



DEPARTMENT OF ECONOMICS Working Paper Series

Exploring the gap between perfect Bayesian equilibrium and sequential equilibrium

Giacomo Bonanno
University of California, Davis

May 25, 2016

Paper # 16-3

In (5) a solution concept for extensive-form games was introduced, called perfect Bayesian equilibrium (PBE), and shown to be a strict refinement of subgame-perfect equilibrium; it was also shown that, in turn, sequential equilibrium (SE) is a strict refinement of PBE. In (6) the notion of PBE was used to provide a characterization of SE in terms of a strengthening the two defining components of PBE (besides sequential rationality), namely AGM consistency and Bayes consistency. In this paper we explore the gap between PBE and SE by identifying solution concepts that lie strictly between PBE and SE; these solution concepts embody a notion of “conservative” belief revision. Furthermore, we provide a method for determining if a plausibility order on the set of histories is choice measurable, which is a necessary condition for a PBE to be a SE.

Department of Economics
One Shields Avenue
Davis, CA 95616
(530)752-0741

<http://wp.econ.ucdavis.edu/>

Exploring the gap between perfect Bayesian equilibrium and sequential equilibrium

Giacomo Bonanno

Department of Economics, University of California, Davis, USA
gfbonanno@ucdavis.edu

May 2016

Abstract

In (5) a solution concept for extensive-form games was introduced, called perfect Bayesian equilibrium (PBE), and shown to be a strict refinement of subgame-perfect equilibrium; it was also shown that, in turn, sequential equilibrium (SE) is a strict refinement of PBE. In (6) the notion of PBE was used to provide a characterization of SE in terms of a strengthening the two defining components of PBE (besides sequential rationality), namely AGM consistency and Bayes consistency. In this paper we explore the gap between PBE and SE by identifying solution concepts that lie strictly between PBE and SE; these solution concepts embody a notion of “conservative” belief revision. Furthermore, we provide a method for determining if a plausibility order on the set of histories is choice measurable, which is a necessary condition for a PBE to be a SE.

1 Introduction

In (5) a solution concept for general extensive-form games was introduced, called perfect Bayesian equilibrium (PBE), and it was shown that (i) PBE is a strict refinement of subgame-perfect equilibrium ((21)) and (ii) in general, the set of sequential equilibria ((15)) is a proper subset of the set of perfect Bayesian equilibria. The definition of PBE is based on two notions (besides sequential rationality): (1) the qualitative property of AGM-consistency relative to a plausibility order¹ and (2) the quantitative property of Bayes consistency. In (6) the notion of PBE was used to provide a new characterization of sequential equilibrium, in terms of a strengthening of both AGM consistency and Bayes consistency. In this paper we continue the analysis of PBE by exploring the gap between PBE and sequential equilibrium: we identify solution concepts that lie strictly between PBE and sequential equilibrium. These solution concepts capture the notion of revising one's beliefs in a "conservative" or "minimal" way.

The paper is organized as follows. The next section reviews the definition of PBE. In Section 3 we review the characterization of sequential equilibrium provided in (6) and offer a method for determining whether a plausibility order satisfies the property of "choice measurability", which is one of the two necessary and sufficient conditions for a PBE to be a sequential equilibrium. In Section 4 we introduce properties of the plausibility order that can be used to define solution concepts that lie between PBE and sequential equilibrium. Section 5 discusses related literature and Section 6 concludes. The proofs are given in the Appendix.

2 Perfect Bayesian equilibrium

In this section we recall the notion of perfect Bayesian equilibrium introduced in (5); we employ the same notation, which makes use of the history-based definition of extensive-form game (see, for example, (18)). As in (5), we restrict attention to *finite* extensive-form games with *perfect recall*.

¹As shown in (3), AGM-consistency can be derived from the primitive concept of a player's epistemic state, which encodes the player's initial beliefs and her disposition to revise those beliefs upon receiving (possibly unexpected) information. The existence of a plausibility order that rationalizes the epistemic state of each player guarantees that the belief revision policy of each player satisfies the so-called AGM axioms for rational belief revision, which were introduced in (1).

A total pre-order on a set H is a binary relation \lesssim which is complete² and transitive³. We write $h \sim h'$ as a short-hand for the conjunction: $h \lesssim h'$ and $h' \lesssim h$, and write $h < h'$ as a short-hand for the conjunction $h \lesssim h'$ and not $h' \lesssim h$.

Definition 1. *Given an extensive form, a plausibility order is a total pre-order \lesssim on the finite set of histories H that satisfies the following properties: $\forall h \in D$ [D is the set of decision histories],*

- PL1. $h \lesssim ha, \forall a \in A(h)$
 $[A(h)$ is the set of actions available at $h]$,
- PL2. (i) $\exists a \in A(h)$ such that $h \sim ha$,
(ii) $\forall a \in A(h)$, if $h \sim ha$ then, $\forall h' \in I(h), h' \sim h'a$,
 $[I(h)$ is the information set that contains $h]$,
- PL3. if history h is assigned to chance, then $h \sim ha, \forall a \in A(h)$.

The interpretation of $h \lesssim h'$ is that history h is at least as plausible as history h' ; thus $h < h'$ means that h is more plausible than h' and $h \sim h'$ means that h is just as plausible as h' .⁴ Property PL1 says that adding an action to a decision history h cannot yield a more plausible history than h itself. Property PL2 says that at every decision history h there is at least one action a which is “plausibility preserving” in the sense that adding a to h yields a history which is as plausible as h ; furthermore, any such action a performs the same role with any other history that belongs to the same information set as h . Property PL3 says that all the actions at a history assigned to chance are plausibility preserving.

An *assessment* is a pair (σ, μ) where σ is a behavior strategy profile and μ is a system of beliefs.⁵

² $\forall h, h' \in H$, either $h \lesssim h'$ or $h' \lesssim h$.

³ $\forall h, h', h'' \in H$, if $h \lesssim h'$ and $h' \lesssim h''$ then $h \lesssim h''$.

⁴As in (5) we use the notation $h \lesssim h'$ rather than the, perhaps more natural, notation $h \geq h'$, for two reasons: (1) it is the standard notation in the extensive literature that deals with AGM belief revision (for a recent survey of this literature see the special issue of the *Journal of Philosophical Logic*, Vol. 40 (2), April 2011) and (2) when representing the order \lesssim numerically it is convenient to assign lower values to more plausible histories. An alternative reading of $h \lesssim h'$ is “history h (weakly) precedes h' in terms of plausibility”.

⁵A behavior strategy profile is a list of probability distributions, one for every information set, over the actions available at that information set. A system of beliefs is a collection of probability distributions, one for every information set, over the histories in that information set.

Definition 2. Given an extensive-form, an assessment (σ, μ) is AGM-consistent if there exists a plausibility order \lesssim on the set of histories H such that:

(i) the actions that are assigned positive probability by σ are precisely the plausibility-preserving actions: $\forall h \in D, \forall a \in A(h),$

$$\sigma(a) > 0 \text{ if and only if } h \sim ha, \quad (P1)$$

(ii) the histories that are assigned positive probability by μ are precisely those that are most plausible within the corresponding information set: $\forall h \in D,$

$$\mu(h) > 0 \text{ if and only if } h \lesssim h', \forall h' \in I(h). \quad (P2)$$

If \lesssim satisfies properties P1 and P2 with respect to (σ, μ) , we say that \lesssim rationalizes (σ, μ) .

An assessment (σ, μ) is sequentially rational if, for every player i and every information set I of hers, player i 's expected payoff - given the strategy profile σ and her beliefs at I (as specified by μ) - cannot be increased by unilaterally changing her choice at I and possibly at information sets of hers that follow I .⁶

⁶ The precise definition is as follows. Let Z denote the set of terminal histories and, for every player i , let $U_i : Z \rightarrow \mathbb{R}$ be player i 's von Neumann-Morgenstern utility function. Given a decision history h , let $Z(h)$ be the set of terminal histories that have h as a prefix. Let $\mathbb{P}_{h,\sigma}$ be the probability distribution over $Z(h)$ induced by the strategy profile σ , starting from history h [that is, if z is a terminal history and $z = ha_1 \dots a_m$ then $\mathbb{P}_{h,\sigma}(z) = \prod_{j=1}^m \sigma(a_j)$]. Let I be an information set of player i and let $u_i(I|\sigma, \mu) = \sum_{h \in I} \mu(h) \sum_{z \in Z(h)} \mathbb{P}_{h,\sigma}(z) U_i(z)$ be player i 's expected utility at I if σ is played, given

her beliefs at I (as specified by μ). We say that player i 's strategy σ_i is *sequentially rational at I* if $u_i(I|\sigma_i, \sigma_{-i}, \mu) \geq u_i(I|\tau_i, \sigma_{-i}, \mu)$ for every strategy τ_i of player i (where σ_{-i} denotes the strategy profile of the players other than i). An assessment (σ, μ) is *sequentially rational* if, for every player i and for every information set I of player i , σ_i is sequentially rational at I . Note that there are two definitions of sequential rationality: the *weakly local* one - which is the one adopted here - according to which at an information set a player can contemplate changing her choice not only there but possibly also at subsequent information sets of hers, and a *strictly local* one, according to which at an information set a player contemplates changing her choice only there. If the definition of perfect Bayesian equilibrium (Definition 4 below) is modified by using the strictly local definition of sequential rationality, then an extra condition needs to be added, namely the "pre-consistency" condition identified in (13) and (19) as being necessary and sufficient for the equivalence of the two notions. For simplicity we have chosen the weakly local definition.

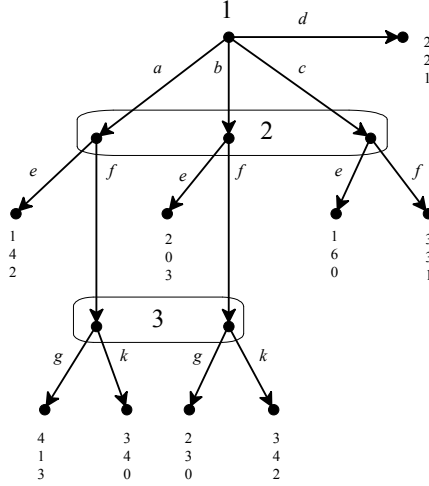


Figure 1:

Consider the extensive-form game shown in Figure 1⁷ and the assessment (σ, μ) where $\sigma = (d, e, g)$ and μ is the following system of beliefs: $\mu(a) = 0$, $\mu(b) = \frac{1}{3}$, $\mu(c) = \frac{2}{3}$ and $\mu(af) = \mu(bf) = \frac{1}{2}$. This assessment is AGM-consistent, since it is rationalized by the following plausibility order:⁸

$$\left(\begin{array}{cc} \text{most plausible} & \emptyset, d \\ & b, c, be, ce \\ & a, ae \\ & af, bf, cf, afg, bfg \\ \text{least plausible} & afk, bfk \end{array} \right) \quad (1)$$

Furthermore (σ, μ) is sequentially rational.⁹

⁷Rounded rectangles represent information sets and the payoffs are listed in the following order: Player 1's payoff at the top, Player 2's payoff in the middle and Player 3's payoff at the bottom.

⁸We use the following convention to represent a total pre-order: if the row to which history h belongs is above the row to which h' belongs, then $h < h'$ (h is more plausible than h') and if h and h' belong to the same row then $h \sim h'$ (h is as plausible as h'). \emptyset denotes the empty history, which corresponds to the root of the tree. In (1) the plausibility-preserving actions are d, e and g ; the most plausible histories in the information set $\{a, b, c\}$ are b and c and the two histories in the information set $\{af, bf\}$ are equally plausible.

⁹Given σ , for Player 1 d yields a payoff of 2 while a and c yield 1 and b yields 2; thus d is

The property of AGM-consistency imposes restrictions on the support of the behavior strategy σ and on the support of the system of beliefs μ . The following property imposes constraints on how probabilities can be distributed over those supports.

Definition 3. Given an extensive form, let \lesssim be a plausibility order that rationalizes the assessment (σ, μ) . We say that (σ, μ) is Bayes consistent (or Bayesian) relative to \lesssim if, for every equivalence class E of \lesssim that contains some decision history h with $\mu(h) > 0$ [that is, $E \cap D_\mu^+ \neq \emptyset$, where $D_\mu^+ = \{h \in D : \mu(h) > 0\}$], there exists a probability density function $v_E : H \rightarrow [0, 1]$ (recall that H is a finite set) such that:

- B1. $v_E(h) > 0$ if and only if $h \in E \cap D_\mu^+$.
- B2. If $h, h' \in E \cap D_\mu^+$ and $h' = ha_1...a_m$ (that is, h is a prefix of h') then $v_E(h') = v_E(h) \times \sigma(a_1) \times \dots \times \sigma(a_m)$.
- B3. If $h \in E \cap D_\mu^+$, then, $\forall h' \in I(h)$, $\mu(h') = v_E(h' | I(h)) \stackrel{\text{def}}{=} \frac{v_E(h')}{\sum_{h'' \in I(h)} v_E(h'')}$.

Property B1 requires that $v_E(h) > 0$ if and only if $h \in E$ and $\mu(h) > 0$. Property B2 requires v_E to be consistent with the strategy profile σ in the sense that if $h, h' \in E$, $\mu(h) > 0$, $\mu(h') > 0$ and $h' = ha_1...a_m$ then the probability that v_E assigns to h' is equal to the probability that v_E assigns to h multiplied by the probabilities (according to σ) of the actions that lead from h to h' .¹⁰ Property B3 requires the system of beliefs μ to satisfy Bayes' rule in the sense that if $h \in E$ and $\mu(h) > 0$ (so that E is the equivalence class of the most plausible elements of $I(h)$) then, for every history $h' \in I(h)$, $\mu(h')$ (the probability assigned to h' by μ) coincides with the probability of h' conditional on $I(h)$ using the probability density function v_E .¹¹

Consider again the game of Figure 1, and the assessment (σ, μ) where $\sigma = (d, e, g)$ and $\mu(a) = 0$, $\mu(b) = \frac{1}{3}$, $\mu(c) = \frac{2}{3}$ and $\mu(af) = \mu(bf) = \frac{1}{2}$. Let \lesssim be the plausibility order (1) given above, which rationalizes (σ, μ) . Then (σ, μ) is Bayes consistent relative to \lesssim . In fact, we have that $D_\mu^+ = \{\emptyset, b, c, af, bf\}$ and the equivalence classes of \lesssim that have a non-empty intersection with D_μ^+ are $E_1 = \{\emptyset, d\}$, $E_2 = \{b, c, be, ce\}$ and $E_3 = \{af, bf, cf, afg, bfg\}$. Let $v_{E_1}(\emptyset) = 1$,

sequentially rational. Given σ and μ , at her information set $\{a, b, c\}$ with e Player 2 obtains an expected payoff of 4 while with f her expected payoff is 3; thus e is sequentially rational. Given μ , at his information set $\{af, bf\}$, Player 3's expected payoff from playing with g is 1.5 while his expected payoff from playing with k is 1; thus g is sequentially rational.

¹⁰Note that if $h, h' \in E$ and $h' = ha_1...a_m$, then $\sigma(a_j) > 0$, for all $j = 1, \dots, m$. In fact, since $h' \sim h$, every action a_j is plausibility preserving and therefore, by Property P1 of Definition 2, $\sigma(a_j) > 0$.

¹¹For an interpretation of the probabilities $v_E(h)$ see (6).

$v_{E_2}(b) = \frac{1}{3}$, $v_{E_2}(c) = \frac{2}{3}$ and $v_{E_3}(af) = v_{E_3}(bf) = \frac{1}{2}$. Then the three probability density functions v_{E_1} , v_{E_2} and v_{E_3} satisfy the properties of Definition 3 and hence (σ, μ) is Bayes consistent relative to \preceq .

Definition 4. An assessment (σ, μ) is a perfect Bayesian equilibrium (PBE) if it is sequentially rational, it is rationalized by a plausibility order on the set of histories and is Bayes consistent relative to it.

We saw above that, for the game illustrated in Figure 1, the assessment (σ, μ) where $\sigma = (d, e, g)$ and $\mu(a) = 0$, $\mu(b) = \frac{1}{3}$, $\mu(c) = \frac{2}{3}$ and $\mu(af) = \mu(bf) = \frac{1}{2}$ is sequentially rational, it is rationalized by the plausibility order (1) and is Bayes consistent relative to it. Thus it is a perfect Bayesian equilibrium.

Remark 1. It is proved in (5) that if (σ, μ) is a perfect Bayesian equilibrium then σ is a subgame-perfect equilibrium and that every sequential equilibrium is a perfect Bayesian equilibrium. Furthermore, the notion of PBE is a strict refinement of subgame-perfect equilibrium and sequential equilibrium is a strict refinement of PBE.

In the next section we review the characterization of sequential equilibrium in terms of PBE provided in (6).

3 Sequential equilibrium

First we recall the definition of sequential equilibrium. An assessment (σ, μ) is *KW-consistent* ('KW' stands for 'Kreps-Wilson') if there is an infinite sequence $\langle \sigma^1, \dots, \sigma^m, \dots \rangle$ of completely mixed behavior strategy profiles such that, letting μ^m be the unique system of beliefs obtained from σ^m by applying Bayes' rule,¹² $\lim_{m \rightarrow \infty} (\sigma^m, \mu^m) = (\sigma, \mu)$. An assessment (σ, μ) is a *sequential equilibrium* if it is KW-consistent and sequentially rational. It was shown in (6) that sequential equilibrium can be characterized as a strengthening of PBE based on two properties: (1) a property of the plausibility order that constrains the supports of the belief system and (2) a strengthening of the notion of Bayes consistency, that imposes constraints on how the probabilities can be distributed over those supports. The details are given below.

¹²That is, for every $h \in D \setminus \{\emptyset\}$, $\mu^m(h) = \frac{\prod_{a \in A_h} \sigma^m(a)}{\sum_{h' \in I(h)} \prod_{a \in A_{h'}} \sigma^m(a)}$ (where A_h is the set of actions that occur in history h). Since σ^m is completely mixed, $\sigma^m(a) > 0$ for every $a \in A$ and thus $\mu^m(h) > 0$ for all $h \in D \setminus \{\emptyset\}$.

Given a plausibility order \lesssim on the finite set of histories H , a function $F : H \rightarrow \mathbb{N}$ (where \mathbb{N} denotes the set of non-negative integers) is said to be an *ordinal integer-valued representation* of \lesssim if, for every $h, h' \in H$,

$$F(h) \leq F(h') \text{ if and only if } h \lesssim h'. \quad (2)$$

Since H is finite, the set of ordinal integer-valued representations is non-empty. A particular ordinal integer-valued representation, which we will call *canonical* and denote by ρ , is defined as follows.

Definition 5. Let $H_0 = \{h \in H : h \lesssim x, \forall x \in H\}$, $H_1 = \{h \in H \setminus H_0 : h \lesssim x, \forall x \in H \setminus H_0\}$ and, in general, for every integer $k \geq 1$, $H_k = \{h \in H \setminus H_0 \cup \dots \cup H_{k-1} : h \lesssim x, \forall x \in H \setminus H_0 \cup \dots \cup H_{k-1}\}$. Thus H_0 is the equivalence class of \lesssim containing the most plausible histories, H_1 is the equivalence class containing the most plausible among the histories left after removing those in H_0 , etc.¹³ The canonical ordinal integer-valued representation of \lesssim , $\rho : H \rightarrow \mathbb{N}$, is given by

$$\rho(h) = k \text{ if and only if } h \in H_k. \quad (3)$$

We call $\rho(h)$ the rank of history h .

Instead of an ordinal integer-valued representation of the plausibility order one could seek a *cardinal* representation which, besides (2), satisfies the following property: if h and h' belong to the same information set [that is, $h' \in I(h)$] and $a \in A(h)$, then

$$F(h') - F(h) = F(h'a) - F(ha). \quad (CM)$$

If we think of F as measuring the “plausibility distance” between histories, then we can interpret CM as a distance-preserving condition: the plausibility distance between two histories in the same information set is preserved by the addition of the same action.

Definition 6. A plausibility order \lesssim on the set of histories H is choice measurable if it has at least one integer-valued representation that satisfies property CM.

For example, the plausibility order (1) is not choice measurable, since any integer-valued representation F of it must be such that $F(a) - F(b) > 0$ and $F(af) - F(bf) = 0$.

¹³Since H is finite, there is an $m \in \mathbb{N}$ such that $\{H_0, \dots, H_m\}$ is a partition of H and, for every $j, k \in \mathbb{N}$, with $j < k \leq m$, and for every $h, h' \in H$, if $h \in H_j$ and $h' \in H_k$ then $h < h'$.

Let (σ, μ) be an assessment which is rationalized by a plausibility order \lesssim . As before, let D_μ^+ be the set of decision histories to which μ assigns positive probability: $D_\mu^+ = \{h \in D : \mu(h) > 0\}$. Let \mathcal{E}_μ^+ be the set of equivalence classes of \lesssim that have a non-empty intersection with D_μ^+ . Clearly \mathcal{E}_μ^+ is a non-empty, finite set. Suppose that (σ, μ) is Bayesian relative to \lesssim and let $\{v_E\}_{E \in \mathcal{E}_\mu^+}$ be a collection of probability density functions that satisfy the properties of Definition 3. We call a probability density function $v : D \rightarrow (0, 1]$ (recall that D is the set of all decision histories) a *full-support common prior* of $\{v_E\}_{E \in \mathcal{E}_\mu^+}$ if, for every $E \in \mathcal{E}_\mu^+$, $v_E(\cdot) = v(\cdot \mid E \cap D_\mu^+)$, that is, for all $h \in E \cap D_\mu^+$, $v_E(h) = \frac{v(h)}{\sum_{h' \in E \cap D_\mu^+} v(h')}$. Note that a full support common prior assigns positive probability to *all* decision histories, not only to those in D_μ^+ .

Definition 7. Consider an extensive form. Let (σ, μ) be an assessment which is rationalized by the plausibility order \lesssim and is Bayesian relative to it and let $\{v_E\}_{E \in \mathcal{E}_\mu^+}$ be a collection of probability density functions that satisfy the properties of Definition 3. We say that (σ, μ) is *uniformly Bayesian relative to \lesssim* if there exists a full-support common prior $v : D \rightarrow (0, 1]$ of $\{v_E\}_{E \in \mathcal{E}_\mu^+}$ that satisfies the following properties.

- UB1. If $a \in A(h)$ and $ha \in D$, then
(i) $v(ha) \leq v(h)$ and (ii) if $\sigma(a) > 0$ then $v(ha) = v(h) \times \sigma(a)$.
- UB2. If $a \in A(h)$, h and h' belong to the same information set and $ha, h'a \in D$
then $\frac{v(h)}{v(h')} = \frac{v(ha)}{v(h'a)}$.

We call such a function v a *uniform full-support common prior* of $\{v_E\}_{E \in \mathcal{E}_\mu^+}$.

UB1 requires that the common prior v be consistent with the strategy profile σ , in the sense that if $\sigma(a) > 0$ then $v(ha) = v(h) \times \sigma(a)$ (thus extending Property B2 of Definition 3 from D_μ^+ to D). UB2 requires that the relative probability, according to the common prior v , of any two histories that belong to the same information set remain unchanged by the addition of the same action.

It is shown in (6) that choice measurability and uniform Bayesian consistency are independent properties. The following proposition is proved in (6).

Proposition 8. (I) and (II) below are equivalent:

- (I) (σ, μ) is a perfect Bayesian equilibrium which is rationalized by a choice measurable plausibility order and is uniformly Bayesian relative to it.
- (II) (σ, μ) is a sequential equilibrium.

We now provide a method for determining if a plausibility order is choice measurable. More generally, will provide a necessary and sufficient condition that applies not only to plausibility orders over sets of histories in a game but to a more general class of structures.

Let S be an arbitrary finite set and let \preceq be a total pre-order on S . Let $S\backslash\sim$ be the set of equivalence classes of S . If $s \in S$, the equivalence class of s is denoted by $[s] = \{t \in S : s \sim t\}$ (where, as before, $s \sim t$ is a short-hand for “ $s \preceq t$ and $t \preceq s$ ”); thus $S\backslash\sim = \{[s] : s \in S\}$. Let \doteq be an equivalence relation on $S\backslash\sim \times S\backslash\sim$. The interpretation of $([s_1], [s_2]) \doteq ([t_1], [t_2])$ is that the distance between the equivalence classes $[s_1]$ and $[s_2]$ is required to be equal to the distance between the equivalence classes $[t_1]$ and $[t_2]$.

Remark 2. In the special case of a plausibility order \preceq on the set of histories H of a game, if E_1, E_2, F_1 and F_2 are equivalence classes of \preceq then $(E_1, E_2) \doteq (F_1, F_2)$ if and only if there exist two decision histories $h, h' \in H$ that belong to the same information set [$h' \in I(h)$] and a non-plausibility-preserving action $a \in A(h)$ such that either $h \in E_1, h' \in E_2, ha \in F_1$ and $h'a \in F_2$ or $ha \in E_1, h'a \in E_2, h \in F_1$ and $h' \in F_2$ (see property CM above).

The general problem that we are addressing is the following.

Problem 9. Given a pair (\preceq, \doteq) , where \preceq is a total pre-order on a finite set S and \doteq is an equivalence relation on the set of pairs of equivalence classes of \preceq , determine whether there exists a function $F : S \rightarrow \mathbb{N}$ such that, for all $s, t, x, y \in S$, (1) $F(s) \leq F(t)$ if and only if $s \preceq t$ and (2) if $([s], [t]) \doteq ([x], [y])$, with $s < t$ and $x < y$, then $F(t) - F(s) = F(y) - F(x)$.

Instead of expressing the equivalence relation \doteq in terms of pairs of elements of $S\backslash\sim$, we shall express it in terms of pairs of numbers (j, k) obtained by using the canonical ordinal representation ρ of \preceq .¹⁴ That is, if $s_1, s_2, t_1, t_2 \in S$ and $([s_1], [s_2]) \doteq ([t_1], [t_2])$ then we shall write this as $(\rho(s_1), \rho(s_2)) \doteq (\rho(t_1), \rho(t_2))$.

¹⁴As in Definition 5, let $S_0 = \{s \in S : s \preceq t, \forall t \in S\}$, and, for every integer $k \geq 1$, $S_k = \{h \in S \setminus S_0 \cup \dots \cup S_{k-1} : s \preceq t, \forall t \in S \setminus S_0 \cup \dots \cup S_{k-1}\}$. The canonical ordinal integer-valued representation of \preceq , $\rho : S \rightarrow \mathbb{N}$, is given by $\rho(s) = k$ if and only if $s \in S_k$.

For example, let $S = \{a, b, c, d, e, f, g, h, \ell, m\}$ and let \lesssim be as shown in (4) below, together with the corresponding canonical representation ρ :¹⁵

$$\left(\begin{array}{cc} \lesssim : & \rho : \\ a & 0 \\ b, c & 1 \\ d & 2 \\ e & 3 \\ f, g & 4 \\ h, \ell & 5 \\ m & 6 \end{array} \right) \quad (4)$$

If the equivalence relation \doteq contains the following pairs:¹⁶

$$\begin{array}{ll} ([a], [b]) \doteq ([h], [m]) & (0, 1) \doteq (5, 6) \\ ([a], [b]) \doteq ([e], [f]) & (0, 1) \doteq (3, 4) \\ ([a], [d]) \doteq ([f], [m]) & (0, 2) \doteq (4, 6) \\ ([b], [e]) \doteq ([e], [f]) & (1, 3) \doteq (3, 4) \\ ([b], [e]) \doteq ([f], [m]) & (1, 3) \doteq (4, 6) \end{array} \quad \text{then we express them (using } \rho) \text{ as} \quad (5)$$

A *bag* (or *multiset*) is a generalization of the notion of set in which members are allowed to appear more than once. An example of a bag is $\{1, 2, 2, 3, 4, 4, 5, 6\}$. Given two bags B_1 and B_2 their *union*, denoted by $B_1 \uplus B_2$, is the bag that contains those elements that occur in either B_1 or B_2 and, furthermore, the number of times that each element occurs in $B_1 \uplus B_2$ is equal to the number of times it occurs in B_1 plus the number of times it occurs in B_2 . For instance, if $B_1 = \{1, 2, 2, 3, 4, 4, 5, 6\}$ and $B_2 = \{2, 3, 6, 6\}$ then $B_1 \uplus B_2 = \{1, 2, 2, 2, 3, 3, 4, 4, 5, 6, 6, 6\}$. We say that B_1 is a *proper sub-bag* of B_2 , denoted by $B_1 \sqsubset B_2$, if $B_1 \neq B_2$ and each element that occurs in B_1 occurs also, and at least as many times, in B_2 . For example, $\{1, 2, 4, 4, 5, 6\} \sqsubset \{1, 1, 2, 4, 4, 5, 5, 6\}$.

Given a pair (i, j) with $i < j$, we associate with it the set $B_{(i,j)} = \{i+1, i+2, \dots, j\}$. For example, $B_{(2,5)} = \{3, 4, 5\}$. Given a set of pairs $P = \{(i_1, j_1), (i_2, j_2), \dots, (i_m, j_m)\}$ (with $i_k < j_k$, for every $k = 1, \dots, m$) we associate with it the bag $B_P = B_{(i_1, j_1)} \uplus$

¹⁵Thus $a < x$ for every $x \in S \setminus \{a\}$, $[b] = \{b, c\}$, $b < d$, etc.

¹⁶For example, \doteq is the smallest reflexive, symmetric and transitive relation that contains the pairs given in (5).

$B_{(i_2, j_2)} \uplus \dots \uplus B_{(i_m, j_m)}$. For example, if $P = \{(0, 2), (1, 4), (2, 5)\}$ then $B_P = \{1, 2\} \uplus \{2, 3, 4\} \uplus \{3, 4, 5\} = \{1, 2, 2, 3, 3, 4, 4, 5\}$.

Definition 10. For every element of \doteq , expressed (using the canonical representation ρ) as $(i, j) \doteq (k, \ell)$ (with $i < j$ and $k < \ell$), the equation corresponding to it is $x_{i+1} + x_{i+2} + \dots + x_j = x_{k+1} + x_{k+2} + \dots + x_\ell$. By the system of equations corresponding to \doteq we mean the set of all such equations.

For example, consider the total pre-order given in (4) and the following equivalence relation \doteq (expressed in terms of ρ and omitting the reflexive pairs):

$$\{(0, 3) \doteq (2, 4), (2, 4) \doteq (0, 3), (2, 4) \doteq (3, 5), (3, 5) \doteq (2, 4), (0, 3) \doteq (3, 5), (3, 5) \doteq (0, 3)\}$$

Then the corresponding system of equations is given by:

$$\begin{aligned} x_1 + x_2 + x_3 &= x_3 + x_4 \\ x_3 + x_4 &= x_1 + x_2 + x_3 \\ x_3 + x_4 &= x_4 + x_5 \\ x_4 + x_5 &= x_3 + x_4 \\ x_1 + x_2 + x_3 &= x_4 + x_5 \\ x_4 + x_5 &= x_1 + x_2 + x_3 \end{aligned} \tag{6}$$

We are now ready to state the solution to Problem 9. The proof is given in Appendix A.

Proposition 11. Given a pair (\lesssim, \doteq) , where \lesssim is a total pre-order on a finite set S and \doteq is an equivalence relation on the set of pairs of equivalence classes of \lesssim , (A), (B) and (C) below are equivalent.

(A) There is a function $F : S \rightarrow \mathbb{N}$ such that, for all $s, t, x, y \in S$, (1) $F(s) \leq F(t)$ if and only if $s \lesssim t$ and (2) if $([s], [t]) \doteq ([x], [y])$, with $s < t$ and $x < y$, then $F(t) - F(s) = F(y) - F(x)$,

(B) The system of equations corresponding to \doteq (Definition 10) has a solution consisting of positive integers.

(C) There is no sequence $\langle ((i_1, j_1) \doteq (k_1, \ell_1)), \dots, ((i_m, j_m) \doteq (k_m, \ell_m)) \rangle$ in \doteq (expressed in terms of the canonical representation ρ of \lesssim) such that $B_{\text{left}} \sqsubset B_{\text{right}}$ where $B_{\text{left}} = B_{(i_1, j_1)} \uplus \dots \uplus B_{(i_m, j_m)}$ and $B_{\text{right}} = B_{(k_1, \ell_1)} \uplus \dots \uplus B_{(k_m, \ell_m)}$.

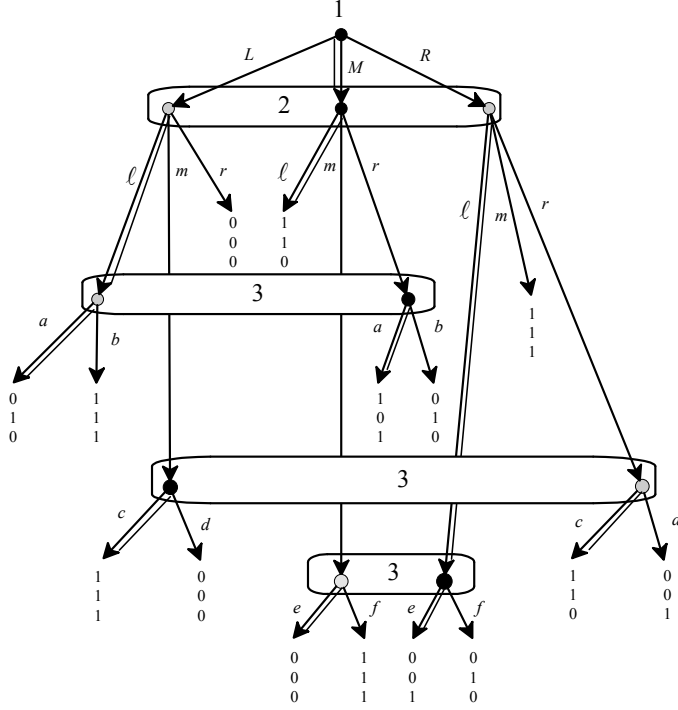


Figure 2:

As an application of Proposition 11 consider the game of Figure 2 (which is based on an example discussed in (2) and (14)) and the assessment (σ, μ) where $\sigma = (M, \ell, a, c, e)$ (highlighted in Figure 2 by doubling the corresponding edges) and $\mu(x) = 1$ for $x \in \{M, Mr, Lm, R\ell\}$ (and $\mu(x) = 0$ for every other decision history $x \neq \emptyset$; the decision histories x such that $\mu(x) > 0$ are shown as black nodes and the decision histories x such that $\mu(x) = 0$ are shown as gray nodes). This assessment is rationalized by the following plausibility order (shown

together with the canonical integer-valued representation ρ):

$$\left(\begin{array}{cc} \lesssim : & \rho : \\ \text{most plausible} & \emptyset, M, M\ell & 0 \\ & R, R\ell, R\ell e & 1 \\ & Mm, Mm e & 2 \\ & Mr, Mr a & 3 \\ & L, L\ell, L\ell a & 4 \\ & Rm & 5 \\ & Lm, Lm c & 6 \\ & Rr, Rr c & 7 \\ & Lr & 8 \\ & R\ell f & 9 \\ & Mm f & 10 \\ & Lm d & 11 \\ & Rr d & 12 \\ & Mr b & 13 \\ \text{least plausible} & L\ell b & 14 \end{array} \right) \quad (7)$$

Two elements of \doteq are $([M], [R]) \doteq ([Mr], [Rr])$ and $([Mm], [Lm]) \doteq ([M], [L])$ (see Remark 2), which - expressed in terms of the canonical ordinal representation ρ - can be written as

$$\begin{aligned} (0, 1) &\doteq (3, 7) \\ (2, 6) &\doteq (0, 4) \end{aligned}$$

Then $B_{left} = \{1\} \cup \{3, 4, 5, 6\} = \{1, 3, 4, 5, 6\}$ and $B_{right} = \{4, 5, 6, 7\} \cup \{1, 2, 3, 4\} = \{1, 2, 3, 4, 5, 6, 7\}$. Thus, since $B_{left} \subset B_{right}$, by Part (C) of Proposition 11 \lesssim is not choice measurable.

As a further application of Proposition 11 consider the total pre-order \lesssim given in (4) together with the subset of the equivalence relation \doteq given in (5). Then there is no cardinal representation of \lesssim that satisfies the constraints expressed by \doteq , because of Part (C) of the above proposition and the following sequence:¹⁷

$$\langle ((0, 1) \doteq (3, 4)), ((1, 3) \doteq (4, 6)), ((3, 4) \doteq (1, 3)), ((4, 6) \doteq (0, 2)) \rangle$$

¹⁷By symmetry of \doteq , we can express the third and fourth constraints as $(4, 6) \doteq (0, 2)$ and $(3, 4) \doteq (1, 3)$ instead of $(0, 2) \doteq (4, 6)$ and $(1, 3) \doteq (3, 4)$, respectively.

where $B_{left} = \{1\} \cup \{2, 3\} \cup \{4\} \cup \{5, 6\} = \{1, 2, 3, 4, 5, 6\} \sqsubset B_{right} = \{4\} \cup \{5, 6\} \cup \{2, 3\} \cup \{1, 2\} = \{1, 2, 2, 3, 4, 5, 6\}$.

In fact, the above sequence corresponds to the following system of equations:

$$\begin{array}{ll}
 x_1 = x_4 & \text{corresponding to } (0, 1) \doteq (3, 4) \\
 x_2 + x_3 = x_5 + x_6 & \text{corresponding to } (1, 3) \doteq (4, 6) \\
 x_4 = x_2 + x_3 & \text{corresponding to } (3, 4) \doteq (1, 3) \\
 x_5 + x_6 = x_1 + x_2 & \text{corresponding to } (4, 6) \doteq (0, 2)
 \end{array}$$

Adding the four equations we get $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = x_1 + 2x_2 + x_3 + x_4 + x_5 + x_6$ which simplifies to $0 = x_2$, which is incompatible with a positive solution.

Remark 3. In (9) an algorithm is provided for determining whether a system of linear equations has a positive solution and for calculating such a solution if one exists. Furthermore, if the coefficients of the equations are integers and a positive solution exists, then the algorithm yields a solution consisting of positive integers.

4 Exploring the gap between PBE and sequential equilibrium

The notion of perfect Bayesian equilibrium incorporates (through the property of AGM-consistency) a belief revision policy which can be interpreted either as the epistemic state of an external observer¹⁸ or as a belief revision policy which is shared by all the players.¹⁹ For example, the perfect Bayesian equilibrium considered in Section 2 for the game of Figure 1, namely $\sigma = (d, e, g)$ and $\mu(a) = 0$, $\mu(b) = \frac{1}{3}$, $\mu(c) = \frac{2}{3}$, $\mu(af) = \mu(bf) = \frac{1}{2}$ reflects the following belief revision policy: the initial beliefs are that Player 1 will play d ; conditional on learning that Player 1 did not play d , the observer would become convinced that Player 1 played either b or c (that is, she would judge a to be less plausible than b and she would consider c to be as plausible as b) and would expect Player 2 to play e ; upon learning that (Player 1 did not play d and) Player 2 played f , the observer would become convinced that Player 1 played either a or b , hence judging af to be as plausible as bf , thereby modifying her earlier judgment that a was less plausible than b . Although such a belief revision policy does not violate the rationality constraints introduced in (1), it does

¹⁸For example, (14) adopts this interpretation.

¹⁹For such an interpretation see (3).

involve a belief change that is not “minimal” or “conservative”. Such “non-minimal” belief changes can be ruled out by imposing the following restriction on the plausibility order: if h and h' belong to the same information set [that is, $h' \in I(h)$] and a is an action available at h [$a \in A(h)$], then

$$h \lesssim h' \text{ if and only if } ha \lesssim h'a. \quad (IND_1)$$

IND_1 says that if h is deemed to be at least as plausible as h' then the addition of any available action a must preserve this judgment, in the sense that ha must be deemed to be at least as plausible as $h'a$, and *vice versa*; it can also be viewed as an “independence” condition, in the sense that observation of a new action cannot lead to a change in the relative plausibility of previous histories.²⁰ Any plausibility order that rationalizes the assessment $\sigma = (d, e, g)$ and $\mu(a) = 0$, $\mu(b) = \frac{1}{3}$, $\mu(c) = \frac{2}{3}$, $\mu(af) = \mu(bf) = \frac{1}{2}$ for the game of Figure 1 must violate IND_1 .

We can obtain a strengthening of the notion of perfect Bayesian equilibrium (Definition 4) by (1) adding property IND_1 and (2) strengthening Bayes consistency (Definition 3) to uniform Bayesian consistency (Definition 7).

Definition 12. Consider an extensive-form game. An assessment (σ, μ) is a weakly independent perfect Bayesian equilibrium if it is sequentially rational, it is rationalized by a plausibility order that satisfies IND_1 and is uniformly Bayesian relative to that plausibility order.

As an example of a weakly independent PBE consider the game of Figure 3 and the assessment (σ, μ) where $\sigma = (c, d, g, \ell)$ and $\mu(b) = \mu(ae) = \mu(bf) = 1$ (thus $\mu(a) = \mu(af) = \mu(be) = 0$). This assessment is sequentially rational and is rationalized by the following plausibility order:

²⁰Note, however, that IND_1 is compatible with the following: $a < b$ [with $b \in I(a)$] and $bc < ad$ [with $bc \in I(ad)$, $c, d \in A(ad)$, $c \neq d$], which embodies a reversal of judgment on the relative plausibility of a and b .

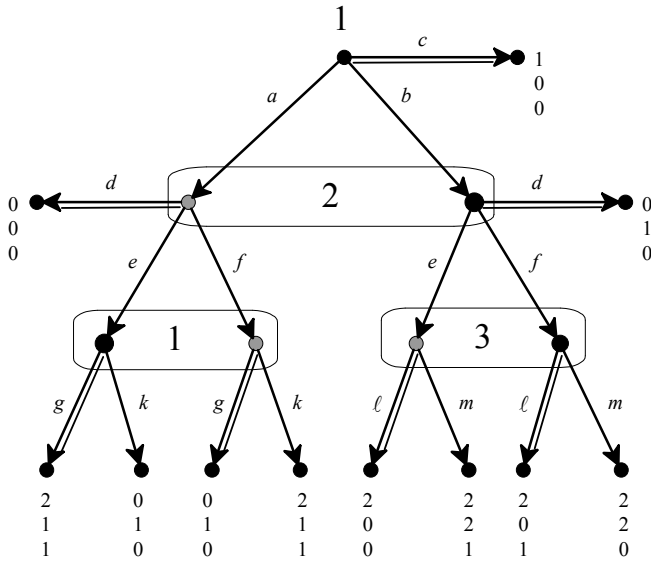


Figure 3:

$$\left(\begin{array}{l} \text{most plausible} \\ \emptyset, c \\ b, bd \\ a, ad \\ bf, bfl \\ be, bel \\ ae, aeg \\ af, afg \\ bfm \\ bem \\ aek \\ \text{least plausible} \\ afk \end{array} \right) \quad (8)$$

It is straightforward to check that plausibility order (8) satisfies IND_1 .²¹ To

²¹For instance, $b < a$ and we have that both $be < ae$ and $bf < af$.

see that (σ, μ) is uniformly Bayesian relative to plausibility order (8), note that $D_\mu^+ = \{\emptyset, b, ae, bf\}$ and thus the only equivalence classes that have a non-empty intersection with D_μ^+ are $E_1 = \{\emptyset, c\}$, $E_2 = \{b, bd\}$, $E_3 = \{ae, aeg\}$ and $E_4 = \{bf, bfl\}$. Letting $v_{E_1}(\emptyset) = 1$, $v_{E_2}(b) = 1$, $v_{E_3}(ae) = 1$ and $v_{E_4}(bf) = 1$, this collection of probability distributions satisfies the Properties of Definition 3. Let v be the uniform distribution over the set of decision histories $D = \{\emptyset, a, b, ae, af, be, bf\}$ (thus $v(h) = \frac{1}{7}$ for every $h \in D$). Then v is a full support common prior of the collection $\{v_{E_i}\}_{i \in \{1,2,3,4\}}$ and satisfies Properties UB1 and UB2 of Definition 7. Note, however, that (σ, μ) is not a sequential equilibrium. This can be shown by appealing to the following lemma, which is proved in the Appendix.

Lemma 13. *Let \lesssim be a plausibility order over the set H of histories of an extensive-form game and let $F : H \rightarrow \mathbb{N}$ be an integer-valued representation of \lesssim (that is, for all $h, h' \in H$, $F(h) \leq F(h')$ if and only if $h \lesssim h'$). Then the following are equivalent:*

(A) *F satisfies Property CM (Definition 6)*

(B) *F satisfies the following property: for all $h, h' \in H$ and $a, b \in A(h)$, if $h' \in I(h)$ then*

$$F(hb) - F(ha) = F(h'b) - F(h'a). \quad (CM')$$

Using Lemma 13 we can prove that the assessment (σ, μ) where $\sigma = (c, d, g, \ell)$ and $\mu(b) = \mu(ae) = \mu(bf) = 1$, for the game of Figure 3, is not a sequential equilibrium. By Proposition 8 it will be sufficient to show that (σ, μ) cannot be rationalized by a choice measurable plausibility order (Definition 6). Let \lesssim be a plausibility order that rationalizes (σ, μ) and let F be an integer-valued representation of \lesssim . Then, by (P2) of Definition 2, it must be that $ae < af$ (because $\mu(ae) > 0$ and $\mu(af) = 0$) and $bf < be$ (because $\mu(bf) > 0$ and $\mu(be) = 0$); thus $F(ae) - F(af) < 0$ and $F(be) - F(bf) > 0$, so that F violates property CM'; hence, by Lemma 13, F violates property CM and thus \lesssim is not choice measurable.

The ordinal counterpart to Property CM' is Property IND₂ below, which can be viewed as another "independence" condition: it says that if action a is implicitly judged to be at least as plausible as action b conditional on history h ($ha \lesssim hb$) then the same judgment must be made conditional on any other history that belongs to the same information set as h : if $h' \in I(h)$ and $a, b \in A(h)$, then

$$ha \lesssim hb \text{ if and only if } h'a \lesssim h'b. \quad (IND_2)$$

Note that Properties IND₁ and IND₂ are independent. An example of a plausibility order that violates IND₁ but satisfies IND₂ is plausibility order (1)

for the game of Figure 1: IND_1 is violated because $b < a$ but $bf \sim af$ and IND_2 is satisfied because at every non-singleton information set there are only two choices, one of which is plausibility preserving and the other is not. An example of a plausibility order that satisfies IND_1 but violates IND_2 is plausibility order (8) for the game of Figure 3.²² Adding Property IND_2 to the properties given in Definition 12 we obtain a refinement of the notion of weakly independent perfect Bayesian equilibrium.

Definition 14. *Consider an extensive-form game. An assessment (σ, μ) is a strongly independent perfect Bayesian equilibrium if it is sequentially rational, it is rationalized by a plausibility order that satisfies Properties IND_1 and IND_2 , and is uniformly Bayesian relative to that plausibility order.*

The following proposition states that the notions of weakly/strongly independent PBE identify two (nested) solution concepts that lie strictly in the gap between PBE and sequential equilibrium. The proof of the first part of Proposition 15 is given in the Appendix.

Proposition 15. *Consider an extensive-form game and an assessment (σ, μ) . If (σ, μ) is a sequential equilibrium then it is a strongly independent PBE. Furthermore, there are games where there is a strongly independent PBE which is not a sequential equilibrium.*

To see that the notion of strongly independent PBE is weaker than the notion of sequential equilibrium, consider the game of Figure 2 and the assessment (σ, μ) where $\sigma = (M, \ell, a, c, e)$ (highlighted in Figure 2 by doubling the corresponding edges) and $\mu(x) = 1$ for $x \in \{M, Mr, Lm, R\ell\}$ (and $\mu(x) = 0$ for every other decision history $x \neq \emptyset$; the decision histories x such that $\mu(x) > 0$ are shown as black nodes and the decision histories x such that $\mu(x) = 0$ are shown as gray nodes). This assessment is rationalized by plausibility order (7). It was shown above that (σ, μ) is a weakly independent PBE. To show that it is in fact a strongly independent PBE we only need to check that plausibility order (7) satisfies Property IND_2 . This is easily verified.²³ Next we show that (σ, μ) is not a sequential equilibrium, by appealing to Proposition 8 and showing that any plausibility order that rationalizes (σ, μ) is not choice measurable.²⁴ Let \preceq

²² IND_1 is satisfied because (1) $b < a$, $be < ae$ and $bf < af$, (2) $ae < af$, $aeg < afg$ and $aek < afk$, (3) $bf < be$, $bfl < bel$ and $bfm < bem$. IND_2 is violated because $b \in I(a)$ and $bf < be$ but $ae < af$.

²³The important observation is that $Mm < Mr$ and $Lm < Lr$ and $Rm < Rr$. The other comparisons involve a plausibility-preserving action *versus* a non-plausibility-preserving action and thus IND_2 is trivially satisfied.

²⁴To prove that (σ, μ) is not a sequential equilibrium it not sufficient to recall that plausibility order (7) was shown above to be not choice measurable, because there could be *another* plausibility order which is choice measurable and rationalizes (σ, μ) .

be a plausibility order that rationalizes (σ, μ) ; then it must satisfy the following properties:

- $Lm < Rr$ (because they belong to the same information set and $\mu(Lm) > 0$ while $\mu(Rr) = 0$). Thus if F is any integer-valued representation of \lesssim it must be that

$$F(Lm) < F(Rr). \quad (9)$$

- $Mr < L\ell \sim L$ (because Mr and $L\ell$ belong to the same information set and $\mu(Mr) > 0$ while $\mu(L\ell) = 0$; furthermore, ℓ is a plausibility-preserving action since $\sigma(\ell) > 0$). Thus if F is any integer-valued representation of \lesssim it must be that

$$F(Mr) < F(L). \quad (10)$$

- $R \sim R\ell < Mm$ (because ℓ is a plausibility-preserving action, $R\ell$ and Mm belong to the same information set and $\mu(R\ell) > 0$ while $\mu(Mm) = 0$). Thus if F is any integer-valued representation of \lesssim it must be that

$$F(R) < F(Mm). \quad (11)$$

Suppose that \lesssim is choice measurable and let F be an integer-valued representation of it that satisfies Property CM. From (9) and (10) we get that

$$F(Lm) - F(L) < F(Rr) - F(Mr) \quad (12)$$

and by Property CM it must be that

$$F(Rr) - F(Mr) = F(R) - F(M). \quad (13)$$

It follows from (12) and (13) that

$$F(Lm) - F(L) < F(R) - F(M). \quad (14)$$

Subtracting $F(M)$ from both sides of (11) we obtain

$$F(R) - F(M) < F(Mm) - F(M). \quad (15)$$

It follows from (14) and (15) that $F(Lm) - F(L) < F(Mm) - F(M)$, which can be written as $F(M) - F(L) < F(Mm) - F(Lm)$, yielding a contradiction, because Property CM requires that $F(M) - F(L) = F(Mm) - F(Lm)$.

Are the notions of weakly/strongly independent PBE “better” or “more natural” than the basic notion of PBE? This issue will be discussed briefly in Section 6.

5 Related literature

Since its introduction in 1982 ((15)), sequential equilibrium has been the most widely used solution concept for extensive-form games. However, in applications, checking the KW-consistency requirement has proved to be rather complex; thus, similarly motivated – but simpler – notions of equilibrium have been sought. The simplest solution concept is “weak sequential equilibrium” ((16, 17)) which is defined as an assessment that is sequentially rational and satisfies with Bayes’ rule at information sets that are reached with positive probability by the strategy profile (while no restrictions are imposed on the beliefs at information sets that have zero probability of being reached). However, this solution concept is too weak in that it is possible for an assessment (σ, μ) to be a weak sequential equilibrium without σ being a subgame-perfect equilibrium. Hence the search in the literature for a “simple” (yet sufficiently strong) solution concept that lies in the gap between subgame-perfect equilibrium and sequential equilibrium. The minimal desired properties of such a solution concepts, which is usually referred to as “perfect Bayesian equilibrium”, are sequential rationality and the “persistent” application of Bayes’ rule. The exact meaning of the latter requirement has not been easy to formalize. In ((11)) a notion of perfect Bayesian equilibrium was provided for a small subset of extensive-form games (namely the class of multi-stage games with observed actions and independent types), but extending that notion to arbitrary games proved to be problematic.²⁵ In (2) a notion of perfect Bayesian equilibrium is provided that can be applied to general extensive-form games (although it was defined only for games without chance moves); however, the proposed definition is in terms of a more complex object, namely a “tree-extended assessment” (ν, σ, μ) where ν is a conditional probability system on the set of terminal nodes. The main idea underlying the notion of perfect Bayesian equilibrium proposed in (2) is what the author calls “strategic independence”: when forming beliefs, the strategic choices of different players should be regarded as independent events. Several recent contributions ((5, 12, 23)) have re-addressed the issue of providing a definition of perfect Bayesian equilibrium that applies to general extensive-form games. Since (5) was reviewed in Section 2, here we shall briefly discuss (12, 23). In (12) the notion of “simple perfect Bayesian equilibrium” is introduced and it is shown to lie strictly between subgame-perfect equilibrium and sequential

²⁵The main element of the notion of PBE put forward in (11) is the “no signaling what you don’t know” condition on beliefs. For example, if Player 2 observes Player 1’s action and Player 1 has observed nothing about a particular move of Nature, then Player 2 should not update her beliefs about Nature’s choice based on Player 1’s action.

equilibrium. This notion is based on an extension of the definition of sub-tree, called “quasi sub-tree”, which consists of an information set I together with all the histories that are successors of histories in I (that is, Γ_I is a quasi-subtree that starts at I if $h' \in \Gamma_I$ if and only if there exists an $h \in I$ such that h is a prefix of h'). A quasi sub-tree Γ_I is called regular if it satisfies the following property: if $h \in \Gamma_I$ and $h' \in I(h)$ then $h' \in \Gamma_I$ (that is, every information set that has a non-empty intersection with Γ_I is entirely included in Γ_I). An information set I is called regular if the quasi-subtree that starts at I is regular. For example, in the game of Figure 4, the singleton information set $\{b\}$ of Player 2 is *not* regular. An assessment (σ, μ) is defined to be a “simple perfect Bayesian equilibrium” if it is sequentially rational and, for every regular quasi-subtree Γ_I , Bayes’ rule is satisfied at every information set that is reached with positive probability by σ in Γ_I (in other words, if the restriction of (σ, μ) to every regular quasi-subtree is a weak sequential equilibrium of the quasi-subtree). This notion of perfect Bayesian equilibrium is weaker than the notion proposed in (5). For example, in the game of Figure 4, the pure-strategy profile $\sigma = (c, d, f)$ (highlighted by double edges), together with the system of beliefs $\mu(a) = \mu(bd) = 0$, $\mu(be) = 1$, is a simple perfect Bayesian equilibrium, while (as shown in (5)) there is no system of beliefs μ' such that (σ, μ') is a perfect Bayesian equilibrium. *A fortiori*, the notion of simple perfect Bayesian equilibrium is weaker than the refinements of PBE discussed in the previous section.

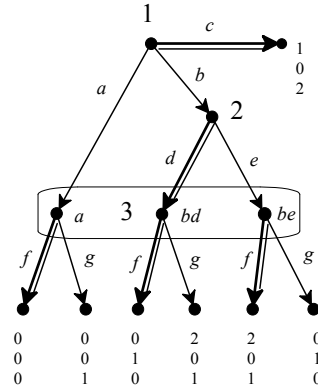


Figure 4:

(23) proposes a definition of perfect Bayesian equilibrium which is framed not in terms of assessments but in terms of “appraisals”. Each player is assumed to have a (possibly artificial) information set representing the beginning of the game and an appraisal for player i is a map that associates with every information set of player i a probability distribution over the set of pure-strategy profiles that reach that information set. Thus an appraisal for player i captures, for every information set of hers, her conjecture about how the information set was reached and what will happen from this point in the game. An appraisal system is defined to be “plainly consistent” if, whenever an information set of player i has a product structure (each information set is identified with the set of pure-strategy profiles that reach that information set), the player’s appraisal at that information set satisfies independence.²⁶ A strategy profile σ is defined to be a perfect Bayesian equilibrium if there is a plainly consistent appraisal system P that satisfies sequential rationality and is such that, at their “initial” information sets all the players assign probability 1 to σ ; (23, p.15) summarizes the notion of PBE as being based on “a simple foundation: sequential rationality and preservation of independence and Bayes updating where applicable” (that is, on subsets of strategy profiles that have the appropriate product structure and independence property). Despite the fact that the notion of PBE suggested in (23) incorporates a notion of independence, it can be weaker than the notion of PBE discussed in Section 2 (Definition 4) and thus, *a fortiori*, weaker than the notion of weakly independent PBE (Definition 12, Section 4). This can be seen from the game of Figure 5, which essentially reproduces Figure 9 of (23, p.24). The strategy profile $\sigma = (b, d, e)$ (highlighted by double edges), together with any system of beliefs μ such that $\mu(ac) > 0$ cannot be a PBE according to Definition 12 (Section 2). In fact, since $\sigma(d) > 0$ while $\sigma(c) = 0$, any plausibility order that rationalizes (σ, μ) must be such that $a \sim ad < ac$, which implies that $\mu(ac) = 0$ (because ac cannot be most plausible in the set $\{ac, ad, bc\}$). On the other hand, σ can be a PBE according to the definition given in (23, p.15), since the information set of Player 3 does not have a product structure so that Player 3 is not able to separate the actions of Players 1 and 2. For example, consider the appraisal system P where, initially, all the players assign probability 1 to σ and, at his information set, Player 2 assigns probability 1 to the strategy profile (b, e) of Players 1 and 3 and, at her information set, Player 3 assigns probability $\frac{1}{3}$ to each of the strategy profiles (a, c) , (a, d) and (b, c) of Players 1 and 2. Then P is plainly consistent and sequentially rational, so that σ is a PBE as defined in

²⁶Intuitively, on consecutive information sets, a player does not change her beliefs about the actions of other players, if she has not received information about those actions.

(23).

Thus, *a fortiori*, the notion of perfect Bayesian equilibrium given in (23) can be weaker than the notions of weakly/strongly independent PBE introduced in Section 4.

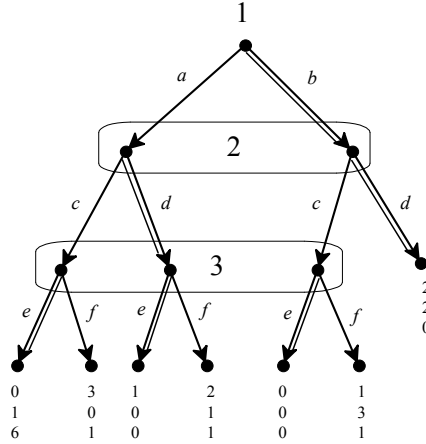


Figure 5:

6 Conclusion

Besides sequential rationality, the notion of perfect Bayesian equilibrium (Definition 4) introduced in (5) is based on two elements: (1) rationalizability of the assessment by a plausibility order (Definition 2) and (2) the notion of Bayesian consistency relative to the plausibility order. The first property identifies the set of decision histories that can be assigned positive conditional probability by the system of beliefs, while the second property imposes constraints on how conditional probabilities can be distributed over that set in order to guarantee “Bayesian updating as long as possible”.²⁷ In (6) it was shown that by

²⁷By “Bayesian updating as long as possible” we mean the following: (1) when information causes no surprises, because the play of the game is consistent with the most plausible play(s) (that

strengthening these two conditions one obtains a “limit free” characterization of sequential equilibrium. The strengthening of the first condition is that the plausibility order that rationalizes the given assessment be choice measurable, that it, that there be a cardinal representation of it (which can be interpreted as measuring the plausibility distance between histories in a way that is preserved by the addition of a common action). The strengthening of the second condition imposes “uniform consistency” on the conditional probability density functions on the equivalence classes of the plausibility order, by requiring that there be a full-support common prior that preserves the relative probabilities of two decision histories in the same information set when a common action is added. There is a “substantial” gap between the notion of PBE and that of sequential equilibrium. In this paper we identified two solution concepts that lie in this gap. The first notion, weakly independent PBE (Definition 12), is obtained by adding a restriction on the belief revision policy encoded in the plausibility order that rationalizes the given assessment (together with strengthening Bayes consistency to uniform Bayes consistency). This restriction says that observation of a new action at an information set should not alter the relative plausibility of any two histories in that information set (condition IND_1); it can be interpreted as an “independence” or “conservative” principle, in the sense that observation of a new action should not lead to a reversal of judgment of plausibility concerning past histories. The second notion, strongly independent PBE (Definition 14), is obtained by adding to the first notion a further restriction, according to which the implicit plausibility ranking of two actions available at the same information set should be independent of the history at which the actions are taken.

A further contribution of this paper has been to provide a method for determining if a plausibility order is choice measurable, which is one of the two conditions which are necessary and sufficient for a PBE to be a sequential equilibrium.

This paper highlights the need to conceptualize refinements of subgame-perfect equilibrium in extensive form games in terms of principles of belief revision. Through the notion of plausibility order and AGM-consistency we have appealed to the principles of belief revision that underlie the so-called AGM theory ((1)). However, in a dynamic game, beliefs typically need to

is, when information sets are reached that have positive prior probability), then beliefs should be updated using Bayes’ rule and (2) when information is surprising (that is, when an information set is reached that had zero prior probability) then new beliefs can be formed in an arbitrary way, but from then on Bayes’ rule should be used to update those new beliefs, whenever further information is received that is consistent with those beliefs.

be revised several times in a sequence as new information sets are reached. Thus the relevant notion of belief revision is *iterated belief revision*. There is an extensive literature on the topic of iterated belief revision (see, for example, (4, 7, 8, 22) and the special issue of the *Journal of Philosophical Logic*, Vol. 40 (2), April 2011). An exploration of different solution concepts in the gap between PBE and sequential equilibrium, based on different principles of iterated belief revision, seems to be a promising area of research.

A Proofs

Proof of Proposition 11. $(A) \Rightarrow (B)$. Let $F' : S \rightarrow \mathbb{N}$ satisfy the properties of Part (A). Select an arbitrary $s_0 \in S_0 = \{s \in S : s \preceq t, \forall t \in S\}$ and define $F : S \rightarrow \mathbb{N}$ by $F(s) = F'(s) - F'(s_0)$. Then F is also a function that satisfies the properties of Part (A) (note that since, for all $s \in S$, $F'(s_0) \leq F'(s)$, $F(s) \in \mathbb{N}$; furthermore, $F(s') = 0$ for all $s' \in S_0$). Let $K = \{k \in \mathbb{N} : k = \rho(s) \text{ for some } s \in S\}$ (where ρ is the canonical ordinal representation of \preceq : see Footnote 14). For every $k \in K$, define

$$\begin{aligned} \hat{x}_0 &= 0 \\ \text{and, for } k > 0, \\ \hat{x}_k &= F(t) - F(s) \quad \text{for some } s, t \in S \text{ such that } \rho(t) = k \text{ and } \rho(s) = k - 1. \end{aligned} \tag{16}$$

Thus \hat{x}_k is the distance, as measured by F , between the equivalence class of some t such that $\rho(t) = k$ and the immediately preceding equivalence class (that is, the equivalence class of some s such that $\rho(s) = k - 1$).²⁸ Note that \hat{x}_k is well defined since, if $x, y \in S$ are such that $\rho(y) = k$ and $\rho(x) = k - 1$, then $x \sim s$ and $y \sim t$ and thus, by (1) of Property (A), $F(x) = F(s)$ and $F(y) = F(t)$. Note also that, for all $k \in K \setminus \{0\}$, \hat{x}_k is a positive integer, since $\rho(t) = k$ and $\rho(s) = k - 1$ imply that $s < t$ and thus, by (1) of Property (A), $F(s) < F(t)$. We want to show that the values $\{\hat{x}_k\}_{k \in K \setminus \{0\}}$ defined in (16) provide a solution to the system of equations corresponding to \doteq (Definition 10). Select an arbitrary element of \doteq , $([s_1], [s_2]) \doteq ([t_1], [t_2])$ (with $s_1 < s_2$ and $t_1 < t_2$) and express it, using the canonical ordinal representation ρ (see Footnote 14), as $(i_1, i_2) \doteq (j_1, j_2)$ (thus $i_1 = \rho(s_1)$, $i_2 = \rho(s_2)$, $j_1 = \rho(t_1)$, $j_2 = \rho(t_2)$, $i_1 < i_2$ and $j_1 < j_2$). Then the corresponding equation (see Definition 10) is: $x_{i_1+1} + x_{i_1+2} + \dots + x_{i_2} = x_{j_1+1} + x_{j_1+2} + \dots + x_{j_2}$. By (2) of Property (A),

²⁸For example, if $S = \{a, b, c, d, e, f\}$ and \preceq is given by $a \sim b < c < d \sim e < f$ then $\rho(a) = \rho(b) = 0$, $\rho(c) = 1$, $\rho(d) = \rho(e) = 2$ and $\rho(f) = 3$; if F is given by $F(a) = F(b) = 0$, $F(c) = 3$, $F(d) = F(e) = 5$ and $F(f) = 9$ then $\hat{x}_0 = 0$, $\hat{x}_1 = 3$, $\hat{x}_2 = 2$ and $\hat{x}_3 = 4$.

$$F(s_2) - F(s_1) = F(t_2) - F(t_1) \quad (17)$$

Using (16), $F(s_2) - F(s_1) = \hat{x}_{i_1+1} + \hat{x}_{i_1+2} + \dots + \hat{x}_{i_2}$. To see this, for every $k \in \{i_1 + 1, i_1 + 2, \dots, i_2 - 1\}$, select an arbitrary $r_k \in S$ such that $\rho(r_k) = k$; then, by (16),

$$F(s_2) - F(s_1) = \underbrace{\hat{x}_{i_1+1}}_{=F(r_{i_1+1})-F(s_1)} + \underbrace{\hat{x}_{i_2+2}}_{=F(r_{i_1+2})-F(r_{i_1+1})} + \dots + \underbrace{\hat{x}_{i_2}}_{=F(s_2)-F(r_{i_2-1})}.$$

Similarly, $F(t_2) - F(t_1) = \hat{x}_{j_1+1} + \hat{x}_{j_1+2} + \dots + \hat{x}_{j_2}$. Thus, by (17), $\hat{x}_{i_1+1} + \hat{x}_{i_1+2} + \dots + \hat{x}_{i_2} = \hat{x}_{j_1+1} + \hat{x}_{j_1+2} + \dots + \hat{x}_{j_2}$.

(B) \Rightarrow (A). Assume that the system of equations corresponding to \doteq has a solution consisting of positive integers $\hat{x}_1, \dots, \hat{x}_m$. Define $F : S \rightarrow \mathbb{N}$ as follows: if $\rho(s) = 0$ (equivalently, $s \in S_0$) then $F(s) = 0$ and if $\rho(s) = k > 0$ (equivalently, $s \in S_k$ for $k > 0$) then $F(s) = \hat{x}_1 + \hat{x}_2 + \dots + \hat{x}_k$ (where ρ and the sets S_k are as defined in Footnote 14). We need to show that F satisfies the properties of Part (A). Select arbitrary $s, t \in S$ with $s \preceq t$. Then $\rho(s) \leq \rho(t)$ and thus $F(s) = \hat{x}_1 + \hat{x}_2 + \dots + \hat{x}_{\rho(s)} \leq F(t) = \hat{x}_1 + \hat{x}_2 + \dots + \hat{x}_{\rho(s)} + \hat{x}_{\rho(s)+1} + \dots + \hat{x}_{\rho(t)}$. Conversely, suppose that $s, t \in S$ are such that $F(s) \leq F(t)$. Then $\hat{x}_1 + \hat{x}_2 + \dots + \hat{x}_{\rho(s)} \leq \hat{x}_1 + \hat{x}_2 + \dots + \hat{x}_{\rho(t)}$ and thus $\rho(s) \leq \rho(t)$, so that $s \preceq t$. Thus Property (1) of Part (A) is satisfied. Now let $s, t, x, y \in S$ be such that $s < t$, $x < y$ and $([s], [t]) \doteq ([x], [y])$. Let $\rho(s) = i$, $\rho(t) = j$, $\rho(x) = k$ and $\rho(y) = \ell$ (thus $i < j$ and $k < \ell$). Then, by (16), $F(t) - F(s) = \hat{x}_{i+1} + \hat{x}_{i+2} + \dots + \hat{x}_j$ and $F(y) - F(x) = \hat{x}_{k+1} + \hat{x}_{k+2} + \dots + \hat{x}_\ell$. Since $x_{i+1} + x_{i+2} + \dots + x_j = x_{k+1} + x_{k+2} + \dots + x_\ell$ is the equation corresponding to $([s], [t]) \doteq ([x], [y])$ (which - using ρ - can be expressed as $(i, j) \doteq (k, \ell)$), by our hypothesis $\hat{x}_{i+1} + \hat{x}_{i+2} + \dots + \hat{x}_j = \hat{x}_{k+1} + \hat{x}_{k+2} + \dots + \hat{x}_\ell$ and thus $F(t) - F(s) = F(y) - F(x)$, so that (2) of Property (A) is satisfied.

not (B) \Rightarrow *not* (C). Suppose that there is a sequence in \doteq (expressed in terms of the canonical representation ρ of \preceq) $\langle ((i_1, j_1) \doteq (k_1, \ell_1)), \dots, ((i_m, j_m) \doteq (k_m, \ell_m)) \rangle$ such that

$$B_{left} \sqsubset B_{right} \quad (18)$$

where $B_{left} = B_{(i_1, j_1)} \uplus \dots \uplus B_{(i_m, j_m)}$ and $B_{right} = B_{(k_1, \ell_1)} \uplus \dots \uplus B_{(k_m, \ell_m)}$. Let $E = \{E_1, \dots, E_m\}$ be the system of equations corresponding to the above sequence (for example, E_1 is the equation $x_{i_1+1} + x_{i_1+2} + \dots + x_{j_1} = x_{k_1+1} + x_{k_1+2} + \dots + x_{\ell_1}$). Let L be the sum of the left-hand-side and R be the sum of the right-hand-side of the equations E_1, \dots, E_m . Note that for every integer i , nx_i is a summand of L if and only if i

appears in B_{left} exactly n times and similarly nx_i is a summand of R if and only if i appears in B_{right} exactly n times. By (18), if nx_i is a summand of L then mx_i is a summand of R with $m \geq n$ and, furthermore, $L \neq R$. Thus there cannot be a positive solution of \mathbb{E} , because it would be incompatible with $L = R$. Since \mathbb{E} is a subset of the system of equations corresponding to \doteq , it follows that the latter cannot have a positive solution either.

It only remains to prove that $not(C) \Rightarrow not(B)$. We will return to this below after providing an additional result. ■

First some notation. Given two vectors $x, y \in \mathbb{R}^m$ we write (1) $x \leq y$ if $x_i \leq y_i$, for every $i = 1, \dots, m$, (2) $x < y$ if $x \leq y$ and $x \neq y$ and (3) $x \ll y$ if $x_i < y_i$, for every $i = 1, \dots, m$.

Lemma 16. *Let A be the $m \times n$ matrix such that the system of equations corresponding to \doteq (Definition 10) can be expressed as $Ax = 0$ (note that each entry of A is either -1 , 0 or 1 ; furthermore, by symmetry of \doteq , for each row a_i of A there is another row a_k such that $a_k = -a_i$).²⁹ If the system of equations $Ax = 0$ does not have a positive integer solution then there exist r rows of A , a_{i_1}, \dots, a_{i_r} with $1 < r \leq \frac{m}{2}$ and r positive integers $\alpha_1, \dots, \alpha_r \in \mathbb{N} \setminus \{0\}$ such that if B is the submatrix of A consisting of the r rows a_{i_1}, \dots, a_{i_r} (thus for every $k = 1, \dots, r$, $b_k = a_{i_k}$, where b_k is the k^{th} row of B) then $\sum_{k=1}^r \alpha_k b_k < 0$.*

Proof. By Stiemke's theorem³⁰ if the system of equations $Ax = 0$ does not have a positive integer solution then there exists a $y \in \mathbb{Z}^m$ (where \mathbb{Z} denotes the set of integers) such that $yA < 0$ (that is, $\sum_{i=1}^m y_i a_i < 0$). Let $K = \{k \in \mathbb{Z} : y_k \neq 0\}$. Let r be the cardinality of K ; then, without loss of generality, we can assume that $r \leq \frac{m}{2}$.³¹ Furthermore, again without loss of generality, we can assume that

²⁹For example, the system of equations (6) can be written as $Ax = 0$, where $x = (x_1, \dots, x_5)$ and

$$A = \begin{pmatrix} 1 & 1 & 0 & -1 & 0 \\ -1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 0 & 1 \\ 1 & 1 & 1 & -1 & -1 \\ -1 & -1 & -1 & 1 & 1 \end{pmatrix} \quad (19)$$

³⁰See, for example, (20, p. 216) or (10, Theorem 1.1, p. 65).

³¹Proof. Recall that for each row a_i of A there is a row a_k such that $a_i = -a_k$. If $y_i \neq 0$ and $y_k \neq 0$ for some i and k such that $a_i = -a_k$ then

for every $k \in K$, $y_k > 0$.³² Let B be the $r \times n$ submatrix of A consisting of those rows a_k of A such that $k \in K$ and for $i = 1, \dots, r$ let $\alpha = (\alpha_1, \dots, \alpha_r)$ be the vector corresponding to $(y_k)_{k \in K}$.³³ Then $\alpha B = \sum_{j=1}^r \alpha_j b_j = yA < 0$ and $\alpha_i \in \mathbb{N} \setminus \{0\}$ for all $i = 1, \dots, r$. ■

Completion of proof of Proposition 11. It remains to prove that *not* (C) \Rightarrow *not* (B). Let A be the $m \times n$ matrix such that the system of equations corresponding to \doteq can be expressed as $Ax = 0$ and assume that $Ax = 0$ does not have a positive integer solution. Let B be the $r \times n$ submatrix of A and $\alpha = (\alpha_1, \dots, \alpha_r)$ the vector of positive integers of Lemma 16 such that $\alpha B = \sum_{j=1}^r \alpha_j b_j < 0$. Define two $r \times n$ matrices $C = (c_{ij})_{i=1, \dots, r; j=1, \dots, n}$ and $D = (d_{ij})_{i=1, \dots, r; j=1, \dots, n}$ as follows (recall that each entry of B is either $-1, 0$ or 1):

$$c_{ij} = \begin{cases} 1 & \text{if } b_{ij} = 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad d_{ij} = \begin{cases} 1 & \text{if } b_{ij} = -1 \\ 0 & \text{otherwise} \end{cases}.$$

Then, for every $i = 1, \dots, r$, $b_i = c_i - d_i$ and thus (since $\sum_{i=1}^r \alpha_i b_i < 0$)

$$\sum_{i=1}^r \alpha_i c_i < \sum_{i=1}^r \alpha_i d_i. \quad (20)$$

Let C' be the matrix obtained from C by replacing each row c_i of C with α_i copies of it and let D' be constructed from D similarly. Then, letting $s = \sum_{i=1}^r \alpha_i$, C' and D' are $s \times n$ matrices whose entries are either 0 or 1. It follows from (20) that

$$\sum_{i=1}^s c'_i < \sum_{i=1}^s d'_i. \quad (21)$$

$$y_i a_i + y_k a_k = \begin{cases} 0 & \text{if } y_i = y_k \\ (y_k - y_i) a_k & \text{if } 0 < y_i < y_k \\ (y_i - y_k) a_i & \text{if } 0 < y_k < y_i \\ (|y_i| + y_k) a_k & \text{if } y_i < 0 < y_k \\ (y_i + |y_k|) a_i & \text{if } y_k < 0 < y_i \\ (|y_k| - |y_i|) a_i & \text{if } y_i < y_k < 0 \\ (|y_i| - |y_k|) a_k & \text{if } y_k < y_i < 0 \end{cases}$$

where all the multipliers (of a_i or a_k) are positive. Thus one can set one of the two values of y_i and y_k to zero and replace the other value with the relevant of the above values while keeping yA unchanged. For example, if $y_k < y_i < 0$ then one can replace y_i with 0 and y_k with $(|y_i| - |y_k|)$ thereby reducing the cardinality of K by one. This process can be repeated until the multipliers of half of the rows of A have been replaced by zero.

³²Proof. Suppose that $y_k < 0$ for some $k \in K$. Recall that there exists an i such that $a_k = -a_i$. By the argument of the previous footnote, $y_i = 0$. Then replace y_k by 0 and replace $y_i = 0$ by $\tilde{y}_i = -y_k$.

³³For example, if $K = \{3, 6, 7\}$ and $y_3 = 2$, $y_6 = 1$, $y_7 = 3$, then B is the $3 \times n$ matrix where $b_1 = a_3$, $b_2 = a_6$ and $b_3 = a_7$ and $\alpha_1 = 2$, $\alpha_2 = 1$ and $\alpha_3 = 3$.

Consider the system of equations

$$C'x = D'x. \quad (22)$$

For every $j = 1, \dots, n$, the j^{th} coordinate of $\sum_{i=1}^s c'_i$ is the number of times that the variable x_j appears on the left-hand-side of (22) and the j^{th} coordinate of $\sum_{i=1}^s d'_i$ is the number of times that the variable x_j appears on the right-hand-side of (22). Hence, by (21), for every $j = 1, \dots, n$, the number of times that the variable x_j appears on the left-hand-side of (22) is less than or equal to the number of times that it appears on the right-hand-side of (22) and for at least one j it is less. Thus, letting $\langle ((i_1, j_1) \doteq (k_1, \ell_1)), \dots, ((i_s, j_s) \doteq (k_s, \ell_s)) \rangle$ be the sequence of elements of \doteq corresponding to the equations in (22), we have that $B_{\text{left}} \sqsubset B_{\text{right}}$ where $B_{\text{left}} = B_{(i_1, j_1)} \sqcup \dots \sqcup B_{(i_m, j_m)}$ and $B_{\text{right}} = B_{(k_1, \ell_1)} \sqcup \dots \sqcup B_{(k_m, \ell_m)}$. ■

Proof of Lemma 13. Let \preceq be a plausibility order on the set of histories H and let $F : H \rightarrow \mathbb{N}$ be an integer-valued representation of \preceq . We want to show that properties CM and CM' below are equivalent (for arbitrary $h, h' \in H$ and $a, b \in A(h)$, with $h' \in I(h)$)

$$F(h') - F(h) = F(h'a) - F(ha). \quad (\text{CM})$$

$$F(hb) - F(ha) = F(h'b) - F(h'a). \quad (\text{CM}')$$

First of all, note that, without loss of generality, we can assume that $F(\emptyset) = 0$.³⁴ First we show that $\text{CM} \Rightarrow \text{CM}'$. Let F be an integer-valued representation of \preceq that satisfies CM. For every decision history h and action $a \in A(h)$, define

$$\lambda(a) = F(ha) - F(h). \quad (23)$$

The function $\lambda : A \rightarrow \mathbb{N}$ is well defined, since, by CM, $h' \in I(h)$ implies that $F(h'a) - F(h') = F(ha) - F(h)$. Then, for every history $h = \langle a_1, a_2, \dots, a_m \rangle$, $F(h) = \sum_{i=1}^m \lambda(a_i)$. In fact,

$$\begin{aligned} \lambda(a_1) + \lambda(a_2) + \dots + \lambda(a_m) &= \\ &= (F(a_1) - F(\emptyset)) + (F(a_1a_2) - F(a_1)) + \dots + (F(a_1a_2\dots a_m) - F(a_1a_2\dots a_{m-1})) = \\ &= F(a_1a_2\dots a_m) = F(h) \text{ (recall the hypothesis that } F(\emptyset) = 0). \end{aligned}$$

Thus, for every $h \in D$ and $a \in A(h)$, $F(ha) = F(h) + \lambda(a)$. Hence, $F(hb) - F(ha) = F(h) + \lambda(b) - F(h) - \lambda(a) = \lambda(b) - \lambda(a)$ and $F(h'b) - F(h'a) = F(h') + \lambda(b) - F(h') - \lambda(a) = \lambda(b) - \lambda(a)$ and, therefore, $F(hb) - F(ha) = F(h'b) - F(h'a)$.

³⁴It is straightforward to check that if $F' : H \rightarrow \mathbb{N}$ is an integer-valued representation of \preceq then so is $F : H \rightarrow \mathbb{N}$ defined by $F(h) = F'(h) - F'(\emptyset)$; furthermore if F' satisfies property CM (CM') then so does F .

Next we show that $CM' \Rightarrow CM$. Let \lesssim be a plausibility order and let $F : H \rightarrow \mathbb{N}$ be an integer-valued representation that satisfies CM' . Select arbitrary $h' \in I(h)$ and $a \in A(h)$. Let $b \in A(h)$ be a plausibility-preserving action at h (there must be at least one such action: see Definition 1); then, $h \sim hb$ and $h' \sim h'b$. Hence, since F is a representation of \lesssim , $F(hb) = F(h)$ and $F(h'b) = F(h')$ and thus

$$F(h') - F(h) = F(h'b) - F(hb). \quad (24)$$

By CM' , $F(h'b) - F(hb) = F(h'a) - F(ha)$. From this and (24) it follows that $F(h') - F(h) = F(h'a) - F(ha)$. ■

Proof of Proposition 15. Let (σ, μ) be a sequential equilibrium. We want to show that (σ, μ) is a strongly independent PBE (Definition 14). By Proposition 8, it is sufficient to show that (σ, μ) is rationalized by a plausibility order \lesssim that satisfies Properties IND_1 and IND_2 . By Proposition 8 there is a choice measurable plausibility order \lesssim that rationalizes (σ, μ) . Let F be an integer-valued representation of \lesssim that satisfies Property CM . Let h and h' be decision histories that belong to the same information set and let $a \in A(h)$. Then, by CM ,

$$F(h) - F(h') = F(ha) - F(h'a) \quad (25)$$

If $h \lesssim h'$ then $F(h) \leq F(h')$; by (25), it follows that $F(ha) \leq F(h'a)$ and thus $ha \lesssim h'a$. Conversely, if $ha \lesssim h'a$ then $F(ha) \leq F(h'a)$ and thus, by (25), $F(h) \leq F(h')$ so that $h \lesssim h'$. Hence \lesssim satisfies IND_1 .

Let h and h' be decision histories that belong to the same information set and let $a, b \in A(h)$. We want to show that IND_2 holds, that is, that $ha \lesssim hb$ if and only if $h'a \lesssim h'b$. Let F be an integer-valued representation of \lesssim that satisfies Property CM . By Lemma 13 F satisfies Property CM' , that is,

$$F(ha) - F(hb) = F(h'a) - F(h'b) \quad (26)$$

If $ha \lesssim hb$ then $F(ha) \leq F(hb)$ and thus, by (26), $F(h'a) \leq F(h'b)$, that is, $h'a \lesssim h'b$. Conversely, if $h'a \lesssim h'b$ then $F(h'a) \leq F(h'b)$ and thus, by (26), $F(ha) \leq F(hb)$, so that $ha \lesssim hb$. ■

References

- [1] Alchourrón, Carlos, Peter Gärdenfors and David Makinson, On the logic of theory change: partial meet contraction and revision functions, *The Journal of Symbolic Logic*, 1985, 50: 510-530.

- [2] Battigalli, Pierpaolo, Strategic independence and perfect Bayesian equilibria, *Journal of Economic Theory*, 1996, 70: 201-234.
 - [3] Bonanno, Giacomo, AGM belief revision in dynamic games, in: K. R. Apt (ed.), *Proceedings of the 13th conference on theoretical aspects of rationality and knowledge (TARK XIII)*, ACM, New York, 2011, 37-45.
 - [4] Bonanno, Giacomo, Belief change in branching time: AGM-consistency and iterated revision, *Journal of Philosophical Logic*, 2012, 41: 201-236.
 - [5] Bonanno, Giacomo, AGM-consistency and perfect Bayesian equilibrium. Part I: definition and properties, *International Journal of Game Theory*, 2013, 42: 567-592.
 - [6] Bonanno, Giacomo, AGM-consistency and perfect Bayesian equilibrium. Part II: from PBE to sequential equilibrium, *International Journal of Game Theory*, forthcoming.
 - [7] Boutilier, Craig, Iterated revision and minimal change of conditional beliefs. *Journal of Philosophical Logic*, 25: 262305.
 - [8] Darwiche, Adnan and Judea Pearl, On the logic of iterated belief revision, *Artificial Intelligence*, 1997, 89: 1-29.
 - [9] Dines, L., On positive solutions of a system of linear equations, *Annals of Mathematics*, 1926-1927, 28: 386-392.
 - [10] Fishburn, P. C., Finite linear qualitative probability, *Journal of Mathematical Psychology*, 1996, 40: 64-77.
 - [11] Fudenberg, Drew and Jean Tirole, Perfect Bayesian equilibrium and sequential equilibrium, *Journal of Economic Theory*, 1991, 53: 236-260.
 - [12] González-Díaz, J. and M.A. Meléndez-Jiménez, On the notion of perfect Bayesian equilibrium, *TOP Journal of the Spanish Society of Statistics and Operations Research*, 2014, 22:128-143.
 - [13] Hendon, Ebbe, Jörgen Jacobsen and Brigitte Sloth, The one-shot-deviation principle for sequential rationality, *Games and Economic Behavior*, 1996, 12: 274-282.
 - [14] Kohlberg, Elon and Philip Reny, Independence on relative probability spaces and consistent assessments in game trees, *Journal of Economic Theory*, 1997, 75: 280-313.
-

-
- [15] Kreps, David and Robert Wilson, Sequential equilibrium, *Econometrica*, 1982, 50: 863-894.
 - [16] Mas-Colell A., Whinston M.D., Green JR (1995) *Microeconomic theory*, Oxford University Press, 1995.
 - [17] Myerson, Roger, *Game theory: analysis of conflict* , Harvard University Press, 1991.
 - [18] Osborne, Martin and Ariel Rubinstein, *A course in game theory*, MIT Press, 1994.
 - [19] Perea, Andrés, A note on the one-deviation property in extensive form games, *Games and Economic Behavior*, 2002, 40: 322-338.
 - [20] Schrijver, A., *Theory of linear and integer programming*, John Wiley & Sons, 1986.
 - [21] Selten, Reinhard, Re-examination of the perfectness concept for equilibrium points in extensive games, *International Journal of Game Theory*, 1975, 4: 25-55.
 - [22] Stalnaker, Robert, Iterated belief revision, *Erkenntnis*, 2009, 70: 189-209.
 - [23] Watson, Joel, Perfect Bayesian equilibrium: general definitions and illustrations, working paper, University of California San Diego, 2016.
-