{\hat{\ell}}) \geq^\mu S_{-i}(J_{\hat{m}})$. By transitivity, $S_{-i}(I) \geq^\mu S_{-i}(J_m)$ for all m . Furthermore, since $G \supset [S_{-i}(I)]_\mu$, $\nu(G|[S_{-i}(I)]_\mu) = 1 > 0$, so Lemma 2 part (iii) implies that there are $\tilde{\ell}, \tilde{m}$ such that $\mu(S_{-i}(J_{\tilde{m}})|S_{-i}(I_{\tilde{\ell}})) > 0$, and so $S_{-i}(J_{\tilde{m}}) \geq^\mu S_{-i}(I_{\tilde{\ell}})$. Again, transitivity implies that $S_{-i}(J_m) \geq^\mu S_{-i}(I)$ for all m . Therefore, $G \subseteq [S_{-i}(I)]_\mu$, contradiction. Hence, $\nu([S_{-i}(I)]_\mu|G) = 0$. Since $G \in S_{-i}(\mathcal{J}_i; \mu)$ was an arbitrary superset, $[S_{-i}(I)]_\mu$ is basic.

Conversely, consider $F \in S_{-i}(\mathcal{J}_i; \mu)$. By definition $F = S_{-i}(I_1) \cup \dots \cup S_{-i}(I_L)$, where $S_{-i}(I_\ell) \equiv^\mu S_{-i}(I_m)$ for all $\ell, m \in \{1, \dots, L\}$. Suppose that there is $I \in \mathcal{J}_i$ such that $S_{-i}(I) \equiv^\mu S_{-i}(I_1)$. Let $G = F \cup S_{-i}(I)$. By part 1, $\nu(S_{-i}(I)|[S_{-i}(I)]_\mu) > 0$ and $\nu(S_{-i}(I_1)|[S_{-i}(I)]_\mu) > 0$, so also $\nu(F|[S_{-i}(I)]_\mu) > 0$ and $\nu(G|[S_{-i}(I)]_\mu) > 0$. By the chain rule, $\nu(F|[S_{-i}(I)]_\mu) = \nu(F|G)\nu(G|[S_{-i}(I)]_\mu)$; but then, $\nu(F|G) > 0$. Since $G \supset F$, F is not basic for ν .

3: assume first that $S_{-i}(I) \geq^\mu S_{-i}(J)$. Since $[S_{-i}(I)]_\mu \supseteq S_{-i}(I)$, $\nu([S_{-i}(I)]_\mu | S_{-i}(I)) > 0$, so $[S_{-i}(I)]_\mu \geq^\nu S_{-i}(I)$. By part 1, $\nu(S_{-i}(J) | [S_{-i}(J)]_\mu) > 0$, so $S_{-i}(J) \geq^\nu [S_{-i}(J)]_\mu$. Finally, since $S_{-i}(\mathcal{I}_i) \subseteq S_{-i}(\mathcal{I}_i; \mu)$, $S_{-i}(I) \geq^\mu S_{-i}(J)$ implies $S_{-i}(I) \geq^\nu S_{-i}(J)$. By transitivity, $[S_{-i}(I)]_\mu \geq^\nu [S_{-i}(J)]_\mu$.

Conversely, assume that $[S_{-i}(I)]_\mu \geq^\nu [S_{-i}(J)]_\mu$. Let $F_1, \dots, F_L \in S_{-i}(\mathcal{I}_i; \mu)$ be a ν -sequence with $F_1 = [S_{-i}(J)]_\mu$ and $F_L = [S_{-i}(I)]_\mu$. By definition, for every $\ell = 2, \dots, L-1$ there is a collection $I_1^\ell, \dots, I_{M(\ell)}^\ell \in \mathcal{I}_i$ such that $S_{-i}(I_m^\ell) =^\mu S_{-i}(I_n^\ell)$ for all $m, n \in \{1, \dots, M(\ell)\}$, and $F_\ell = \cup_m S_{-i}(I_m^\ell)$. For $\ell = 1$ (resp. $\ell = L$), we can take $I_1^\ell, \dots, I_{M(\ell)}^\ell$ to be such that $\{S_{-i}(I_1^\ell), \dots, S_{-i}(I_{M(\ell)}^\ell)\}$ is the \geq^μ -equivalence class containing $S_{-i}(J)$ (resp. $S_{-i}(I)$). Finally, for each $\ell = 1, \dots, L-1$, $\nu(\cup_{m=1}^{M(\ell+1)} S_{-i}(I_m^{\ell+1}) | \cup_{m=1}^{M(\ell)} S_{-i}(I_m^\ell)) > 0$. Part (iii) of Lemma 2 then yields $m(\ell) \in \{1, \dots, M(\ell)\}$ for each $\ell = 1, \dots, L-1$, and $n(\ell) \in \{1, \dots, M(\ell)\}$ for each $\ell = 2, \dots, L$, such that $\mu(S_{-i}(I_{n(\ell+1)}^{\ell+1}) | S_{-i}(I_{m(\ell)}^\ell)) > 0$ for each $\ell = 1, \dots, L-1$; thus, $S_{-i}(I_{n(\ell+1)}^{\ell+1}) \geq^\mu S_{-i}(I_{m(\ell)}^\ell)$ for each $\ell = 1, \dots, L-1$. By assumption, $S_{-i}(I_{m(\ell)}^\ell) =^\mu S_{-i}(I_{n(\ell)}^\ell)$ for each $\ell = 2, \dots, L-1$. Furthermore, $S_{-i}(I) =^\mu S_{-i}(I_{n(L)}^L)$ and $S_{-i}(J) =^\mu S_{-i}(I_{m(1)}^1)$. Summing up,

$$S_{-i}(I) =^\mu S_{-i}(I_{n(L)}^L) \geq^\mu S_{-i}(I_{m(L-1)}^{L-1}) =^\mu S_{-i}(I_{n(L-1)}^{L-1}) \geq^\mu S_{-i}(I_{m(L-2)}^{L-2}) =^\mu S_{-i}(I_{n(L-2)}^{L-2}) \geq^\mu \dots \geq^\mu S_{-i}(I_{m(1)}^1) =^\mu S_{-i}(J).$$

By transitivity, $S_{-i}(I) \geq^\mu S_{-i}(J)$. ■

The following result is key to characterizing extensible CPSs.

Proposition 1 Fix a non-empty collection $\mathcal{C}_i \subseteq 2^{S_{-i}} \setminus \{\emptyset\}$ and a CPS $\mu \in \Delta(S_{-i}, \mathcal{C}_i)$ for player $i \in N$. The following are equivalent:

1. μ is congruent;
2. for every μ -sequence $F_1, \dots, F_K \in \mathcal{C}_i$, there exists $p \in \Delta(S_{-i})$ with $p(\cup_k F_k) = 1$, such that for every $\ell = 1, \dots, K$ and $E \subseteq F_\ell$,

$$p(E) = \mu(E | F_\ell) p(F_\ell). \quad (7)$$

If a probability p that satisfies the property in (2) exists, it is unique; furthermore, $p(F_K) > 0$, and for all $\ell = 1, \dots, K-1$, $p(F_\ell) > 0$ iff $\mu(F_k | F_{k+1}) > 0$ for all $k = \ell + 1, \dots, K$.

Note that, in part 2, the μ -sequence F_1, \dots, F_K μ -supports p .

Proof: (1) \Rightarrow (2): assume that μ is congruent. Let $F_1, \dots, F_K \in \mathcal{C}_i$ be a μ -sequence.

Define $G_1 = F_1$ and, inductively, $G_k = F_k \setminus (F_1 \cup \dots \cup F_{k-1})$ for $k = 2, \dots, K$. Note that $F_1 \cup \dots \cup F_k = G_1 \cup \dots \cup G_k$ for all $k = 1, \dots, K$, [for $k = 1$ this is by definition. By induction, $G_1 \cup \dots \cup G_{k+1} = (G_1 \cup \dots \cup G_k) \cup G_{k+1} = (F_1 \cup \dots \cup F_k) \cup G_{k+1} = (F_1 \cup \dots \cup F_k) \cup [F_{k+1} \setminus (F_1 \cup \dots \cup F_k)] = F_1 \cup \dots \cup F_{k+1}$] and $G_k \cap G_\ell = \emptyset$ for all $k \neq \ell$. [Let $\ell > k$: then $G_\ell = F_\ell \setminus (F_1 \cup \dots \cup F_{\ell-1}) = F_\ell \setminus (G_1 \cup \dots \cup G_{\ell-1})$, and $k \in \{1, \dots, \ell - 1\}$.] Also, $G_k \subseteq F_k$ for all $k = 1, \dots, K$.

I now define a set function $\rho : 2^{S_i} \rightarrow \mathbb{R}$. For every $\ell = 1, \dots, K$ and $E \subseteq S_i$ with $E \subseteq G_\ell$, let

$$\rho(E) \equiv \mu(E|F_\ell) \cdot \prod_{k=\ell}^{K-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)},$$

with the usual convention that the product over an empty set of indices equals 1. By assumption, the denominators of the above fractions are all strictly positive. Also, since the sets G_1, \dots, G_k are disjoint by construction, if $\emptyset \neq E \subseteq G_\ell$ for some ℓ then $E \not\subseteq G_k$ for $k \neq \ell$, so $\rho(E)$ is uniquely defined; furthermore, $\emptyset \subseteq G_k$ for all k , but $\rho(\emptyset)$ is still well-defined and equal to 0.

To complete the definition of $\rho(\cdot)$, for all events $E \subseteq S_i$ such that $E \not\subseteq G_k$ for $k = 1, \dots, K$ [i.e., E intersects two or more events G_k , or none], let

$$\rho(E) = \sum_{k=1}^K \rho(E \cap G_k).$$

The function $\rho(\cdot)$ thus defined takes non-negative values. I claim that $\rho(\cdot)$ is additive. Consider an ordered list $E_1, \dots, E_M \subseteq S_i$ such that $E_m \cap E_{\bar{m}} = \emptyset$ for $m \neq \bar{m}$. If there is $\ell \in \{1, \dots, K\}$ such that $E_m \subseteq G_\ell$ for all m , then by additivity of $\mu(\cdot|F_\ell)$,

$$\begin{aligned} \rho\left(\bigcup_m E_m\right) &= \mu\left(\bigcup_m E_m \middle| F_\ell\right) \cdot \prod_{k=\ell}^{K-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)} = \left(\sum_m \mu(E_m|F_\ell)\right) \cdot \prod_{k=\ell}^{K-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)} = \\ &= \sum_m \left(\mu(E_m|F_\ell) \cdot \prod_{k=\ell}^{K-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)}\right) = \sum_m \rho(E_m). \end{aligned}$$

Thus, for a general ordered list $E_1, \dots, F_M \subseteq S_{-i}$ of pairwise disjoint events,²²

$$\begin{aligned} \rho\left(\bigcup_m E_m\right) &= \sum_k \rho\left(\left[\bigcup_m E_m\right] \cap G_k\right) = \sum_k \rho\left(\bigcup_m [E_m \cap G_k]\right) = \\ &= \sum_k \sum_m \rho(E_m \cap G_k) = \sum_m \sum_k \rho(E_m \cap G_k) = \sum_m \rho(E_m). \end{aligned}$$

Now consider $E \subseteq S_{-i}$ with $E \subseteq F_m$ and $E \subseteq G_\ell$ for some $\ell, m \in \{1, \dots, K\}$ with $\ell \neq m$. Since $F_m \subseteq F_1 \cup \dots \cup F_m = G_1 \cup \dots \cup G_m$, it must be the case that $\ell < m$. Consider the ordered list $F_\ell, \dots, F_m \in \mathcal{C}_i$: since F_1, \dots, F_K is a μ -sequence, so is F_ℓ, \dots, F_m , so by congruence, since by assumption $E \subseteq F_m \cap G_\ell \subseteq F_m \cap F_\ell$,

$$\mu(E|F_\ell) \prod_{k=\ell}^{m-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)} = \mu(E|F_m).$$

Multiply both sides by the positive quantity $\prod_{k=m}^{K-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)}$ to get

$$\rho(E) = \mu(E|F_\ell) \prod_{k=\ell}^{K-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)} = \mu(E|F_m) \prod_{k=m}^{K-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)}.$$

Therefore, for all $E \subseteq S_{-i}$ with $E \subseteq F_m$ for some $m \in \{1, \dots, K\}$,

$$\begin{aligned} \mu(E|F_m) \prod_{k=m}^{K-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)} &= \sum_{\ell=1}^K \mu(E \cap G_\ell|F_m) \prod_{k=m}^{K-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)} = \\ &= \sum_{\ell=1}^K \rho(E \cap G_\ell) = \rho(E). \end{aligned}$$

It follows that, for all $m \in \{1, \dots, K\}$ and $E \subseteq S_{-i}$ with $E \subseteq F_m$,

$$\rho(F_m) = \mu(F_m|F_m) \prod_{k=m}^{K-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)} = \prod_{k=m}^{K-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)}, \quad (8)$$

²²For future reference, if the set S_{-i} is an arbitrary measurable space, and probabilities in i 's CPS are countably additive, this step of the proof still holds, and shows that ρ is countably additive. Specifically, the derivation holds as written for a countable collection F_1, F_2, \dots (I purposely omitted limits from the summations). In particular, interchanging the order of the summation in the second line is allowed because all summands are non-negative and the derivation shows that $\sum_k \sum_m \rho(E_m \cap G_k) = \sum_k \rho([\cup_m E_m] \cap G_k)$, a sum of finitely many finite terms.

and therefore

$$\rho(E) = \mu(E|F_m)\rho(F_m). \quad (9)$$

Finally, notice that $\rho(\cup_k G_k) = \rho(\cup_k F_k) \geq \rho(F_K) = 1$; thus, one can define a probability measure $p \in \Delta(S_{-i})$ by letting

$$\forall E \subseteq S_{-i}, \quad p(E) = \frac{\rho(E)}{\rho(\cup_k G_k)} = \frac{\rho(E)}{\rho(\cup_k F_k)}.$$

For every $\ell \in \{1, \dots, K\}$ and every event $E \subseteq F_\ell$, p satisfies Eq. (7), as asserted.

To show that p is uniquely defined, let $q \in \Delta(S_{-i})$ be a measure that satisfies Eq.(7). I first claim that, for every $m = 1, \dots, K$,

$$q(F_m) = \prod_{k=m}^{K-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)} \cdot q(F_N) = \rho(F_m)q(F_K).$$

The claim is trivially true for $m = K$, so consider $m \in \{1, \dots, K-1\}$ and assume that the claim holds for $m+1$. By Eq.(7),

$$\mu(F_m \cap F_{m+1}|F_{m+1})q(F_{m+1}) = q(F_m \cap F_{m+1}) = \mu(F_m \cap F_{m+1}|F_m)q(F_m);$$

since $\mu(F_m \cap F_{m+1}|F_m) > 0$ by assumption, solving for $q(F_m)$ and invoking the inductive hypothesis yields

$$q(F_m) = \frac{\mu(F_m \cap F_{m+1}|F_{m+1})}{\mu(F_m \cap F_{m+1}|F_m)} q(F_{m+1}) = \frac{\mu(F_m \cap F_{m+1}|F_{m+1})}{\mu(F_m \cap F_{m+1}|F_m)} \cdot \prod_{k=m+1}^{K-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)} \cdot q(F_K) = \prod_{k=m}^{K-1} \frac{\mu(F_k \cap F_{k+1}|F_{k+1})}{\mu(F_k \cap F_{k+1}|F_k)} \cdot q(F_K).$$

Since $G_m \subseteq F_m$, Eq. (7) implies that

$$q(G_m) = \mu(G_m|F_m)q(F_m) = \mu(G_m|F_m) \cdot \rho(F_m) \cdot q(F_K) = \rho(G_m) \cdot q(F_K),$$

where the last equality follows from Eq.(9). Since $\sum_k q(G_k) = q(\cup_k G_k) = q(\cup_k F_k)$, if in addition q satisfies $q(\cup_k F_k) = 1$, then

$$1 = \sum_m \rho(G_m) \cdot q(F_K) = q(F_K)\rho(\cup_m G_m)$$

which immediately implies that $q(F_K) > 0$, and indeed that

$$q(F_K) = \frac{1}{\rho(\cup_m G_m)} = \frac{\rho(F_K)}{\rho(\cup_m G_m)} = p(F_K).$$

so also $p(F_K) > 0$, as claimed. Furthermore, for $m = 1, \dots, K-1$,

$$q(F_m) = \rho(F_m)q(F_N) = \rho(F_m) \frac{1}{\rho(\cup_m G_m)} = p(F_m).$$

Furthermore, let $k_0 \in \{1, \dots, K-1\}$ be such that $\mu(F_k \cap F_{k+1} | F_{k+1}) > 0$ for all $k > k_0$, and $\mu(F_{k_0} \cap F_{k_0+1} | F_{k_0+1}) = 0$. By inspecting Eq. (8), it is clear that $\rho(F_k) = 0$ for $k = 1, \dots, k_0$, and $\rho(F_k) > 0$ for $k = k_0+1, \dots, K$. Then, $p(F_k) = 0$ for $k = 1, \dots, k_0$, and $p(F_k) > 0$ for $k = k_0+1, \dots, K$. From the above argument, it follows that the same is true for any $q \in \Delta(S_{-i})$ that satisfies Eq. (7) and $q(\cup_k F_k) = 1$. Thus, the last claim of the Proposition follows.

Finally, if $q \in \Delta(S_{-i})$ satisfies Eq.(7) and $q(\cup_k F_k) = 1$, for every $k = k_0+1, \dots, K$ and $E \subseteq S_{-i}$ such that $E \subset F_k$,

$$q(E) = \mu(E|F_k)q(F_k) = \mu(E|F_k)p(F_k) = p(E)$$

and therefore, for every $E \subseteq S_{-i}$,

$$q(E) = \sum_k q(E \cap G_k) = \sum_{k=k_0+1}^K q(E \cap G_k) = \sum_{k=k_0+1}^K p(E \cap G_k) = \sum_k p(E \cap G_k) = p(E).$$

In other words, p is the unique probability measure that satisfies Eq. (7) and $p(\cup_k F_k) = 1$.

(2) \Rightarrow (1): assume that (2) holds. Consider a μ -sequence F_1, \dots, F_K . Fix an event $E \subseteq F_1 \cap F_K$. By assumption, there exists $p \in \Delta(S_{-i})$ that satisfies Eq. (7) for $k = 1, \dots, K$, with $p(\cup_{k=1}^K F_k) = 1$.

Since $p(F_K) > 0$, $\mu(E|F_K) = \frac{p(E)}{p(F_K)}$. If $p(F_1) = 0$, then a fortiori $p(E) = 0$, so $\mu(E|F_K) = 0$; on the other hand, $p(F_1) = 0$ implies that there is $k = 1, \dots, K-1$ such that $\mu(F_k \cap F_{k+1} | F_{k+1}) = 0$, so

$$\mu(E|F_1) \cdot \prod_{k=1}^{K-1} \frac{\mu(F_k \cap F_{k+1} | F_{k+1})}{\mu(F_k \cap F_{k+1} | F_k)} = \mu(E|F_1) \cdot 0 = 0 = \mu(E|F_K).$$

If instead $p(F_1) > 0$, then $\mu(E|F_1) = \frac{p(E)}{p(F_1)}$; furthermore, by the above argument $p(F_k) > 0$ for all $k = 2, \dots, K-1$ as well, so

$$\mu(E|F_1) \cdot \prod_{k=1}^{K-1} \frac{\mu(F_k \cap F_{k+1} | F_{k+1})}{\mu(F_k \cap F_{k+1} | F_k)} = \frac{p(E)}{p(F_1)} \cdot \prod_{k=1}^{K-1} \frac{p(F_k \cap F_{k+1})}{p(F_{k+1})} \cdot \frac{p(F_k)}{p(F_k \cap F_{k+1})} = \frac{p(E)}{p(F_1)} \cdot \frac{p(F_1)}{p(F_K)} = \frac{p(E)}{p(F_K)} = \mu(E|F_K).$$

■

Corollary 3 *If μ is congruent, then for every μ -sequence F_1, \dots, F_K such that $\mu(F_1|F_K) > 0$, the reverse-ordered list F_K, F_{K-1}, \dots, F_1 is also a μ -sequence: that is, $\mu(F_k|F_{k+1}) > 0$ for all $k = 1, \dots, K-1$.*

In particular, this Corollary applies if $F_1 = F_K$.

Proof: Let F_1, \dots, F_K be as in the statement, and consider the ordered list F_1, \dots, F_K, F_{K+1} with $F_{k+1} = F_1$. Then F_1, \dots, F_{K+1} is also a μ -sequence. Let p be the unique measure in (2) of Proposition 1. The last claim of that Proposition shows that necessarily $p(F_{K+1}) > 0$, but since $F_{K+1} = F_1$, also $p(F_1) > 0$. Again, the last claim in Proposition 1 implies that then $p(F_k) > 0$ for all $k = 1, \dots, K$.

Then, for all $k = 1, \dots, K-1$, $\mu(F_k \cap F_{k+1}|F_k) > 0$ implies that $p(F_k \cap F_{k+1}) > 0$, and so

$$\mu(F_k|F_{k+1}) = \mu(F_k \cap F_{k+1}|F_{k+1}) = \frac{p(F_k \cap F_{k+1})}{p(F_{k+1})} > 0.$$

■

Corollary 4 *Let G_1, \dots, G_N be a μ -sequence and p the measure in (2) of Proposition 1; consider $F \in S_{-i}(\mathcal{G}_i)$ such that $F \subset \cup_{k=1}^K G_k$. Then, for every $E \subseteq F$, $p(E) = \mu(E|F)p(F)$.*

Proof: It is enough to consider the case $p(F) > 0$.

Let $k \in \{1, \dots, K\}$ be such that $p(G_k) > 0$ and $\mu(F|G_k) = \mu(F \cap G_k|G_k) > 0$. One such k must exist, because $p(F) > 0$ implies $p(F \cap G_m) > 0$ for some $m \in \{1, \dots, K\}$, and by construction $p(F \cap G_m) = p(G_m)\mu(F \cap G_m|G_m)$.

I claim that, for any such k , $\mu(G_k|F) > 0$. Since $F \subseteq \cup_m G_m$ and $\mu(F|F) = 1$, $\mu(G_m|F) > 0$ for at least one $m \in \{1, \dots, K\}$. If $m = k$, the claim is true. If $m < k$, then the ordered list $F, G_m, G_{m+1}, \dots, G_k, F$ is a μ -sequence that satisfies the conditions of Corollary 3, so that in

particular $\mu(G_k|F) > 0$, as claimed. Finally, suppose $m > k$. Since $p(G_k) > 0$, by the last claim of Proposition 1, $\mu(G_\ell|G_{\ell+1}) > 0$ for $\ell = k, \dots, K-1$. Hence, since $\mu(G_m|F) > 0$, the ordered list $F, G_m, G_{m-1}, \dots, G_{k+1}, G_k, F$ is a μ -sequence that satisfies the conditions in Corollary 3, so in particular $\mu(G_k|F) > 0$, as claimed.

This implies that the ordered list $G_1, \dots, G_k, F, G_k, \dots, G_K$ is a μ -sequence. Let p' be the measure delivered by Proposition 1 for this μ -sequence. Notice that $p(F \cup \bigcup_k G_k) = p'(F \cup \bigcup_k G_k) = 1$, and for all $\ell \in \{1, \dots, K\}$ and $E \subseteq S_{-i}$ with $E \subseteq G_\ell$, $p'(E) = p'(G_\ell)\mu(E|G_\ell)$. Since p is the unique probability with these properties, $p = p'$. But then, for $E \subseteq S_{-i}$ with $E \subseteq F$,

$$p(E) = p'(E) = p'(F)\mu(E|F) = p(F)\mu(E|F),$$

as claimed. ■

Corollary 5 *Let G_1, \dots, G_K and F_1, \dots, F_M be μ -sequences with $\cup_m F_m \subseteq \cup_k G_k$. Let p and q be the probabilities associated with G_1, \dots, G_K and F_1, \dots, F_M respectively. Consider $E \subseteq \cup_m F_m$. Then $p(E) = p(\cup_m F_m)q(E)$.*

Proof: It is enough to consider the case $p(\cup_m F_m) > 0$.

Since, for every m , $F_m \subseteq \cup_k G_k$, Corollary 4 implies that, for every $E' \subseteq S_{-i}$ with $E' \subseteq F_m$,

$$p(E') = \mu(E'|F_m)p(F_m).$$

Hence, the measure $p' \in \Delta(S_{-i})$ defined by $p'(E) = p(E \cap \cup_m F_m)/p(\cup_m F_m)$ satisfies

$$\forall E' \subseteq S_{-i}, E' \subseteq F_m, \quad p(E') = \mu(E'|F_m)p'(F_m) \quad \text{and} \quad p'(\cup_m F_m) = 1.$$

Therefore, $p' = q$, or $p(E') = p(\cup_m F_m)q(E')$ for every m and $E' \subseteq S_{-i}$ with $E' \subseteq F_m$. In particular, let $\bar{F}_1 = F_1$ and, for $m = 2, \dots, M$, let $\bar{F}_m = F_m \setminus (F_1 \cup \dots \cup F_{m-1})$. Then, for every m ,

$$p(E \cap \bar{F}_m) = p(\cup_\ell F_\ell)q(E \cap \bar{F}_m)$$

and so, since $\bar{F}_1, \dots, \bar{F}_M$ is a partition of $\cup_m F_m$ and $E \subseteq \cup_m F_m$, summing over all m yields $p(E) = p(\cup_m F_m)q(E)$, as required. ■

Finally, I prove Theorem 3.

Proof: I show (1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (1).

(1) \Rightarrow (2): by Lemma 3, it is enough to show that μ can be uniquely extended to a CPS $\mu^+ \in \Delta(S_{-i}, S_{-i}(\mathcal{S}_i) \cup [S_{-i}(\mathcal{S}_i)]_\mu)$. Let $\{F_1, \dots, F_L\}$ be a \geq^μ -equivalence class. By Lemma 1, there is a μ -sequence G_1, \dots, G_M , with $G_1 = G_M = F_1$. By Proposition 1, there is a unique $p \in \Delta(S_{-i})$ such that G_1, \dots, G_M μ -supports p —that is, $p(\cup_m G_m) = 1$ and $p(E) = \mu(E|G_m)p(G_m)$ for all m and all $E \subseteq G_m$. Furthermore, there is $\bar{m} \in \{1, \dots, M\}$ such that $m \geq \bar{m}$ if and only if $p(G_m) > 0$; since surely $p(G_M) > 0$ and $G_1 = G_M = F_1$, $\bar{m} = 1$. Therefore $p(G_m) > 0$ for all m , so $p(F_\ell) > 0$ for all ℓ . Finally, suppose that $\bar{F}_1, \dots, \bar{F}_L$ is another \geq^μ -equivalence class, with $[F_1]_\mu = [\bar{F}_1]_\mu$. Again, by Lemma 1, there is a μ -sequence $\bar{G}_1, \dots, \bar{G}_M$ with $\{\bar{G}_1, \dots, \bar{G}_M\} = \{\bar{F}_1, \dots, \bar{F}_L\}$, and Corollary 5 implies that $\bar{F}_1, \dots, \bar{F}_L$ μ -supports the same probability p . Therefore, one can define an array $\mu^+ \in \Delta(S_{-i}^{S_{-i}(\mathcal{S}_i) \cup [S_{-i}(\mathcal{S}_i)]_\mu})$ by letting $\mu^+(\cdot|F) = \mu(\cdot|F)$ for $F \in S_{-i}(\mathcal{S}_i)$, and $\mu^+(\cdot|[F]_\mu) = p$, where p is the unique probability that μ -support the \geq^μ -equivalence class containing $F \in S_{-i}(\mathcal{S}_i)$.

It remains to be shown that μ^+ is a CPS. Its construction immediately implies that $\mu^+(G|G) = 1$ for all $G \in S_{-i}(\mathcal{S}_i) \cup [S_{-i}(\mathcal{S}_i)]_\mu$. To show that the chain rule holds, consider $F, G \in S_{-i}(\mathcal{S}_i) \cup [S_{-i}(\mathcal{S}_i)]_\mu$ and $E \subseteq F$. If $F, G \in S_{-i}(\mathcal{S}_i)$, then the conclusion follows from the fact that $\mu^+(\cdot|F) = \mu(\cdot|F)$ and $\mu^+(\cdot|G) = \mu(\cdot|G)$, because μ is a CPS. If $F = [S_{-i}(I)]_\mu$ and $G = [S_{-i}(J)]_\mu$, and $F \neq G$ (otherwise the conclusion is immediate), then I claim that $\mu^+([S_{-i}(I)]_\mu|[S_{-i}(J)]_\mu) = 0$, which again that the chain rule holds. To see this, let F_1, \dots, F_L and G_1, \dots, G_M be the \geq^μ -equivalence classes containing $S_{-i}(I)$ and $S_{-i}(J)$ respectively. If $\mu^+(\cup_\ell F_\ell|\cup_m G_m) > 0$, then Lemma 2 part (iii) implies that $\mu(F_{\bar{\ell}}|G_{\bar{m}}) > 0$ for some $\bar{\ell}, \bar{m}$, so $F_{\bar{\ell}} \geq^\mu G_{\bar{m}}$ and thus, by transitivity, $S_{-i}(I) \geq^\mu S_{-i}(J)$. Since $\mu^+(\cup_m G_m|\cup_\ell F_\ell) \geq \mu^+(\cup_\ell F_\ell|\cup_\ell F_\ell) = 1$, a symmetric argument shows that $S_{-i}(J) \geq^\mu S_{-i}(I)$, so $S_{-i}(I) =^\mu S_{-i}(J)$ and thus $F = [S_{-i}(I)]_\mu = [S_{-i}(J)]_\mu = G$, contradiction. This proves the claim.

Finally, consider the case of $F = S_{-i}(I)$ and $G = [S_{-i}(J)]_\mu$. Lemma 2 implies that there is \bar{m} with $\mu(F|G_{\bar{m}}) > 0$, so $F \geq^\mu G_{\bar{m}}$ and thus $F \geq^\mu S_{-i}(J)$. Since $\mu^+(G|F) \geq \mu^+(F|F) = 1$, the same Lemma implies that there is G' with $G' \equiv^\mu G$ and $\mu(G'|F) > 0$, so $G' \geq^\mu F$ and $S_{-i}(J) \geq^\mu F$. Therefore, F is an element of the μ -equivalence class for $S_{-i}(J)$. But then, since this collection of events supports $\mu^+(\cdot|G)$, $\mu^+(E|G) = \mu(E|F)\mu^+(F|G) = \mu^+(E|F)\mu^+(F|G)$, as required.

(2) \Rightarrow (3): fix the unique extension $\nu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i; \mu))$ of μ . The probabilities $\{\nu(\cdot|[F]_\mu) : F \in S_{-i}(\mathcal{I}_i)\}$ can be partially ordered as follows: $\nu(\cdot|[F]_\mu) \geq \nu(\cdot|[G]_\mu)$ iff $F \geq^\mu G$. [The ordering is clearly reflexive and transitive because so is \geq^μ . To see that it is antisymmetric, if $\nu(\cdot|[F]_\mu) \geq \nu(\cdot|[G]_\mu)$ and $\nu(\cdot|[G]_\mu) \geq \nu(\cdot|[F]_\mu)$, then $F \geq^\mu G$ and $G \geq^\mu F$, i.e. $F \equiv^\mu G$; but then $[F]_\mu = [G]_\mu$ and so $\nu(\cdot|[F]_\mu) = \nu(\cdot|[G]_\mu)$.]

Let p_1, \dots, p_L be an enumeration of $\{\nu(\cdot|[F]_\mu) : F \in S_{-i}(\mathcal{I}_i)\}$ such that, for all ℓ, m , $p_\ell \geq p_m$ implies $\ell \leq m$. [This can be obtained by considering any completion of the partial order \geq , and assigning indices consistently with this completion, with $\ell = 1$ being the greatest element.] For every $F \in S_{-i}(\mathcal{I}_i)$, let $\ell(F)$ denote the index ℓ such that $p_\ell = \nu(\cdot|[F]_\mu)$. Finally, define a sequence $(p^n) \subset \Delta(S_{-i})$ by letting

$$p^n = \sum_{\ell=1}^L \frac{\frac{1}{n^{\ell-1}}}{\sum_{m=1}^L \frac{1}{n^{m-1}}} p_\ell.$$

For every $n \geq 1$ and $F \in S_{-i}(\mathcal{I}_i)$, $p_{\ell(F)}(F) = \nu(F|[F]_\mu)(F) > 0$ by the last statement in Lemma 2, and so $p^n(F) > 0$. Furthermore, consider $F \in S_{-i}(\mathcal{I}_i)$ and an event $E \subseteq F$. Suppose there is $G \in S_{-i}(\mathcal{I}_i)$ such that $\nu(E|[G]_\mu) > 0$; then $\nu(F|[G]_\mu) > 0$, so by Lemma 2 part (iii) there is $G' \in S_{-i}(\mathcal{I}_i)$ with $G' \equiv^\mu G$ and $\mu(F|G') > 0$, so $F \geq^\mu G'$. By transitivity, $F \geq^\mu G$. Hence, $\nu(\cdot|[F]_\mu) \geq \nu(\cdot|[G]_\mu)$, so either $\nu(\cdot|[G]_\mu) = \nu(\cdot|[F]_\mu)$, or $\ell(F) < \ell(G)$. Thus,

$$p^n(E) = \sum_{\ell=\ell(F)}^L \frac{\frac{1}{n^{\ell-1}}}{\sum_{m=1}^L \frac{1}{n^{m-1}}} p_\ell(E).$$

This holds in particular for $E = F$. Hence,

$$\begin{aligned} \frac{p^n(E)}{p^n(F)} &= \frac{\sum_{\ell=\ell(F)}^L \frac{\frac{1}{n^{\ell-1}}}{\sum_{m=1}^L \frac{1}{n^{m-1}}} p_\ell(E)}{\sum_{\ell=\ell(F)}^L \frac{\frac{1}{n^{\ell-1}}}{\sum_{m=1}^L \frac{1}{n^{m-1}}} p_\ell(F)} = \frac{\sum_{\ell=\ell(F)}^L \frac{1}{n^{\ell-1}} p_\ell(E)}{\sum_{\ell=\ell(F)}^L \frac{1}{n^{\ell-1}} p_\ell(F)} = \frac{n^{\ell(F)-1} \sum_{\ell=\ell(F)}^L \frac{1}{n^{\ell-\ell(F)}} p_\ell(E)}{n^{\ell(F)-1} \sum_{\ell=\ell(F)}^L \frac{1}{n^{\ell-\ell(F)}} p_\ell(F)} = \\ &= \frac{p_{\ell(F)}(E) + \sum_{\ell=\ell(F)+1}^L \frac{1}{n^{\ell-\ell(F)}} p_\ell(E)}{p_{\ell(F)}(F) + \sum_{\ell=\ell(F)+1}^L \frac{1}{n^{\ell-\ell(F)}} p_\ell(F)} \rightarrow \frac{p_{\ell(F)}(E)}{p_{\ell(F)}(F)} = \frac{\nu(E|[F]_\mu)}{\nu(F|[F]_\mu)} = \mu(E|F). \end{aligned}$$

(3) \Rightarrow (1): consider a μ -sequence F_1, \dots, F_L and an event $E \subseteq F_1 \cap F_L$. Let $(p^n) \subseteq \Delta(S_{-i})$ generate μ in the sense of condition (3). Since $\mu(F_{\ell+1}|F_\ell) > 0$ for all $\ell = 1, \dots, L-1$, there is \bar{n} such that $n \geq \bar{n}$ implies $p^n(F_{\ell+1} \cap F_\ell)/p(F_\ell) > 0$. For every such n and event $E \subseteq F_1 \cap F_L$,

$$\frac{p^n(E)}{p^n(F_1)} \cdot \prod_{\ell=1}^{L-1} \frac{p^n(F_\ell \cap F_{\ell+1})}{p^n(F_\ell)} = \frac{p^n(E)}{p^n(F_1)} \cdot \prod_{\ell=1}^{L-1} \frac{p^n(F_\ell)}{p^n(F_{\ell+1})} = \frac{p^n(E)}{p^n(F_L)}$$

Since $p^n(E)/p^n(F_1) \rightarrow \mu(E|F_1)$, $p^n(F_\ell \cap F_{\ell+1})/p^n(F_{\ell+1}) \rightarrow \mu(F_\ell \cap F_{\ell+1}|F_{\ell+1})$, $p^n(F_\ell \cap F_{\ell+1})/p^n(F_\ell) \rightarrow \mu(F_\ell \cap F_{\ell+1}|F_\ell) > 0$, and $p^n(E)/p^n(F_L) \rightarrow \mu(E|F_L)$, it follows that Congruence holds. ■

C Appendix: Nested strategic information

The equivalence of Definition 6 with the general one, Definition 7, follows from the next lemma.

Lemma 5 *Consider a dynamic game with nested strategic information. For every $i \in N$ and $J \in \mathcal{I}_i$, there is a unique $I \in \mathcal{I}_i$ such that $S_{-i}(I) \in S_{-i}(\mathcal{I}_i)$ is basic for μ , $S_{-i}(I) \supseteq S_{-i}(J)$, and $\mu(S_{-i}(J)|S_{-i}(I)) > 0$. Furthermore, $S_{-i}(I) = [S_{-i}(J)]_\mu$.*

Proof: Let \mathcal{J} the collection of information sets $\hat{I} \in \mathcal{I}_i$ such that $S_{-i}(J) \subseteq S_{-i}(\hat{I})$. Since $S_{-i}(\hat{I}) \cap S_{-i}(\hat{I}') \supseteq S_{-i}(J) \neq \emptyset$ for every pair $\hat{I}, \hat{I}' \in \mathcal{J}$, by nested strategic information either $S_{-i}(\hat{I}) \supseteq S_{-i}(\hat{I}')$ or $S_{-i}(\hat{I}') \supseteq S_{-i}(\hat{I})$. Hence, the events in $S_{-i}(\mathcal{J})$ can be linearly ordered by set inclusion: write $\mathcal{J} = \{I_1, \dots, I_L\}$, with $S_{-i}(I_1) \supseteq S_{-i}(I_2) \dots \supseteq S_{-i}(I_L) \supseteq S_{-i}(J)$.

If $\mu(S_{-i}(J)|S_{-i}(I_\ell)) = 0$ for all ℓ , then $S_{-i}(J)$ is itself basic. Otherwise, there is ℓ such that $\mu(S_{-i}(J)|S_{-i}(I_\ell)) > 0$. Then $\mu(S_{-i}(I_m)|S_{-i}(I_\ell)) \geq \mu(S_{-i}(J)|S_{-i}(I_\ell)) > 0$ for all $m > \ell$, so by the chain rule $\mu(S_{-i}(J)|S_{-i}(I_m)) = \mu(S_{-i}(J)|S_{-i}(I_\ell))/\mu(S_{-i}(I_m)|S_{-i}(I_\ell)) > 0$ as well. Thus, let $\bar{\ell} = \min\{\ell : \mu(S_{-i}(J)|S_{-i}(I_\ell)) > 0\}$, and let $I \equiv I_{\bar{\ell}}$. By construction, $S_{-i}(I) \supseteq S_{-i}(J)$ and $\mu(S_{-i}(J)|S_{-i}(I)) > 0$; it remains to be shown that $S_{-i}(I)$ is basic for μ . To see this, consider $K \in \mathcal{S}_i$ such that $S_{-i}(K) \supset S_{-i}(I)$. Then also $S_{-i}(K) \supset S_{-i}(J)$, so $K = I_m$ for some $m \in \{1, \dots, L\}$. Furthermore, by the way the information sets I_1, \dots, I_L are numbered, $m < \bar{\ell}$; by the definition of $\bar{\ell}$, $\mu(S_{-i}(J)|S_{-i}(K)) = \mu(S_{-i}(J)|S_{-i}(I_m)) = 0$. If $\mu(S_{-i}(I)|S_{-i}(K)) > 0$, then the chain rule would imply $\mu(S_{-i}(J)|S_{-i}(K)) = \mu(S_{-i}(J)|S_{-i}(I)) \cdot \mu(S_{-i}(I)|S_{-i}(K)) > 0$, contradiction. Hence, $\mu(S_{-i}(I)|S_{-i}(K)) = 0$. Since $K \in \mathcal{S}_i$ was arbitrary, $S_{-i}(I)$ is basic.

Finally, suppose that $\tilde{I} \in \mathcal{S}_i$ is also such that $S_{-i}(\tilde{I})$ is basic, $S_{-i}(\tilde{I}) \supseteq S_{-i}(J)$, and $\mu(S_{-i}(J)|S_{-i}(\tilde{I})) > 0$, but $S_{-i}(\tilde{I}) \neq S_{-i}(I)$. Then $S_{-i}(I) \cap S_{-i}(\tilde{I}) \supseteq S_{-i}(J) \neq \emptyset$, so either $S_{-i}(I) \supset S_{-i}(\tilde{I})$ or $S_{-i}(\tilde{I}) \supset S_{-i}(I)$. If $S_{-i}(I) \supset S_{-i}(\tilde{I})$, then $\mu(S_{-i}(\tilde{I})|S_{-i}(I)) \geq \mu(S_{-i}(J)|S_{-i}(I)) > 0$, so $S_{-i}(\tilde{I})$ is not basic for μ . Similarly, if $S_{-i}(\tilde{I}) \supset S_{-i}(I)$, then $\mu(S_{-i}(I)|S_{-i}(\tilde{I})) \geq \mu(S_{-i}(J)|S_{-i}(\tilde{I})) > 0$, and $S_{-i}(I)$ is not basic for μ .

For the second claim, by Lemma 1 there is a μ -sequence $F_1, \dots, F_M \in S_{-i}(\mathcal{S}_i)$ such that $F_1 = F_M = S_{-i}(J)$ and $G \stackrel{\mu}{=} S_{-i}(J)$ iff $G = F_m$ for some $m \in \{1, \dots, L\}$, so that $[S_{-i}(J)]_\mu = \cup_m F_m$. Since $S_{-i}(I) \supseteq S_{-i}(J)$, $S_{-i}(I) \stackrel{\mu}{\geq} S_{-i}(J)$; and since $\mu(S_{-i}(J)|S_{-i}(I)) > 0$, $S_{-i}(J) \stackrel{\mu}{\geq} S_{-i}(I)$. Thus, $S_{-i}(I) = F_{\bar{m}}$ for some $\bar{m} \in \{1, \dots, M\}$.

Let $F = \cup_{m=1}^M F_m$. Since $\mu(F_{m+1}|F_m) > 0$ for all $m = 1, \dots, M-1$, $F_m \cap F_{m+1} \neq \emptyset$, so either $F_m \supset F_{m+1}$ or $F_m \subset F_{m+1}$. Therefore, $F = F_m$ for some $m \in \{1, \dots, M\}$. By construction $F_m = F \supseteq F_{\bar{m}} = S_{-i}(I)$. Suppose that $F_m \supsetneq S_{-i}(I)$. If $\mu(S_{-i}(J)|F_m) > 0$, then also $\mu(S_{-i}(I)|F_m) > 0$, which contradicts the fact that $S_{-i}(I)$ is basic for μ ; thus, $\mu(S_{-i}(J)|F_m) = 0$. If $m = M$, this contradicts the fact that $F_M = S_{-i}(J)$; thus, assume that $m < M$. I claim that $\mu(F_\ell|F_m) > 0$ for all $\ell = m+1, \dots, M$; for $\ell = M$, this again yields a contradiction, because $F_M = S_{-i}(J)$ by construction. The claim is proved by induction. Since F_1, \dots, F_M is a μ -sequence, $\mu(F_{m+1}|F_m) > 0$, so the claim holds for $\ell = m+1$. Inductively, assume $\mu(F_\ell|F_m) > 0$ for some $\ell \in \{m+1, \dots, M-1\}$. Since

F_1, \dots, F_M is a μ -sequence, $\mu(F_{\ell+1} \cap F_\ell | F_\ell) > 0$; by the choice of m , $F_m \supseteq F_\ell$, so the chain rule yields $\mu(F_{\ell+1} | F_m) \geq \mu(F_{\ell+1} \cap F_\ell | F_m) = \mu(F_{\ell+1} \cap F_\ell | F_\ell) \mu(F_\ell | F_m) > 0$. This proves the claim.

Thus, $F_m = S_{-i}(I)$. By construction $F_m = \cup_\ell F_\ell = [S_{-i}(J)]_\mu$. ■

Corollary 6 *Assume that the dynamic game has nested strategic information. Then every $i \in N$ and $\mu \in \Delta(S_i, S_{-i}(\mathcal{I}_i))$ admits an extension.*

Proof: By Lemma 3, it is enough to show that there exists $\mu^+ \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i) \cup [S_{-i}(\mathcal{I}_i)]_\mu)$ with $\mu^+(\cdot | F) = \mu(\cdot | F)$ for all $F \in S_{-i}(\mathcal{I}_i)$. But this follows immediately the fact that, by Lemma 5, $[S_{-i}(I)]_\mu \in S_{-i}(\mathcal{I}_i)$ for every $I \in \mathcal{I}_i$, so that $[S_{-i}(\mathcal{I}_i)]_\mu \subseteq S_{-i}(\mathcal{I}_i)$, and one can take $\mu^+ = \mu$. ■

The equivalence of the two definitions for games with nested strategic information can now be established. By Corollary 6, μ admits an extension ν . By Lemma 4 part 2, every basic event for ν is of the form $[S_{-i}(I)]_\mu$, for $I \in \mathcal{I}_i$; by Lemma 5, each such event belongs to $S_{-i}(\mathcal{I}_i)$ and is basic for μ . Conversely, if $S_{-i}(I)$ is basic for μ , then $J \equiv I$ trivially satisfies $S_{-i}(J) \supseteq S_{-i}(I)$, $S_{-i}(J)$ is basic, and $\mu(S_{-i}(I) | S_{-i}(J)) > 0$; by Lemma 5, $S_{-i}(J) = S_{-i}(I)$ is thus the unique such conditioning event satisfying these properties, and furthermore $S_{-i}(I) = [S_{-i}(I)]_\mu$, which by Lemma 4 part 2 is basic for ν . Therefore, μ and ν have the same basic events; since ν extends μ , the corresponding conditional probabilities coincide. Finally, by Lemma 4 part 3, if $S_{-i}(I), S_{-i}(J) \in S_{-i}(\mathcal{I}_i)$ are basic events for μ , hence for ν , then $S_{-i}(I) = [S_{-i}(I)]_\mu \geq^\nu [S_{-i}(J)]_\mu = S_{-i}(J)$ if and only if $S_{-i}(I) \geq^\mu S_{-i}(J)$. This implies that Definition 7 reduces to Definition 6 for games with nested strategic information.

D Appendix: Proofs of the main results

D.1 Theorem 1 (structural and sequential rationality)

Suppose that $s_i \in S_i$ is maximal for \succsim^μ , but not sequentially rational for μ . Then there is $I \in \mathcal{I}(s_i)$ and $t_i \in S_i(I)$ such that $E_{\mu(\cdot|S_{-i}(I))} U_i(s_i, \cdot) < E_{\mu(\cdot|S_{-i}(I))} U_i(t_i, \cdot)$.

By the strategic independence property, there is $r_i \in S_i(I)$ such that $U_i(r_i, s_{-i}) = U_i(t_i, s_{-i})$ for all $s_{-i} \in S_{-i}(I)$, and $U_i(r_i, s_{-i}) = U_i(s_i, s_{-i})$ for all $s_{-i} \in S_{-i} \setminus S_{-i}(I)$.

Let $\nu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i; \mu))$ be an extension of μ . By part 2 of Lemma 4, $[S_{-i}(I)]_\mu$ is basic for ν . By part 1 of Lemma 4, $\nu(S_i(I)|[S_{-i}(I)]_\mu) > 0$; by the chain rule and the fact that ν extends μ , $\nu(E|[S_{-i}(I)]_\mu) = \nu(S_{-i}(I)|[S_{-i}(I)]_\mu) \nu(E|S_{-i}(I)) = \nu(S_{-i}(I)|[S_{-i}(I)]_\mu) \mu(E|S_{-i}(I))$ for all $E \subseteq S_{-i}(I)$.

Therefore, $E_{\mu(\cdot|S_{-i}(I))} U_i(s_i, \cdot) < E_{\mu(\cdot|S_{-i}(I))} U_i(t_i, \cdot)$ implies

$$\begin{aligned} & \int_{S_{-i}(I)} U_i(s_i, s_{-i}) d\nu(\cdot|[S_{-i}(I)]_\mu) = \nu(S_{-i}(I)|[S_{-i}(I)]_\mu) \cdot E_{\mu(\cdot|S_{-i}(I))} U_i(s_i, \cdot) < \\ & < \nu(S_{-i}(I)|[S_{-i}(I)]_\mu) \cdot E_{\mu(\cdot|S_{-i}(I))} U_i(t_i, \cdot) = \int_{S_{-i}(I)} U_i(t_i, s_{-i}) d\nu(\cdot|[S_{-i}(I)]_\mu). \end{aligned}$$

Therefore,

$$\begin{aligned} E_{\nu(\cdot|[S_{-i}(I)]_\mu)} U_i(s_i, \cdot) &= \int_{S_{-i}} U_i(s_i, s_{-i}) d\nu(\cdot|[S_{-i}(I)]_\mu) = \\ &= \int_{S_{-i}(I)} U_i(s_i, s_{-i}) d\nu(\cdot|[S_{-i}(I)]_\mu) + \int_{(S_{-i}) \setminus S_{-i}(I)} U_i(s_i, s_{-i}) d\nu(\cdot|[S_{-i}(I)]_\mu) < \\ &< \int_{S_{-i}(I)} U_i(t_i, s_{-i}) d\nu(\cdot|[S_{-i}(I)]_\mu) + \int_{(S_{-i}) \setminus S_{-i}(I)} U_i(s_i, s_{-i}) d\nu(\cdot|[S_{-i}(I)]_\mu) = \\ &= \int_{S_{-i}(I)} U_i(r_i, s_{-i}) d\nu(\cdot|[S_{-i}(I)]_\mu) + \int_{(S_{-i}) \setminus S_{-i}(I)} U_i(r_i, s_{-i}) d\nu(\cdot|[S_{-i}(I)]_\mu) = E_{\nu(\cdot|[S_{-i}(I)]_\mu)} U_i(r_i, \cdot). \end{aligned}$$

Now fix $J \in \mathcal{I}_i$; two cases must be considered.

Case 1: $\nu(S_{-i}(I)|[S_{-i}(J)]_\mu) = 0$. For such J , trivially

$$\int_{S_{-i}(I)} U_i(s_i, s_{-i}) d\nu(\cdot|[S_{-i}(J)]_\mu) = 0 = \int_{S_{-i}(I)} U_i(r_i, s_{-i}) d\nu(\cdot|[S_{-i}(J)]_\mu)$$

and so

$$\begin{aligned}
\mathbb{E}_{\nu(\cdot|[\mathcal{S}_{-i}(J)]_\mu)} U_i(s_i, \cdot) &= \int_{\mathcal{S}_{-i}} U_i(s_i, s_{-i}) d\nu(\cdot|[\mathcal{S}_{-i}(J)]_\mu) = \\
&= \int_{\mathcal{S}_{-i}(I)} U_i(s_i, s_{-i}) d\nu(\cdot|[\mathcal{S}_{-i}(J)]_\mu) + \int_{(\mathcal{S}_{-i}) \setminus \mathcal{S}_{-i}(I)} U_i(s_i, s_{-i}) d\nu(\cdot|[\mathcal{S}_{-i}(J)]_\mu) = \\
&= \int_{\mathcal{S}_{-i}(I)} U_i(r_i, s_{-i}) d\nu(\cdot|[\mathcal{S}_{-i}(J)]_\mu) + \int_{(\mathcal{S}_{-i}) \setminus \mathcal{S}_{-i}(I)} U_i(r_i, s_{-i}) d\nu(\cdot|[\mathcal{S}_{-i}(J)]_\mu) = \mathbb{E}_{\nu(\cdot|[\mathcal{S}_{-i}(J)]_\mu)} U_i(r_i, \cdot).
\end{aligned}$$

Case 2: $\nu(\mathcal{S}_{-i}(I)|[\mathcal{S}_{-i}(J)]_\mu) > 0$. In this case, $\nu([\mathcal{S}_{-i}(I)]_\mu|[\mathcal{S}_{-i}(J)]_\mu) \geq \nu(\mathcal{S}_{-i}(I)|[\mathcal{S}_{-i}(J)]_\mu) > 0$, so $\mathcal{S}_{-i}(I) \geq^\nu [\mathcal{S}_{-i}(J)]_\mu$; since trivially $\nu([\mathcal{S}_{-i}(I)]_\mu|\mathcal{S}_{-i}(I)) > 0$, also $[\mathcal{S}_{-i}(I)]_\mu \geq^\nu \mathcal{S}_{-i}(I)$, and so by transitivity $[\mathcal{S}_{-i}(I)]_\mu \geq^\nu [\mathcal{S}_{-i}(J)]_\mu$.

To conclude the argument, consider $J \in \mathcal{J}_i$ with $[\mathcal{S}_{-i}(J)]_\mu >^\nu [\mathcal{S}_{-i}(I)]_\mu$. If $\nu(\mathcal{S}_{-i}(I)|[\mathcal{S}_{-i}(J)]_\mu) > 0$, then Case 2 holds and $[\mathcal{S}_{-i}(I)]_\mu \geq^\nu [\mathcal{S}_{-i}(J)]_\mu$, contradiction; thus, $\nu(\mathcal{S}_{-i}(I)|[\mathcal{S}_{-i}(J)]_\mu) = 0$. Hence Case 1 applies, so $\mathbb{E}_{\nu(\cdot|[\mathcal{S}_{-i}(J)]_\mu)} U_i(r_i, \cdot) = \mathbb{E}_{\nu(\cdot|[\mathcal{S}_{-i}(J)]_\mu)} U_i(s_i, \cdot)$. Thus, $\mathbb{E}_{\nu(\cdot|[\mathcal{S}_{-i}(I)]_\mu)} U_i(r_i, \cdot) > \mathbb{E}_{\nu(\cdot|[\mathcal{S}_{-i}(I)]_\mu)} U_i(s_i, \cdot)$ and, by Lemma 4 part 2, $\mathbb{E}_{\nu(\cdot|G)} U_i(r_i, \cdot) = \mathbb{E}_{\nu(\cdot|G)} U_i(s_i, \cdot)$ for all $G \in \mathcal{S}_{-i}(\mathcal{J}_i; \mu)$ that are basic for ν and satisfy $G >^\nu [\mathcal{S}_{-i}(I)]_\mu$; hence, $s_i \not\asymp^\mu r_i$.

On the other hand, consider $J \in \mathcal{J}_i$ such that $\mathbb{E}_{\nu(\cdot|[\mathcal{S}_{-i}(J)]_\mu)} U_i(r_i, \cdot) < \mathbb{E}_{\nu(\cdot|[\mathcal{S}_{-i}(J)]_\mu)} U_i(s_i, \cdot)$. Then Case 1 cannot apply, so $\nu(\mathcal{S}_{-i}(I)|[\mathcal{S}_{-i}(J)]_\mu) > 0$ and Case 2 applies instead; this implies $[\mathcal{S}_{-i}(I)]_\mu \geq^\nu [\mathcal{S}_{-i}(J)]_\mu$. If also $[\mathcal{S}_{-i}(J)]_\mu \geq^\nu [\mathcal{S}_{-i}(I)]_\mu$, then by Lemma 4 part 3 $\mathcal{S}_{-i}(I) =^\mu \mathcal{S}_{-i}(J)$, so $[\mathcal{S}_{-i}(I)]_\mu = [\mathcal{S}_{-i}(J)]_\mu$, which contradicts the fact that $\mathbb{E}_{\nu(\cdot|[\mathcal{S}_{-i}(I)]_\mu)} U_i(r_i, \cdot) > \mathbb{E}_{\nu(\cdot|[\mathcal{S}_{-i}(I)]_\mu)} U_i(s_i, \cdot)$. Hence $[\mathcal{S}_{-i}(I)]_\mu >^\nu [\mathcal{S}_{-i}(J)]_\mu$. By Lemma 4 part 2, this implies that, for every $G \in \mathcal{S}_{-i}(\mathcal{J}_i; \mu)$ that is basic for ν and such that $\mathbb{E}_{\nu(\cdot|G)} U_i(r_i, \cdot) < \mathbb{E}_{\nu(\cdot|G)} U_i(s_i, \cdot)$, the basic event $F = [\mathcal{S}_{-i}(I)]_\mu \in \mathcal{S}_{-i}(\mathcal{J}_i; \mu)$ satisfies $F >^\nu G$ and $\mathbb{E}_{\nu(\cdot|F)} U_i(r_i, \cdot) > \mathbb{E}_{\nu(\cdot|F)} U_i(s_i, \cdot)$: therefore, $r_i \not\asymp^\mu s_i$.

Thus, $r_i \succ^\mu s_i$, which contradicts the assumption that s_i was maximal for \succ^μ . ■

D.2 Elicitation

Throughout this section, fix a dynamic game $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$, a questionnaire $Q = (Q_i)_{i \in N}$, and the corresponding elicitation game $(N \cup \{c\}, (S_i^*, \mathcal{I}_i^*, U_i^*)_{i \in N \cup \{c\}}, S^*(\cdot))$, according to Definition 8. It is convenient to let $N^* = N \cup \{c\}$. Also, as in part 1 of Definition 8, for every $i \in N$, let $W_i = \{\emptyset\}$ if $Q_i = \emptyset$ and $W_i = \{b, p\}$ if $Q_i = (I, E, p)$.

D.2.1 Preliminaries

I first verify that the elicitation game satisfies two properties in Section 2. This is necessary to ensure that definitions and results on structural rationality in Section 4 apply.

It is immediate by inspecting Definition 8 that, for every $i \in N^*$ and $I^* \in \mathcal{I}_i^*$, $S^*(I) = S_i^*(I^*) \times S_{-i}^*(I^*)$. Second, fix $i \in N$ (so $i \neq c$) and $I^*, J^* \in \mathcal{I}_i^*$: it must be shown that either $S^*(I^*) \cap S^*(J^*) = \emptyset$, or $S^*(I^*)$ and $S^*(J^*)$ are nested. This is immediate if I^* or J^* equal I_i^1 . Otherwise, $I^* = (s_i, w_i, I)$ and $J^* = (s'_i, w'_i, J)$, where $s_i \in S_i(I)$ and $s'_i \in S_i(J)$; then $S_i^*(I^*) = \{(s_i, w_i)\}$ and $S_i^*(J^*) = \{(s'_i, w'_i)\}$. If either $s_i \neq s'_i$ or $w_i \neq w'_i$, then $S^*(I^*) \cap S^*(J^*) = \emptyset$. Thus, suppose $s_i = s'_i$ and $w_i = w'_i$. By part 4 of Definition 8, $S^*(I^*) = \{(s_i, w_i)\} \times S_{-i}(I) \times W_{-i} \times S_c^*$ and $S^*(J^*) = \{(s_i, w_i)\} \times S_{-i}(J) \times W_{-i} \times S_c^*$. Therefore, $S^*(I^*) \cap S^*(J^*) \neq \emptyset$ implies $S_{-i}(I) \cap S_{-i}(J) \neq \emptyset$, and so $S(I) \cap S(J) \supseteq [\{s_i\} \times S_{-i}(I)] \cap [\{s_i\} \times S_{-i}(J)] \neq \emptyset$. Therefore $S(I)$ and $S(J)$ are nested: say $S(I) \supseteq S(J)$, and so $S_{-i}(I) \supseteq S_{-i}(J)$. But then, part 4 of Definition 8 implies that $S^*(I^*)$ and $S^*(J^*)$ are nested, as required.

Next, it must be verified that the elicitation game satisfies strategic independence. Again, it is enough to focus on $i \in N$ and $I^* = (s_i, w_i, I) \in \mathcal{I}_i^*$, with $s_i \in S_i(I)$, because $S_{-i}^*(I^*) = S_{-i}^*$ for all other I^* (including for $i = c$ and $I^* = \phi$). But part 4 of Definition 8 implies that $S_i^*(I^*) = \{(s_i, w_i)\}$, a singleton set, so strategic independence holds trivially.

Finally, the elicitation game preserves nested strategic information:

Remark 3 *If $(N, (S_i, \mathcal{I}_i, U_i)_{i \in N}, S(\cdot))$ has nested strategic information, so does the associated elicitation game.*

Proof: Suppose the original game has nested strategic information, and fix a player $i \in N$. It is enough to consider information sets of the form $(s_i, w_i, I), (s'_i, w'_i, I') \in \mathcal{I}_i$. Suppose that $((s_{-i}, w_{-i}), s_c^*) \in S_{-i}^*((s_i, w_i, I)) \cap S_{-i}^*((s'_i, w'_i, I'))$; then, by Definition 8 part 4, $s_{-i} \in S_{-i}(I) \cap S_{-i}(I')$. Since the original game has nested strategic information, $S_{-i}(I)$ and $S_{-i}(I')$ are nested; assume that $S_{-i}(I) \subseteq S_{-i}(I')$. Then, Definition 8 part 4 immediately implies that $S_{-i}^*((s_i, w_i, I)) \subseteq S_{-i}^*((s'_i, w'_i, I'))$. ■

D.2.2 Proof of Theorem 2

Throughout this subsection, fix a player $i \in N$ and a CPS $\mu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i))$ with extension $\nu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i; \mu))$. I first show that every CPS for the associated elicitation game that agrees with μ preserves the likelihood ordering of conditioning events.

Lemma 6 *Let $\mu^* \in \Delta(S_{-i}^*, \mathcal{I}_i^*)$ agree with μ (Definition 9). Then:*

- (0) $I^* \in \mathcal{I}_i^*$ if and only if $S_{-i}^*(I^*) = S_{-i}(I) \times W_{-i} \times S_c^*$ for some $I \in \mathcal{I}_i$.
- (1) for any $I^*, J^* \in \mathcal{I}_i^*$, $\mu^*(S_{-i}^*(I^*) | S_{-i}^*(J^*)) = \mu(\text{proj}_{S_{-i}} S_{-i}^*(I^*) | \text{proj}_{S_{-i}} S_{-i}^*(J^*))$.
- (2) for any $I^*, J^* \in \mathcal{I}_i^*$, $S_{-i}^*(I^*) \geq^{\mu^*} S_{-i}^*(J^*)$ if and only if $\text{proj}_{S_{-i}} S_{-i}^*(I^*) \geq^{\mu} \text{proj}_{S_{-i}} S_{-i}^*(J^*)$.
- (3) $S_{-i}^*(\mathcal{I}_i^*; \mu^*) = \{F \times W_{-i} \times S_c^* : F \in S_{-i}(\mathcal{I}_i; \mu)\}$

Proof: (0): Fix $I^* \in \mathcal{I}_i^*$. If $I^* = \phi^*$ or $I^* = I_i^1$, then $S_{-i}^*(I^*) = S_{-i}^* = S_{-i} \times W_{-i} \times S_c^* = S_{-i}(\phi) \times W_{-i} \times S_c^*$. If instead $I^* = (s_i, w_i, I)$ for some $s_i \in S_i$, $w_i \in W_i$ and $I \in \mathcal{I}_i$, then $S_{-i}(I^*) = S_{-i}(I) \times W_{-i} \times S_c^*$. Conversely, for every $I \in \mathcal{I}_i$, $s_i \in S_i(I)$ and $w_i \in W_i$, $I^* = (s_i, w_i, I) \in \mathcal{I}_i^*$ satisfies $S_{-i}^*(I^*) = S_{-i}(I) \times W_{-i} \times S_c^*$.

(1): if $I^* = \phi^*$ or $I^* = I_i^1$, then $S_{-i}^*(I^*) = S_{-i}^*$, so both conditional probabilities equal 1. Otherwise, by part 4 of Definition 8, $S_{-i}^*(I^*) = S_{-i}(I) \times W_{-i} \times S_c^*$ for some $I \in \mathcal{I}_i$, so

$$\begin{aligned} \mu^*(S_{-i}^*(I^*) | S_{-i}^*(J^*)) &= \mu^*(S_{-i}(I) \times W_{-i} \times S_c^* | S_{-i}^*(J^*)) = [\text{marg}_{S_{-i} \times S_c^*} \mu^*(\cdot | S_{-i}^*(J^*))] (S_{-i}(I) \times S_c^*) = \\ &= \mu(S_{-i}(I) | \text{proj}_{S_{-i}} S_{-i}^*(J^*)) = \mu(\text{proj}_{S_{-i}} S_{-i}^*(I^*) | \text{proj}_{S_{-i}} S_{-i}^*(J^*)), \end{aligned}$$

where the second equality follows from the definition of marginalization and the third from Definition 9.

(2): suppose that $S_{-i}^*(I^*) \geq^{\mu^*} S_{-i}^*(J^*)$. Then there are $I_1^*, \dots, I_L^* \in \mathcal{I}_i^*$ such that $I_1^* = J^*$, $I_L^* = I^*$, and $\mu^*(S_{-i}^*(I_{\ell+1}^*) | S_{-i}^*(I_\ell^*)) > 0$ for $\ell = 1, \dots, L-1$. Hence (1) implies that $\mu(\text{proj}_{S_{-i}} S_{-i}^*(I_{\ell+1}^*) | \text{proj}_{S_{-i}} S_{-i}^*(I_\ell^*)) > 0$ for $\ell = 1, \dots, L-1$, which implies that $\text{proj}_{S_{-i}} S_{-i}^*(I^*) = \text{proj}_{S_{-i}} S_{-i}^*(I_L^*) \geq^\mu \text{proj}_{S_{-i}} S_{-i}^*(I_1^*) = \text{proj}_{S_{-i}} S_{-i}^*(J^*)$.

Conversely, suppose that $\text{proj}_{S_{-i}} S_{-i}^*(I^*) \geq^\mu \text{proj}_{S_{-i}} S_{-i}^*(J^*)$, so there are $I_1, \dots, I_L \in \mathcal{I}_i$ such that $S_{-i}(I_1) = \text{proj}_{S_{-i}} S_{-i}^*(J^*)$, $S_{-i}(I_L) = \text{proj}_{S_{-i}} S_{-i}^*(I^*)$, and $\mu(S_{-i}(I_{\ell+1}) | S_{-i}(I_\ell)) > 0$ for all $\ell = 1, \dots, L-1$. Let $I_1^* = J^*$, $I_L^* = I^*$, and $I_\ell^* = (s_i, w_i, I_\ell)$, with $s_i \in S_{-i}(I_\ell)$ and $w_i \in W_i$, for all $\ell = 2, \dots, L-1$. Then, for all $\ell = 1, \dots, L$, $\text{proj}_{S_{-i}} S_{-i}^*(I_\ell^*) = S_{-i}(I_\ell)$, so part (1) implies that $\mu^*(S_{-i}(I_{\ell+1}^*) | S_{-i}(I_\ell^*)) > 0$ for all $\ell = 1, L-1$. Therefore $S_{-i}^*(I^*) \geq^{\mu^*} S_{-i}^*(J^*)$.

(3) Fix $I_1^*, \dots, I_L^* \in \mathcal{I}_i^*$; by part (0), for every $\ell = 1, \dots, L$, there is $I_\ell \in \mathcal{I}_i$ such that $S_{-i}^*(I_\ell^*) = S_{-i}(I_\ell) \times W_{-i} \times S_c^*$, so $\text{proj}_{S_{-i}} S_{-i}^*(I_\ell^*) = S_{-i}(I_\ell)$. Then, by part (2), $S_{-i}^*(I_\ell^*) =^{\mu^\circ} S_{-i}^*(I_m^*)$ for all ℓ, m if and only if $S_{-i}(I_\ell) =^\mu S_{-i}(I_m)$ for all ℓ, m . Furthermore, $\cup_\ell S_{-i}^*(I_\ell^*) = [\cup_\ell S_{-i}(I_\ell)] \times W_{-i} \times S_c^*$. Therefore, $S_{-i}^*(\mathcal{I}_i^*; \mu^*) = \{F \times W_{-i} \times S_c^* : F \in S_{-i}(\mathcal{I}_i; \mu)\}$. ■

Lemma 7 *If μ is extensible, then there is an extensible CPS $\mu^* \in \Delta(S_{-i}^*, \mathcal{I}_i^*)$ that agrees with μ .*

Proof: By Lemma 6 part (3), $S_{-i}^*(\mathcal{I}_i^*; \mu^\circ)$ is the same for all CPSs μ° that agree with μ ; to emphasize this, in this proof only, denote this set by $S_{-i}^*(\mathcal{I}_i^*; \mu)$.

Now recall that $\nu \in \Delta(S_{-i}, S_{-i}(\mathcal{I}_i; \mu))$ is an extension of μ . Define an array $(\nu^*(\cdot | F^*))_{F^* \in S_{-i}^*(\mathcal{I}_i^*; \mu)} \in \Delta(S_{-i}^*)^{S_{-i}^*(\mathcal{I}_i^*; \mu)}$ by fixing an arbitrary element $w_{-i} \in W_{-i}$ (cf. part 1 of Definition 8) and letting, for every $s_{-i} \in S_{-i}$, $s_c^* \in S_c^*$, and $F^* \in S_{-i}^*(\mathcal{I}_i^*; \mu^*)$,

$$\nu^*({(s_{-i}, w_{-i}, s_c^*) | F^*}) = \frac{1}{|S_c^*|} \nu(\{s_{-i}\} | \text{proj}_{S_{-i}} F^*).$$

That is, ν^* assigns probability 1 to every player $j \in N \setminus \{i\}$ choosing the pre-specified $w_j \in W_j$.

I claim that $\nu^* \in \Delta(S_{-i}^*; S_{-i}^*(\mathcal{I}_i^*; \mu))$. Fix $F^* \in S_{-i}^*(\mathcal{I}_i^*; \mu)$, so $F^* = F \times W_{-i} \times S_c^*$ for some $F \in S_{-i}(\mathcal{I}_i; \mu)$: then

$$\nu^*(F^*|F^*) = \sum_{s_c^* \in S_c^*} \nu^*(F \times \{(w_{-i}, s_c^*)|F^*\}) = \sum_{s_c^* \in S_c^*} \frac{1}{|S_c^*|} \nu(F|F) = 1.$$

Furthermore, fix $F^*, G^* \in S_{-i}^*(\mathcal{I}_i^*; \mu)$ with $F^* \subseteq G^*$, and $E^* \subset F^*$; write $F^* = F \times W_{-i} \times S_c^*$ and $G^* = G \times W_{-i} \times S_c^*$, with $F, G \in S_{-i}(\mathcal{I}_i; \mu)$ and $\text{proj}_{S_{-i}} E^* \subseteq F \subseteq G$. Then

$$\begin{aligned} \nu^*(E^*|G^*) &= \sum_{(s_{-i}, s_c^*) \in S_{-i} \times S_c^*: (s_{-i}, w_{-i}, s_c^*) \in E^*} \nu^*({s_{-i}, w_{-i}, s_c^*}|G^*) = \sum_{(s_{-i}, s_c^*) \in S_{-i} \times S_c^*: (s_{-i}, w_{-i}, s_c^*) \in E^*} \frac{1}{|S_c^*|} \nu({s_{-i}}|G) = \\ &= \sum_{(s_{-i}, s_c^*) \in S_{-i} \times S_c^*: (s_{-i}, w_{-i}, s_c^*) \in E^*} \frac{1}{|S_c^*|} \nu({s_{-i}}|F) \nu(F|G) = \nu(F|G) \nu^*(E^*|F^*). \end{aligned}$$

In the particular case $E^* = F^*$, one gets $\nu^*(F^*|G^*) = \nu(F|G) \nu^*(F^*|F^*) = \nu(F|G)$: thus, for general E^* , $\nu^*(E^*|G^*) = \nu^*(E^*|F^*) \nu^*(F^*|G^*)$. This proves that ν^* is indeed a CPS.

Let μ^* be the restriction of ν^* to the conditioning events $S_{-i}^*(\mathcal{I}_i^*) \subseteq S_{-i}^*(\mathcal{I}_i^*; \mu)$; note that $S_{-i}^*(\mathcal{I}_i^*) = \{S_{-i}(I) \times W_{-i} \times S_c^* : I \in \mathcal{I}_i\} = \{F \times W_{-i} \times S_c^* : F \in S_{-i}(\mathcal{I}_i)\}$. Then $\mu^* \in \Delta(S_{-i}^*, S_{-i}^*(\mathcal{I}_i^*))$, so ν^* is an extension of μ^* , and in particular $S_{-i}^*(\mathcal{I}_i^*; \mu) = S_{-i}^*(\mathcal{I}_i^*; \mu^*)$.

Finally, μ^* agrees with μ : for every $(s_{-i}, s_c^*) \in S_{-i} \times S_c^*$ and $F^* \in S_{-i}^*(\mathcal{I}_i)$,

$$\begin{aligned} [\text{marg}_{S_{-i} \times S_c^*} \mu^*(\cdot|F^*)]({s_{-i}, s_c^*}) &= \mu^*({s_{-i}} \times W_{-i} \times {s_c^*}|F^*) = \\ &= \nu^*({s_{-i}} \times W_{-i} \times {s_c^*}|F^*) = \\ &= \nu^*({s_{-i}, w_{-i}, s_c^*}|F^*) = \\ &= \frac{1}{|S_c^*|} \nu({s_{-i}}|\text{proj}_{S_{-i}} F^*) = \frac{1}{|S_c^*|} \mu({s_{-i}}|\text{proj}_{S_{-i}} F^*) \end{aligned}$$

The first equality is the definition of marginalization; the second follows from the fact that ν^* extends μ^* ; the third and fourth from the definition of ν^* ; and the last from the fact that ν extends μ . This completes the proof. ■

The following Lemma identifies a decomposition property that every extension of a CPS that agrees with μ must satisfy. (This property holds by construction for the CPS ν^* defined in the proof of Lemma 7.) It provides the key step in the proof of Theorem 2.

Lemma 8 Let $\mu^* \in \Delta(S_{-i}^*, \mathcal{I}_i^*)$ a CPS that agrees with μ and admits an extension $\nu^* \in \Delta(S_{-i}^*, S_{-i}^*(\mathcal{I}_i^*; \mu^*))$.

Fix $I^* \in \mathcal{I}_i^*$ and let $I \in \mathcal{I}_i$ be such that $S_{-i}^*(I^*) = S_{-i}(I) \times W_{-i} \times S_c^*$ (cf. Lemma 6 part 0). Then

$$[S_{-i}^*(I^*)]_{\mu^*} = [S_{-i}(I)]_{\mu} \times W_{-i} \times S_c^* \quad (10)$$

and, for all $s_{-i} \in S_{-i}$ and $s_c^* \in S_c^*$,

$$\nu^*({s_{-i}} \times W_{-i} \times {s_c^*} | [S_{-i}^*(I^*)]_{\mu^*}) = \frac{1}{|S_c^*|} \nu({s_{-i}} | [S_{-i}(I)]_{\mu}). \quad (11)$$

Furthermore, for every $s_i \in S_i$, if $Q_i = (\hat{I}, E, p)$ then

$$\mathbb{E}_{\nu^*({s_{-i}} | [S_{-i}^*(I^*)]_{\mu^*})} U_i^*((s_i, b), \cdot) = \frac{|S_c^*| - 1}{|S_c^*|} \mathbb{E}_{\nu({s_{-i}} | [S_{-i}(I)]_{\mu})} U_i(s_i, \cdot) + \frac{1}{|S_c^*|} \nu(S_{-i}(\hat{I}) | [S_{-i}(I)]_{\mu}) \mu(E | S_{-i}(\hat{I})) \quad (12)$$

$$\mathbb{E}_{\nu^*({s_{-i}} | [S_{-i}^*(I^*)]_{\mu^*})} U_i^*((s_i, p), \cdot) = \frac{|S_c^*| - 1}{|S_c^*|} \mathbb{E}_{\nu({s_{-i}} | [S_{-i}(I)]_{\mu})} U_i(s_i, \cdot) + \frac{1}{|S_c^*|} \nu(S_{-i}(\hat{I}) | [S_{-i}(I)]_{\mu}) p, \quad (13)$$

whereas, if $Q_i = \emptyset$, then

$$\mathbb{E}_{\nu^*({s_{-i}} | [S_{-i}^*(I^*)]_{\mu^*})} U_i^*((s_i, \emptyset), \cdot) = \mathbb{E}_{\nu({s_{-i}} | [S_{-i}(I)]_{\mu})} U_i(s_i, \cdot). \quad (14)$$

Proof: Fix $I^* \in \mathcal{I}_i^*$ and let $I \in \mathcal{I}_i$ be such that $S_{-i}^*(I^*) = S_{-i}(I) \times W_{-i} \times S_c^*$. By parts (0) and (2) of Lemma 6,

$$\begin{aligned} [S_{-i}^*(I^*)]_{\mu^*} &= \cup \{S_{-i}^*(J^*) : J^* \in \mathcal{I}_i^*, S_{-i}^*(J^*) =^{\mu^*} S_{-i}^*(I^*)\} = \\ &= \cup \{S_{-i}(J) \times W_{-i} \times S_c^* : J \in \mathcal{I}_i, S_{-i}(J) =^{\mu} S_{-i}(I)\} = \\ &= \cup \{S_{-i}(J) : J \in \mathcal{I}_i, S_{-i}(J) =^{\mu} S_{-i}(I)\} \times W_{-i} \times S_c^* = [S_{-i}(I)]_{\mu} \times W_{-i} \times S_c^*. \end{aligned}$$

Consider $s_{-i} \in S_{-i}$, and $s_c^* \in S_c^*$. If $s_{-i} \notin [S_{-i}(I)]_{\mu}$, then $\{{s_{-i}} \times W_{-i} \times {s_c^*}\} \cap [S_{-i}^*(I^*)]_{\mu^*} = \emptyset$, so $\nu^*({s_{-i}} \times W_{-i} \times {s_c^*} | [S_{-i}^*(I^*)]_{\mu^*}) = 0 = \nu({s_{-i}} | [S_{-i}(I)]_{\mu})$. Thus, assume $s_{-i} \in [S_{-i}(I)]_{\mu}$, so there is $J \in \mathcal{I}_i$ such that $s_{-i} \in S_{-i}(J)$ and $S_{-i}(J) =^{\mu} S_{-i}(I)$, so $S_{-i}(J) \subseteq [S_{-i}(I)]_{\mu}$. In addition, by Lemma 4 part 1, $\nu(S_{-i}(J) | [S_{-i}(I)]_{\mu}) > 0$. Finally, $S_{-i}(J) \times W_{-i} \times S_c^* \in S_{-i}^*(\mathcal{I}_i^*) \in S_{-i}(\mathcal{I}_i^*; \mu^*)$ and $S_{-i}(J) \times W_{-i} \times S_c^* \subseteq [S_{-i}^*(I^*)]_{\mu^*}$. Then, by the chain rule, the assumptions that ν^* extends μ^* and

μ^* agrees with μ , and the fact that $\nu(S_{-i}(J)|[S_{-i}(I)]_\mu) > 0$,

$$\begin{aligned}
\nu^*({s_{-i}} \times W_{-i} \times {s_c^*} | [S_{-i}^*(I^*)]_{\mu^*}) &= \nu^*({s_{-i}} \times W_{-i} \times {s_c^*} | S_{-i}(J) \times W_{-i} \times S_c^*) \cdot \nu^*(S_{-i}(J) \times W_{-i} \times S_c^* | [S_{-i}^*(I^*)]_{\mu^*}) = \\
&= \mu^*({s_{-i}} \times W_{-i} \times {s_c^*} | S_{-i}(J) \times W_{-i} \times S_c^*) \cdot \nu^*(S_{-i}(J) \times W_{-i} \times S_c^* | [S_{-i}^*(I^*)]_{\mu^*}) = \\
&= \frac{1}{|S_c^*|} \mu({s_{-i}} | S_{-i}(J)) \cdot \nu^*(S_{-i}(J) \times W_{-i} \times S_c^* | [S_{-i}^*(I^*)]_{\mu^*}) = \\
&= \frac{1}{|S_c^*|} \nu({s_{-i}} | [S_{-i}(I)]_\mu) \cdot \frac{\nu^*(S_{-i}(J) \times W_{-i} \times S_c^* | [S_{-i}^*(I^*)]_{\mu^*})}{\nu(S_{-i}(J) | [S_{-i}^*(I)]_\mu)} \equiv \\
&\equiv \frac{1}{|S_c^*|} \nu({s_{-i}} | [S_{-i}(I)]_\mu) \cdot \kappa_J.
\end{aligned}$$

It must thus be shown that $\kappa_J = 1$ for all $J \in I_i$ such that $S_{-i}(I) =^\mu S_{-i}(J)$.

To do so, let $\{I_1, \dots, I_L\}$ be such that, for any $J \in \mathcal{I}_i$, $S_{-i}(I) =^\mu S_{-i}(J)$ if and only if $J = I_\ell$ for some $\ell = 1, \dots, L$. [Thus $\{S_{-i}(I_1), \dots, S_{-i}(I_L)\}$ is the \geq^μ -equivalence class containing $S_{-i}(I)$.] By Lemma 1, there is a μ -sequence F_1, \dots, F_M such that $\{F_1, \dots, F_M\} = \{S_{-i}(I_1), \dots, S_{-i}(I_L)\}$. Note that, for every $m = 1, \dots, M$, there is $I_{(m)} \in \{I_1, \dots, I_L\}$ such that $F_m = S_{-i}(I_{(m)})$.²³ For every $m = 1, \dots, M-1$, $\mu(F_{m+1}|F_m) > 0$, so there is $s_{-i}^m \in F_m \cap F_{m+1}$. Then also $s_{-i} \in [S_{-i}(I)]_\mu$ and $\{s_{-i}\} \times W_{-i} \times S_c^* \subseteq [S_{-i}^*(I^*)]_{\mu^*}$, so (adding over all $s_c^* \in S_c^*$)

$$\nu(\{s_{-i}^m\} | [S_{-i}(I)]_\mu) \cdot \kappa_{I_{(m)}} = \nu^*({s_{-i}} \times W_{-i} \times S_c^* | [S_{-i}^*(I^*)]_{\mu^*}) = \nu(\{s_{-i}^m\} | [S_{-i}(I)]_\mu) \cdot \kappa_{I_{(m+1)}},$$

which implies that $\kappa_m = \kappa_{m+1}$. Therefore, there is $\kappa \in \mathbb{R}$ such that $\kappa_J = \kappa$ for all $J \in \mathcal{I}_i$ with $S_{-i}(I) =^\mu S_{-i}(J)$. But

$$1 = \sum_{s_{-i} | [S_{-i}(I)]_\mu, s_c^* \in S_c^*} \nu^*({s_{-i}} \times W_{-i} \times {s_c^*} | [S_{-i}^*(I^*)]_{\mu^*}) = \sum_{s_{-i} | [S_{-i}(I)]_\mu, s_c^* \in S_c^*} \frac{1}{|S_c^*|} \nu(\{s_{-i}\} | [S_{-i}(I)]_\mu) \cdot \kappa = \kappa,$$

which completes the proof of Eq. (11).

²³Recall that, to obtain the μ -sequence F_1, \dots, F_M , it may be necessary to rearrange and/or repeat the sets $S_{-i}(I_1), \dots, S_{-i}(I_L)$; hence the need for a separate indexing of the information sets I_1, \dots, I_L .

Finally, fix $s_i \in S_i^*$. If $Q_i = (\hat{I}, E, p)$, then:

$$\begin{aligned}
& E_{\nu^*(\cdot | [S_{-i}^*(I^*)]_{\mu^*})} U_i^*((s_i, b), \cdot) = \\
&= \sum_{s_{-i} \in S_{-i}} \sum_{w_{-i} \in W_{-i}} \sum_{s_c^* \in S_c^*} \nu^* (\{(s_{-i}, w_{-i}, s_c^*)\} | [S_{-i}^*(I^*)]_{\mu^*}) U_i^*((s_i, b), (s_{-i}, w_{-i}, s_c^*)) = \\
&= \sum_{s_{-i} \in S_{-i}} \sum_{w_{-i} \in W_{-i}} \sum_{s_c^* \in S_c^* \setminus \{i\}} \nu^* (\{(s_{-i}, w_{-i}, s_c^*)\} | [S_{-i}^*(I^*)]_{\mu^*}) U_i(s_i, s_{-i}) + \\
&\quad + \sum_{s_{-i} \in S_{-i}} \sum_{w_{-i} \in W_{-i}} \nu^* (\{(s_{-i}, w_{-i}, i)\} | [S_{-i}^*(I^*)]_{\mu^*}) 1_E(s_{-i}) = \\
&= \sum_{s_{-i} \in S_{-i}} \sum_{s_c^* \in S_c^* \setminus \{i\}} \nu^* (\{(s_{-i}\} \times W_{-i} \times \{s_c^*\} | [S_{-i}^*(I^*)]_{\mu^*}) U_i(s_i, s_{-i}) + \\
&\quad + \sum_{s_{-i} \in S_{-i}} \nu^* (\{(s_{-i}\} \times W_{-i} \times \{i\} | [S_{-i}^*(I^*)]_{\mu^*}) 1_E(s_{-i}) = \\
&= \sum_{s_{-i} \in S_{-i}} \sum_{s_c^* \in S_c^* \setminus \{i\}} \frac{1}{|S_c^*|} \nu(\{(s_{-i}\} | [S_{-i}(I)]_{\mu}) U_i(s_i, s_{-i}) + \\
&\quad + \sum_{s_{-i} \in S_{-i}} \frac{1}{|S_c^*|} \nu(\{(s_{-i}\} | [S_{-i}(I)]_{\mu}) 1_E(s_{-i}) = \\
&= \frac{|S_c^*| - 1}{|S_c^*|} E_{\nu(\cdot | [S_{-i}]_{\mu})} U_i(s_i, \cdot) + \frac{1}{|S_c^*|} \nu(E | [S_{-i}(I)]_{\mu}) = \\
&= \frac{|S_c^*| - 1}{|S_c^*|} E_{\nu(\cdot | [S_{-i}]_{\mu})} U_i(s_i, \cdot) + \frac{1}{|S_c^*|} \nu(S_{-i}(\hat{I}) | [S_{-i}(I)]_{\mu}) \mu(E | S_{-i}(\hat{I})),
\end{aligned}$$

i.e., Eq. (12) holds. The other equations are proved similarly. ■

The proof of Theorem 2 can now be completed. Lemma 7 shows that there exists an extensible CPS $\mu^* \in \Delta(S_{-i}^*, S_{-i}^*(\mathcal{I}_i^*))$ that agrees with μ ; call ν^* its extension. By Lemma 4 part 2, $F^* \in S_{-i}^*(\mathcal{I}_i^*, \mu^*)$ is basic for ν^* if and only if $F^* = [S_{-i}^*(I^*)]_{\mu^*}$ for some $I^* \in \mathcal{I}_i^*$; by part 3 of that Lemma, $[S_{-i}^*(I^*)]_{\mu^*} \geq^{\nu^*} [S_{-i}^*(I^*)]_{\mu^*}$ if and only if $S_{-i}^*(I^*) \geq^{\mu^*} S_{-i}^*(J^*)$; by parts (0) and (2) of Lemma 6, this holds if and only if $S_{-i}(I) \geq^{\mu} S_{-i}(J)$, where $I, J \in \mathcal{I}_i$ are such that $S_{-i}^*(I^*) = S_{-i}(I) \times W_{-i} \times S_c^*$ and $S_{-i}^*(J^*) = S_{-i}(J) \times W_{-i} \times S_c^*$; and by applying part 3 of Lemma 4 to ν rather than ν^* , this holds if and only if $[S_{-i}(I)]_{\mu} \geq^{\nu} [S_{-i}(J)]_{\mu}$. Then part (1) of Theorem 2 follows from the definition of structural preferences (Definition 7) and Eqs. (12)-(14).

For part (2), let $Q_i = (I, E, p)$, and fix $s_i \in S_i$. Suppose that $p > \mu(E|S_{-i}(I))$. Let $I^* \in \mathcal{I}_i^*$ be such that $S_{-i}^*(I^*) = S_{-i}(I) \times W_{-i} \times S_c^*$, which exists by part (0) of Lemma 6. By part 2 of Lemma 4, $[S_{-i}^*(I^*)]_{\mu^*}$ and $[S_{-i}(I)]_{\mu}$ are basic events for ν^* and ν respectively. By Eqs. (12) and (13), and the fact that $\nu(S_{-i}(I)|[S_{-i}(I)]_{\mu}) > 0$ by Lemma 4 part 1, $E_{\nu^*(\cdot|[S_{-i}^*(I^*)]_{\mu^*})} U_i^*((s_i, b), \cdot) < E_{\nu^*(\cdot|[S_{-i}^*(I^*)]_{\mu^*})} U_i^*((s_i, p), \cdot)$. Furthermore, consider $J^* \in \mathcal{I}_i^*$ such that $[S_{-i}^*(J^*)]_{\mu^*} >^{\nu^*} [S_{-i}^*(I^*)]_{\mu^*}$. Then $\nu^*([S_{-i}^*(I^*)]_{\mu^*}|[S_{-i}^*(J^*)]_{\mu^*}) = 0$, because otherwise $[S_{-i}^*(I^*)]_{\mu^*} \geq^{\nu^*} [S_{-i}^*(J^*)]_{\mu^*}$, contradiction. A fortiori $\nu^*(S_{-i}^*(I^*)|[S_{-i}^*(J^*)]_{\mu^*}) = 0$; since $U_i^*((s_i, b), (s_{-i}, w_{-i}, s_c^*)) = U_i^*((s_i, p), (s_{-i}, w_{-i}, s_c^*))$ for $s_{-i} \notin S_{-i}(I)$ and all $w_{-i} \in W_{-i}$, $s_c^* \in S_c^*$, it follows that $E_{\nu^*(\cdot|[S_{-i}^*(J^*)]_{\mu^*})} U_i^*((s_i, b), \cdot) = E_{\nu^*(\cdot|[S_{-i}^*(J^*)]_{\mu^*})} U_i^*((s_i, p), \cdot)$. Hence, $\text{not } (s_i, b) \succ^{\mu^*} (s_i, p)$.

Finally, consider $J^* \in \mathcal{I}_i^*$ such that $E_{\nu^*(\cdot|[S_{-i}^*(J^*)]_{\mu^*})} U_i^*((s_i, b), \cdot) > E_{\nu^*(\cdot|[S_{-i}^*(J^*)]_{\mu^*})} U_i^*((s_i, p), \cdot)$. By the argument just given, it must be the case that $\nu^*([S_{-i}^*(I^*)]_{\mu^*}|[S_{-i}^*(J^*)]_{\mu^*}) > 0$: thus $[S_{-i}^*(I^*)]_{\mu^*} \geq^{\nu^*} [S_{-i}^*(J^*)]_{\mu^*}$. If $[S_{-i}^*(I^*)]_{\mu^*} =^{\nu^*} [S_{-i}^*(J^*)]_{\mu^*}$, then by Lemma 4 part 3 $S_{-i}^*(I^*) =^{\mu^*} S_{-i}^*(J^*)$, and so $[S_{-i}^*(I^*)]_{\mu^*} = [S_{-i}^*(J^*)]_{\mu^*}$, which contradicts the fact that $E_{\nu^*(\cdot|[S_{-i}^*(I^*)]_{\mu^*})} U_i^*((s_i, b), \cdot) < E_{\nu^*(\cdot|[S_{-i}^*(I^*)]_{\mu^*})} U_i^*((s_i, p), \cdot)$. Hence $[S_{-i}^*(I^*)]_{\mu^*} >^{\nu^*} [S_{-i}^*(J^*)]_{\mu^*}$. But then, since J^* was arbitrary, $(s_i, p) \succ^{\mu^*} (s_i, b)$. Thus, $(s_i, p) \succ^{\mu^*} (s_i, b)$, as claimed.

The case $\mu(E|S_{-i}(E)) > p$ is analogous, so the proof is omitted.

Finally, suppose that $Q_i = (I, E, p)$ and (s_i, b) is structurally rational in the elicitation game. Suppose that there is $t_i \in S_i$ such that $t_i \succ^{\mu} s_i$. Then, by (1), $(t_i, b) \succ^{\mu^*} (s_i, b)$: contradiction. Thus, s_i is structurally rational in the original game. Furthermore, suppose that $\mu(E|S_{-i}(I)) < p$: then (2) implies that $(s_i, p) \succ^{\mu^*} (s_i, b)$, contradiction. Thus, $\mu(E|S_{-i}(I)) \geq p$. The case of (s_i, p) structurally rational is analogous, so the proof is omitted. ■

References

- Frank J. Anscombe and Robert J. Aumann. A definition of subjective probability. *Annals of Mathematical Statistics*, 34:199–205, 1963.
- Geir B Asheim and Andrés Perea. Sequential and quasi-perfect rationalizability in extensive games. *Games and Economic Behavior*, 53(1):15–42, 2005.
- R.J. Aumann and J.H. Dreze. Assessing strategic risk. *American Economic Journal: Microeconomics*, 1(1):1–16, 2009.
- P. Battigalli and M. Siniscalchi. Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games. *Journal of Economic Theory*, 88(1):188–230, 1999.
- P. Battigalli and M. Siniscalchi. Strong belief and forward induction reasoning. *Journal of Economic Theory*, 106(2):356–391, 2002a.
- P. Battigalli and M. Siniscalchi. Strong Belief and Forward Induction Reasoning. *Journal of Economic Theory*, 106(2):356–391, 2002b.
- G.M. Becker, M.H. DeGroot, and J. Marschak. Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3), 1964.
- E. Ben-Porath. Rationality, Nash equilibrium and backwards induction in perfect-information games. *The Review of Economic Studies*, pages 23–46, 1997.
- L. Blume, A. Brandenburger, and E. Dekel. Lexicographic probabilities and choice under uncertainty. *Econometrica: Journal of the Econometric Society*, 59(1):61–79, 1991a.
- L. Blume, A. Brandenburger, and E. Dekel. Lexicographic probabilities and equilibrium refinements. *Econometrica: Journal of the Econometric Society*, pages 81–98, 1991b.

- Adam Brandenburger. The power of paradox: some recent developments in interactive epistemology. *International Journal of Game Theory*, 35(4):465–492, 2007.
- J. Brandts and G. Charness. The strategy versus the direct-response method: a first survey of experimental comparisons. *Experimental Economics*, 14(3):375–398, 2011.
- David J Cooper and John B Van Huyck. Evidence on the equivalence of the strategic and extensive form representation of games. *Journal of Economic Theory*, 110(2):290–308, 2003.
- Russell Cooper, Douglas V DeJong, Robert Forsythe, and Thomas W Ross. Forward induction in the battle-of-the-sexes games. *American Economic Review*, 83(5):1303–1316, 1993.
- Miguel A Costa-Gomes and Georg Weizsäcker. Stated beliefs and play in normal-form games. *The Review of Economic Studies*, 75(3):729–762, 2008.
- David Dillenberger. Preferences for one-shot resolution of uncertainty and allais-type behavior. *Econometrica*, 78(6):1973–2004, 2010.
- Larry G. Epstein and Stanley E. Zin. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57:937–969, 1989.
- Urs Fischbacher, Simon Gächter, and Simone Quercia. The behavioral validity of the strategy method in public good experiments. *Journal of Economic Psychology*, 33(4):897–913, 2012.
- Itzhak Gilboa and David Schmeidler. A derivation of expected utility maximization in the context of a game. *Games and Economic Behavior*, 44(1):172–182, 2003.
- David M Grether and Charles R Plott. Economic theory of choice and the preference reversal phenomenon. *The American Economic Review*, 69(4):623–638, 1979.
- Steffen Huck and Wieland Müller. Burning money and (pseudo) first-mover advantages: an experimental study on forward induction. *Games and Economic Behavior*, 51(1):109–127, 2005.

- E. Kohlberg and J.F. Mertens. On the strategic stability of equilibria. *Econometrica: Journal of the Econometric Society*, 54(5):1003–1037, 1986.
- David Kreps and Garey Ramey. Consistency, structural consistency, and sequential rationality. *Econometrica: Journal of the Econometric Society*, 55:1331–1348, 1987.
- David M. Kreps and Evan L. Porteus. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 46:185–200, 1978.
- D.M. Kreps and R. Wilson. Sequential equilibria. *Econometrica: Journal of the Econometric Society*, 50(4):863–894, 1982.
- R. Duncan Luce and Howard Raiffa. *Games and Decisions*. Wiley, New York, 1957.
- George J. Mailath, Larry Samuelson, and Jeroen Swinkels. Extensive form reasoning in normal form games. CARESS Working Paper 90–01, University of Pennsylvania, January 1990.
- Roger B Myerson. Refinements of the nash equilibrium concept. *International journal of game theory*, 7(2):73–80, 1978.
- Roger B. Myerson. An axiomatic derivation of subjective probability, utility, and evaluation functions. *Theory and Decision*, 11:339–352, 1979.
- Roger B. Myerson. Axiomatic foundations of bayesian decision theory. Discussion Paper 671, The Center for Mathematical Studies in Economics and Management Science, Northwestern University, January 1986.
- Yaw Nyarko and Andrew Schotter. An experimental study of belief learning using elicited beliefs. *Econometrica*, 70(3):971–1005, 2002.
- Martin J. Osborne and A. Rubinstein. *A Course on Game Theory*. MIT Press, Cambridge, MA, 1994.

- P.J. Reny. Backward induction, normal form perfection and explicable equilibria. *Econometrica*, 60(3):627–649, 1992. ISSN 0012-9682.
- A. Rényi. On a new axiomatic theory of probability. *Acta Mathematica Hungarica*, 6(3):285–335, 1955.
- Pedro Rey-Biel. Equilibrium play and best response to (stated) beliefs in normal form games. *Games and Economic Behavior*, 65(2):572–585, 2009.
- R.W. Rosenthal. Games of Perfect Information, Predatory Pricing and the Chain-Store Paradox. *Journal of Economic Theory*, 25(1):92–100, 1981. ISSN 0022-0531.
- A. Rubinstein. Comments on the interpretation of game theory. *Econometrica*, 59(4):909–924, 1991. ISSN 0012-9682.
- Leonard J. Savage. *The Foundations of Statistics*. Wiley, New York, 1954.
- Andrew Schotter, Keith Weigelt, and Charles Wilson. A laboratory investigation of multiperson rationality and presentation effects. *Games and Economic behavior*, 6(3):445–468, 1994.
- R. Selten. Ein oligopolexperiment mit preisvariation und investition. *Beiträge zur experimentellen Wirtschaftsforschung*, ed. by H. Sauermann, JCB Mohr (Paul Siebeck), Tübingen, pages 103–135, 1967.
- R. Selten. Reexamination of the perfectness concept for equilibrium points in extensive games. *International journal of game theory*, 4(1):25–55, 1975. ISSN 0020-7276.
- Marciano Siniscalchi. Foundations for structural preferences. mimeo, Northwestern University, 2016a.
- Marciano Siniscalchi. Structural rationality: applications to epistemic game theory. mimeo, Northwestern University, 2016b.

John B Van Huyck, Raymond C Battalio, and Richard O Beil. Tacit coordination games, strategic uncertainty, and coordination failure. *The American Economic Review*, 80(1):234–248, 1990.

Paul Weirich. Causal decision theory. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2016 edition, 2016.