UC DAVIS
DEPARTMENT OF ECONOMICS
Working Paper Series

AGM-consistency and perfect Bayesian equilibrium. Part II: from PBE to sequential equilibrium.

Giacomo Bonanno
University of California, Davis

February 04, 2014

Paper # 14-1

In Bonanno (Int. J. Game Theory, 42:567-592, 2013) a general notion of perfect Bayesian equilibrium (PBE) was introduced for extensive-form games and shown to be intermediate between subgame-perfect equilibrium and sequential equilibrium. The essential ingredient of the proposed notion is the existence of a plausibility order on the set of histories that rationalizes a given assessment. In this paper we study restrictions on the belief revision policy encoded in a plausibility order and provide necessary and sufficient conditions for a PBE to be a sequential equilibrium.

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

http://www.econ.ucdavis.edu/working_search.cfm

AGM-consistency and perfect Bayesian equilibrium. Part II: from PBE to sequential equilibrium.

Giacomo Bonanno
Department of Economics,
University of California,
Davis, CA 95616-8578, USA
gfbonanno@ucdavis.edu

January 2014

Abstract

In [5] (*Int. J. Game Theory*, 42:567-592, 2013) a general notion of perfect Bayesian equilibrium (PBE) was introduced for extensive-form games and shown to be intermediate between subgame-perfect equilibrium and sequential equilibrium. The essential ingredient of the proposed notion is the existence of a plausibility order on the set of histories that rationalizes a given assessment. In this paper we study restrictions on the belief revision policy encoded in a plausibility order and provide necessary and sufficient conditions for a PBE to be a sequential equilibrium.
Keywords: plausibility order, belief revision, Bayesian updating, independence, sequential equilibrium, consistency.

1 Introduction

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 ([18]); it was also shown that, in turn, then notion of sequential equilibrium ([11]) is a strict refinement of PBE. The central elements of the definition of perfect Bayesian equilibrium are the qualitative notions of plausibility order and AGM-consistency. As shown in [4], these notions can be derived from the primitive concept of a player's epistemic state, which encodes the player's initial (or prior) 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 axioms for "rational" belief revision introduced in [1]. In this paper we continue the study of PBE, in particular we explore the restrictions on belief revision incorporated in the notion of sequential equilibrium and provide necessary and sufficient conditions for a PBE to be a sequential equilibrium. There are two such conditions. One is the qualitative notion of choice-measurable plausibility order, which was shown in [5] to be implied by sequential equilibrium. Choice measurability requires that the plausibility order $\precsim$ on the set of histories that rationalizes the given assessment have a *cardinal* representation, in the sense that there is an integer-valued representation $F : H \to \mathbb{N}$ of $\precsim$ such that $F(h) \leq F(h')$ if and only if $h \precsim h'$ (this is the ordinal part) and, if a is an action available at h and h' (so that h and h' belong tot he same information set) then $F(h) - F(h') = F(ha) - F(h'a)$ (this is the cardinal part). Thus a cardinal representation F measures the distance between histories and this distance is required to be preserved by the addition of a common action. The notion of choice measurability corresponds to the notion of labelling introduced in [11, p. 887] and further clarified in [19, 20]. The second condition is a strengthening of the notion of "consistent Bayesian updating" which is part of the definition of PBE.

The paper is organized as follows. The next section reviews the definition of PBE. Section 3 discusses choice measurability and provides two qualitative characterizations of it, one of which is known from the theory of qualitative probabilities and makes essential use of the structure of histories, and the other is, as far as I know, new and is more general, in the sense that it applies to pre-orders on arbitrary sets. In Section 4 it is shown that choice measurability together with a strengthening of the notion of "consistent Bayesian updating" are necessary and sufficient for a PBE to be a sequential equilibrium. Section 5 discusses related literature and Section 6 concludes. The proofs are given in Appendix B, while Appendix A reviews the notation of [5], which is used throughout this paper.

2 AGM-consistency and perfect Bayesian equilibrium

We use the history-based definition of extensive-form game (see Appendix A). Recall that an *assessment* is a pair (σ, μ) where σ is a behavior strategy profile and μ is a system of beliefs (that is, a collection of probability distributions, one for every information set, over the elements of that information set). An assessment encodes the beliefs and belief revision policy of the players: the strategy profile σ yields the initial beliefs as well as conditional beliefs about future moves, while μ gives conditional beliefs about past moves. The following definitions are taken from [5].

Given a set H, a total pre-order on H is a binary relation on H which is complete ($\forall h, h' \in H$, either $h \precsim h'$ or $h' \precsim h$) and transitive ($\forall h, h', h'' \in H$, if $h \precsim h'$ and $h' \precsim h''$ then $h \precsim h''$).

Definition 1 *Given an extensive form, a* plausibility order *is a total pre-order on the set of histories H that satisfies the following properties: for every $h \in D$ (D is the set of decision histories),*

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

If $h \precsim h'$ we say that history h is *at least as plausible* as history h'.[1] Property $PL1$ says that adding an action to a decision history h cannot yield a more plausible history than h itself. Property

[1] As in [5] we use the notation $h \precsim h'$ for "h is *at least as plausible* as h'", rather than the perhaps more natural notation $h \succsim h'$, for two reasons: (1) it is the standard notation in the extensive literature that deals with AGM-style ([1]) belief revision and (2) as shown below, it is convenient to assign lower values to more plausible histories (and the value 0 to the most plausible histories). In light of this, an alternative reading of $h \precsim h'$ is "h is *not more implausible* than h'", in which case one would think of $\precsim$ as an "implausibility" order.

$PL2$ says that at every decision history h there is some action a such that adding a to h yields a history which is at least as plausible as h and, furthermore, any such action a performs the same role with any other history that belongs to the same information set. Property $PL3$ says that all the actions at a history assigned to chance are "plausibility preserving". We write $h \sim h'$ (with the interpretation that h is *as plausible as* h') as a short-hand for "$h \precsim h'$ and $h' \precsim h$" (thus $\sim$ is an equivalence relation on H) and we write $h \prec h'$ (with the interpretation that h is *more plausible* than h') as a short-hand for "$h \precsim h'$ and not $h' \precsim h$". It follows from Property $PL1$ that, for every $h, h' \in H$, if h' is a prefix of h then $h' \precsim h$. Furthermore, by Properties $PL1$ and $PL2$, for every decision history h, there is at least one action a at h such that $h \sim ha$, that is, ha is as plausible as h and, furthermore, if h' belongs to the same information set as h, then $h' \sim h'a$. We call such actions *plausibility preserving.*

A plausibility order encodes a qualitative description of the *epistemic state* of each player, that is, the player's *initial beliefs* (or *prior beliefs*: her beliefs before the game is played) and her *disposition to revise those beliefs* when informed that it is her turn to move.

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

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

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

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

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

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

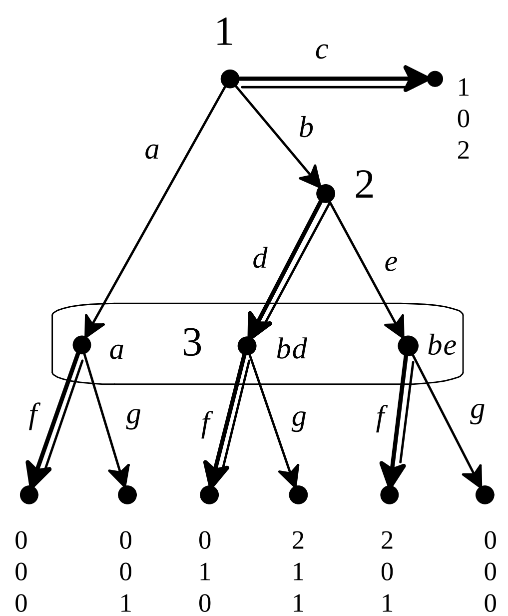
An assessment (σ, μ) is sequentially rational if, for every player i and every information set I of hers, player i's expected payoff - given the strategy profile σ and her beliefs at I - cannot be increased by unilaterally changing her choice at I and possibly at information sets of hers that follow I.[2] In order to define sequential rationality more precisely we need additional notation. Let Z denote the set of terminal histories and, for every player i, let $U_i : Z \to \mathbb{R}$ 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 $z = ha_1...a_m$ 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 payoff 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

[2] 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 6 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 on μ identified in [8] and [15] as being necessary and sufficient for the equivalence of the two notions. For simplicity we have chosen the weakly local definition.

every strategy τ_i of player i (where σ_{-i} denotes the strategy profile of the players other than i, that is, $\sigma_{-i} = (\sigma_1, ..., \sigma_{i-1}, \sigma_{i+1}, ..., \sigma_n)$).

Definition 3 *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*.*

In conjunction with sequential rationality, the notion of AGM-consistency is sufficient to eliminate some Nash equilibria as "implausible". Consider, for example, the extensive game of Figure 1 and the pure-strategy profile $\sigma = (c, d, f)$ (highlighted by double edges), which constitutes a Nash equilibrium of the game (and also a subgame-perfect equilibrium since there are no proper subgames). Can σ be part of a sequentially rational AGM-consistent assessment (σ, μ)? Since, for Player 3, choice f can be rational only if the player assigns (sufficiently high) positive probability to history be, sequential rationality requires that $\mu(be) > 0$; however, any such assessment is *not* AGM-consistent. In fact, if there were a plausibility order $\precsim$ that satisfied Definition 2, then, by $P1$, $b \sim bd$ (since $\sigma(d) = 1 > 0$) and $b \prec be$ (since $\sigma(e) = 0$)[3] and, by $P2$, $be \precsim bd$ (since - by hypothesis - μ assigns positive probability to be). By transitivity of $\precsim$, from $b \sim bd$ and $b \prec be$ it follows that $bd \prec be$, yielding a contradiction.



The Nash equilibrium $\sigma = (c, d, f)$ cannot be part
of a sequentially rational AGM-consistent assessment.
Figure 1

On the other hand, the Nash equilibrium $\sigma' = (b, d, g)$ together with $\mu'(bd) = 1$ forms a sequentially rational, AGM-consistent assessment: it can be rationalized by several plausibility orders, for

[3]By definition of plausibility order, $b \precsim be$ and, by *P1,* it is not the case that $b \sim be$ because e is not assigned positive probability by σ. Thus $b \prec be$.

instance $\begin{pmatrix} \emptyset, b, bd, bdg & \text{most plausible} \\ a, c, be, ag, beg & \\ af, bdf, bef & \text{least plausible} \end{pmatrix}$, where each row represents an equivalence class of $\precsim$.

Throughout this paper we adopt the following representation of a total pre-order: if the row to which history h belongs is above the row to which h' belongs, then $h \prec h'$ (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').

Definition 4 *Fix an extensive form. Let $\precsim$ be a plausibility order that rationalizes the assessment (σ, μ). We say that (σ, μ) is* Bayesian relative to $\precsim$ *if for every $\precsim$-equivalence class E that contains some decision history h with $\mu(h) > 0$, there exists a probability measure $\nu_E : H \to [0, 1]$ such that:*

$B1$. $Supp(\nu_E) = \{h \in E : \mu(h) > 0\}$.

$B2$. If $h, h' \in Supp(\nu_E)$ and $h' = ha_1...a_m$ (that is, h is a prefix of h') then

$$\nu_E(h') = \nu_E(h) \times \sigma(a_1) \times \ ... \ \times \ \sigma(a_m).$$

$B3$. If $h' \in Supp(\nu_E)$, then, $\forall h \in I(h')$

$$\mu(h) = \nu_E (h \mid I(h')) = \frac{\nu_E(h)}{\nu_E(I(h'))}.$$

Property $B1$ requires that $\nu_E(h) > 0$ if and only if $h \in E$ and $\mu(h) > 0$ (thus $\nu_E(h) = 0$ if and only if either $h \in H \backslash E$ or $\mu(h) = 0$). Property $B2$ requires ν_E to be consistent with the strategy profile σ in the sense that if $h, h' \in E$, $\nu_E(h) > 0$, $\nu_E(h') > 0$ and $h' = ha_1...a_m$ then the probability of h' (according to ν_E) is equal to the probability of h multiplied by the probabilities (according to σ) of the actions that lead from h to h'.[4] Property $B3$ requires the system of beliefs μ to satisfy Bayes' rule in the sense that if history h belongs to information set I then $\mu(h)$ (the probability assigned to h by μ) is the probability of h conditional on I using the probability measure ν_E, where E is the equivalence class of the most plausible elements of I. In fact, as shown in [5], Property $B3$ is equivalent to the following (for every $I \subseteq H$, let $Min_{\precsim}\ I$ denote the set of most plausible histories in I, that is, $Min_{\precsim}\ I = \{h \in I : h \precsim h', \forall h' \in I\}$):

$$\begin{array}{c} \text{For every information set } I \text{ such that } Min_{\precsim} I \subseteq E \\ \text{and, for every } h \in I,\ \ \mu(h) = \nu_E (h \mid I) = \frac{\nu_E(h)}{\nu_E(I)}. \end{array}$$

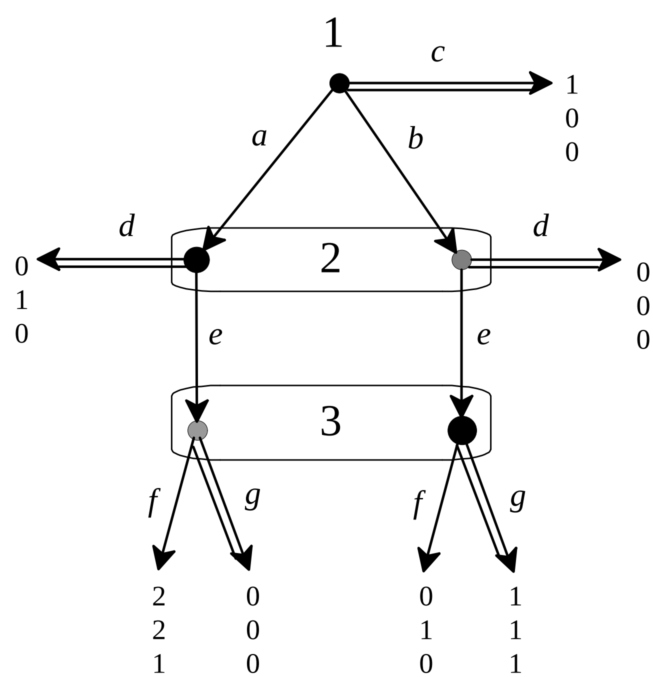
Definition 5 *An assessment (σ, μ) is* Bayesian AGM-consistent *if it is rationalized by a plausibility order $\precsim$ on the set of histories H and it is Bayesian relative to $\precsim$.*

The last ingredient of the definition of perfect Bayesian equilibrium is the standard requirement of sequential rationality.

Definition 6 *An assessment (σ, μ) is a* perfect Bayesian equilibrium *if it is Bayesian AGM-consistent and sequentially rational.*

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

Remark 7 *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 (but the converse is not true). The example of Figure 1 shows that PBE is a strict refinement of subgame-perfect equilibrium (since in that game every Nash equilibrium is subgame-perfect because there are no proper subgames).*



The assessment $\sigma = (c, d, g)$, $\mu(a) = \mu(be) = 1$
is a perfect Bayesian equilibrium.
Figure 2

For the game illustrated in Figure 2, a perfect Bayesian equilibrium is given by $\sigma = (c, d, g)$ and $\mu(a) = \mu(be) = 1$ (σ is highlighted by double edges and the histories that are assigned positive probability by μ are shown as black nodes). In fact (σ, μ) is sequentially rational and, furthermore, it is rationalized by the following plausibility order and is Bayesian relative to it (as can be verified using the following measures on the equivalence classes of the plausibility order that contain histories h with $\mu(h) > 0$):

$$\begin{pmatrix} \emptyset, c \\ a, ad \\ b, bd \\ be, beg \\ ae, aeg \\ bef \\ aef \end{pmatrix}, \qquad \begin{pmatrix} \nu_0(\emptyset) = 1 \\ \nu_1(a) = 1 \\ \nu_3(be) = 1 \end{pmatrix} \tag{1}$$

As explained below, the belief revision policy encoded in the PBE shown in Figure 2 may be criticized as failing to be "minimal". In the following sections we introduce restrictions on belief revision that can be used to refine the notion of perfect Bayesian equilibrium and provide a characterization of sequential equilibrium which is essentially qualitative.

3 Choice-measurable plausibility orders

The belief revision policy encoded in a perfect Bayesian equilibrium can be interpreted either as the epistemic state of an external observer[5] or as a belief revision policy which is shared by all the players. For example, the perfect Bayesian equilibrium illustrated in Figure 2 reflects the following belief revision policy: the initial beliefs are that Player 1 will play c; conditional on learning that Player 1 did not play c, the observer would become convinced that Player 1 played a (that is, she would judge a to be strictly more plausible than b) and would expect Player 2 to play d; upon learning that (Player 1 did not play c and) Player 2 did not play d, the observer would become convinced that Player 1 played b and Player 2 played e, hence judging be to be strictly more plausible than ae, thereby reversing her earlier belief that a was strictly more plausible than b. Although such a belief revision policy does not violate the rationality constraints introduced in [1], it does involve a belief change which is not "minimal". 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 ($h' \in I(h)$) and a is an action available at h ($a \in A(h)$), then

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

IND_1 says that if h is deemed to be at least as plausible as h' then the addition of any 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 actions. Any plausibility order that rationalizes the assessment shown in Figure 2 must violate IND_1.

Another "minimality" or "independence" condition is the following, which says that if action a is implicitly judged to be at least as plausible as action b conditional on history h 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

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

These two properties of plausibility orders are independent of each other: see Lemma 9 below.

Rather than investigating refinements of perfect Bayesian equilibrium induced by properties such as IND_1 and IND_2, we will focus on one property of plausibility orders that implies both IND_1 and IND_2 and is at the core of the notion of sequential equilibrium. The following definitions are taken from [5].

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

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

[5] For example, [9] adopt this interpretation. For a subjective interpretation of perfect Bayesian equilibrium and an epistemic characterization of it see [4].

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 8 *Let $H_0 = \{h \in H : h \precsim x, \quad \forall x \in H\}$, $H_1 = \{h \in H \setminus H_0 : \ h \precsim x, \quad \forall x \in H \backslash H_0\}$ and, in general, for every integer $k \geq 1$, $H_k = \{h \in H \setminus H_0 \cup ... \cup \ H_{k-1} : \ \ h \precsim x, \quad \forall x \in H \setminus H_0 \cup ... \cup \ H_{k-1}\}$. Thus H_0 is the equivalence class of $\precsim$ 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.*[6] *The* canonical ordinal integer-valued representation *of $\precsim$, $\rho : H \to \mathbb{N}$, is given by*

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

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

Using ρ, property IND_1 can be restated as follows: if $h' \in I(h)$ and $a \in A(h)$ then

$$\rho(h') - \rho(h) \geq 0 \text{ if and only if } \rho(h'a) - \rho(ha) \geq 0. \tag{4}$$

Instead of an ordinal representation of the plausibility order $\precsim$ one could seek a *cardinal* representation $F : H \to \mathbb{N}$ which, besides (2), satisfies the following property: if $h' \in I(h)$ and $a \in A(h)$, then

$$F(h') - F(h) = F(h'a) - F(ha). \tag{IND_1C}$$

IND_1C can be viewed as a cardinal version of IND_1 and it clearly implies the latter. If we think of such a function F as measuring the "plausibility distance" between histories, then we can interpret IND_1C 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. The cardinal counterpart to IND_2 is given by: if $h' \in I(h)$ and $a, b \in A(h)$, then

$$F(hb) - F(ha) = F(h'b) - F(h'a) \tag{IND_2C}$$

Clearly IND_2C implies IND_2.

The proof of the following lemma is given in Appendix B.

Lemma 9 *While the qualitative properties IND_1 and IND_2 are independent, IND_1C holds if and only if IND_2C holds.*

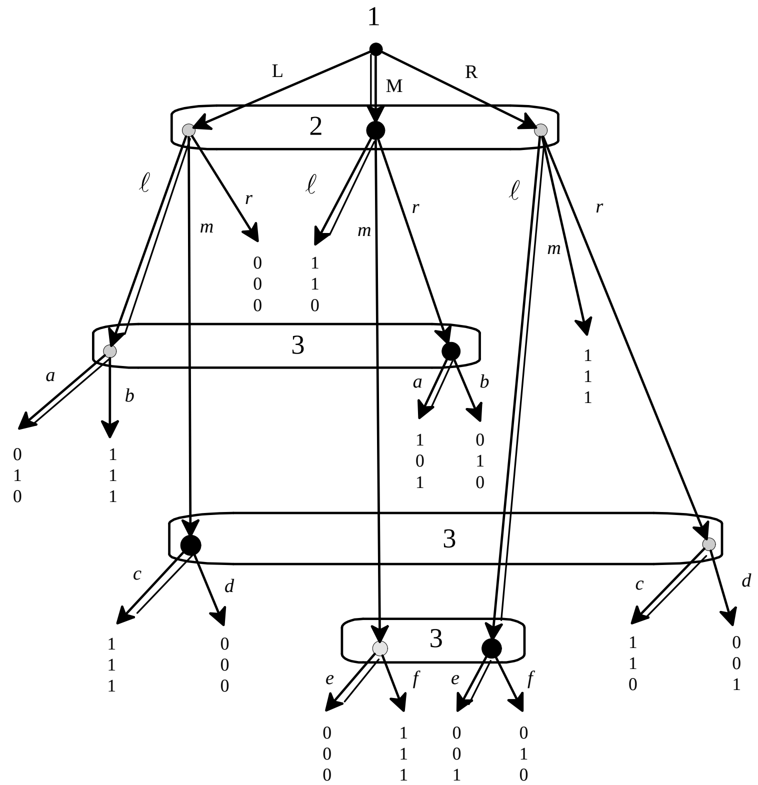
Definition 10 *A plausibility order $\precsim$ on the set of histories H is* choice-measurable *if there exists an $F : H \to \mathbb{N}$ that satisfies (1) $F(h) \leq F(h')$ if and only if $h \precsim h'$ and (2) property IND_1C (which, by the above lemma, is equivalent to IND_2C). We call such a function F a* cardinal integer-valued representation *of $\precsim$. Note that, without loss of generality, we can set $F(\emptyset) = 0$.*[7]

[6] Since H is finite, there is an $m \in \mathbb{N}$ such that $\{H_0, ..., H_m\}$ 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 $h \prec h'$.

[7] If $F(\emptyset) > 0$, define $\hat{F} : H \to \mathbb{N}$ by $\hat{F}(h) = F(h) - F(\emptyset)$. Then F is a cardinal representation of $\precsim$ if and only if $\hat{F}$ is. It is straightforward to show that, within the context of a set of histories, the notion of choice measurability coincides with the notion of "additive plausibility" discussed in [20]. In Proposition 14 we will deal with a more general notion that, unlike the notion of additive plausibility, can be applied to arbitrary sets.

Note that even if a plausibility order is choice measurable there is no guarantee that the canonical representation ρ is also a cardinal representation, that is, that ρ satisfies property IND_1C. This fact is illustrated in [5, Figure 4, p. 577].

IND_1C implies both IND_1 and IND_2, but the converse is not true, as can be seen from the game shown in Figure 3, which is based on an example discussed in [2] and [9].



The assessment $\sigma = (M, \ell, (a, c, e))$, $\mu(M) = \mu(Mr) = \mu(Lm) = \mu(R\ell) = 1$
is a perfect Bayesian equilibrium but cannot be rationalized by a
choice-measurable plausibility order.
Figure 3

The assessment $\sigma = (M, \ell, (a, c, e))$, $\mu(M) = \mu(Mr) = \mu(Lm) = \mu(R\ell) = 1$ is a perfect Bayesian equilibrium, rationalized by the following plausibility order, which satisfies IND_1 and IND_2. For simplicity we have omitted from the order the terminal histories, since they are not relevant for the following discussion.[8]

[8] The complete order is as follows:
$\emptyset \sim M \sim M\ell \prec R \sim R\ell \sim R\ell e \prec Mm \sim Mme \prec Mr \sim Mra \prec L \sim L\ell \sim L\ell a \prec Rm \prec Lm \sim Lmc \prec Rr \sim Rrc \prec Lr \prec Mrb \prec R\ell f \prec Mmf \prec L\ell b \prec Lmd \prec Rrd.$

$$\begin{pmatrix} \emptyset, M \\ R, R\ell \\ Mm \\ Mr \\ L, L\ell \\ Lm \\ Rr \end{pmatrix} \tag{5}$$

However, this plausibility order is *not* choice measurable as will be shown below.

Our objective is to find necessary and sufficient conditions for a plausibility order to be choice measurable. We start with a discussion of the plausibility order (5), which is reproduced below together with the canonical ordinal integer-valued representation ρ and the corresponding values of the cardinal function $F : H \to \mathbb{N}$ that we are seeking. Thus, for example, $F(L) = F(L\ell) = y_4$; in general, if $\rho(h) = i$ then $F(h)$ is denoted by y_i.

$$\begin{pmatrix} \precsim : & \rho : & F : \\ \emptyset, M & 0 & 0 \\ R, R\ell & 1 & y_1 \\ Mm & 2 & y_2 \\ Mr & 3 & y_3 \\ L, L\ell & 4 & y_4 \\ Lm & 5 & y_5 \\ Rr & 6 & y_6 \end{pmatrix} \tag{6}$$

Condition IND_1C imposes the following constraints:

$$\begin{array}{c} F(R) - F(M) = F(Rr) - F(Mr) \\ F(L) - F(M) = F(Lm) - F(Mm) \end{array} \tag{7}$$

which, using the notation of (6), can be written as

$$\begin{array}{c} y_1 = y_6 - y_3 \\ y_4 = y_5 - y_2 \end{array} \tag{8}$$

Thus the problem is to determine whether there exist positive integers y_i $(i = 1, ..., 6)$ such that (1) $y_i < y_{i+1}$ for all $i = 1, ..., 5$ and (2) the system of equations (8) is satisfied. The problem can be further simplified by defining new integers $x_1, ..., x_6$ as follows: for every $i = 1, ..., 6$, $x_i = y_i - y_{i-1}$ (letting $y_0 = 0$), so that, for $k > j$, $y_k - y_j = x_{j+1} + x_{j+2} + ... + x_k$ (for example, $y_5 - y_2 = \underbrace{x_3}_{=y_3-y_2} + \underbrace{x_4}_{=y_4-y_3} + \underbrace{x_5}_{=y_5-y_4}$). With this transformation, the constraints (8) can be rewritten as

$$\begin{array}{c} x_1 = x_4 + x_5 + x_6 \\ x_1 + x_2 + x_3 + x_4 = x_3 + x_4 + x_5 + x_6 \end{array} \tag{9}$$

and the problem of finding a cardinal representation of (5) reduces to determining whether there are positive integers $x_1, ..., x_6$ that satisfy the system of equations (9). The answer is negative, since subtracting the first equation from the second yields $x_2 + x_3 + x_4 = x_3$, that is, $x_2 + x_4 = 0$ which is not compatible with x_2 and x_4 being positive. Thus the plausibility order (5) is not choice measurable.[9]

We now provide necessary and sufficient conditions for a plausibility order to be choice measurable. The first characterization of choice measurability, which is known, is given in Proposition 11 and exploits the structure of the elements of H as sequences (and thus sets) of actions, while the second characterization is more general and, as far as I know, new.

For the first characterization we need to introduce some definitions. Given a finite set of histories H and a total pre-order $\precsim$ on H, let $C = \langle (h_1, h_2), (h_3, h_4), ..., (h_{n-1}, h_n) \rangle$ (with $n \geq 2$ and, clearly, n even), be a sequence in $\precsim$ (thus $h_i \precsim h_{i+1}$, for all odd i). Let $A(C) = \{a \in A : a \in h_i \text{ for some } i = 1, ..., n\}$ be the set of actions that appear in at lest one history in sequence C. For every $a \in A(C)$ define $Odd(a) = \{i : a \in h_i \text{ and } i \text{ is odd}\}$ and $Even(a) = \{i : a \in h_i \text{ and } i \text{ is even}\}$. We say that the sequence is *cancelling* if, for every $a \in A(C)$, the cardinality of $Odd(a)$ is equal to the cardinality of $Even(a)$, that is, if every action in $A(C)$ appears as many times in the odd-numbered histories as in the even-numbered histories. A cancelling sequence is *strict* if, for at least one pair (h_i, h_{i+1}) in C, $h_i \prec h_{i+1}$.

The following proposition is an application of a known result in qualitative probability theory ([10, 17, 7]) to the set of histories H, where each history is viewed as a set of actions. For a proof see [20, Lemma 4.1, p. 19].

Proposition 11 *A plausibility order $\precsim$ on H is choice measurable if and only there are no strict cancelling sequences in $\precsim$.*

As an application of the above proposition, consider the plausibility order (5), which we already showed not to be choice measurable. Then, by Proposition 11, there must be a strict cancelling sequence and indeed one such sequence is $\langle (R, Mm), (Mr, L), (Lm, Rr) \rangle$. In fact, $A(C) = \{L, M, R, m, r\}$, $Odd(R) = \{1\}$, $Even(R) = \{6\}$, $Odd(M) = \{3\}$, $Even(M) = \{2\}$, $Odd(m) = \{5\}$, $Even(m) = \{2\}$, $Odd(r) = \{3\}$, $Even(r) = \{6\}$, $Odd(L) = \{5\}$, $Even(L) = \{4\}$ and $R \prec Mm$.

The test for choice measurability provided by Proposition 11 has two drawbacks, one practical and one conceptual. From a practical point of view, the larger the set H (and thus the more complex the plausibility order $\precsim$ on H) the more difficult it is to determine whether or not a cancelling sequence exists. From a conceptual point of view, the necessary and sufficient condition identified in Proposition 11 makes no reference to (and no use of) the constraints corresponding to condition IND_1C, which are at the core of the notion of choice measurability (for example, the constraints (7) for the plausibility order (5)). The following characterization, on the other hand, makes explicit use of those constraints; furthermore, it applies to the more general case of a total pre-order on an *arbitrary* finite set S (whose elements need not have the structure of sets of atomes, e.g. sequences of actions).

Let S be an arbitrary finite set and let $\precsim$ 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 "$s \precsim t$ and $t \precsim s$"); 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

[9] It can be shown that *any* plausibility order that rationalizes the assessment shown in Figure 3 is not choice measurable.

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 $[t_2]$. The problem we are addressing is the following.

Problem 12 *Given a pair $(\precsim, \doteq)$, where $\precsim$ is a total pre-order on a finite set S and $\doteq$ is an equivalence relation on the set of equivalence classes of $\precsim$, determine whether there exists a function $F : S \to \mathbb{N}$ such that, for all $s, t, x, y \in S$, (1) $F(s) \leq F(t)$ if and only if $s \precsim t$ and (2) if $([s], [t]) \doteq ([x], [y])$, with $s \prec t$ and $x \prec y$, then $F(t) - F(s) = F(y) - F(x)$.*

Instead of expressing $\doteq$ 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 $\precsim$.[10] 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 $(\rho(s_1), \rho(s_2)) \doteq (\rho(t_1), \rho(t_2))$. For example, let $S = \{a, b, c, d, e, f, g, h, \ell, m\}$ and let $\precsim$ be as follows, with the corresponding canonical representation ρ (thus $a \prec x$ for every $x \in S\backslash\{a\}$, $[b] = \{b, c\}$, $b \prec d$, etc.):

$$\begin{pmatrix} \precsim : & \rho : \\ a & 0 \\ b, c & 1 \\ d & 2 \\ e & 3 \\ f, g & 4 \\ h, \ell & 5 \\ m & 6 \end{pmatrix} \tag{10}$$

If the equivalence relation $\doteq$ contains the following pairs:[11]

$$\begin{array}{c} ([a], [b]) \doteq ([h], [m]) \\ ([a], [b]) \doteq ([e], [f]) \\ ([a], [d]) \doteq ([f], [m]) \\ ([b], [e]) \doteq ([e], [f]) \\ ([b], [e]) \doteq ([f], [m]) \end{array} \quad \text{then we express them (using } \rho \text{) as} \quad \begin{array}{c} (0, 1) \doteq (5, 6) \\ (0, 1) \doteq (3, 4) \\ (0, 2) \doteq (4, 6) \\ (1, 3) \doteq (3, 4) \\ (1, 3) \doteq (4, 6) \end{array} \tag{11}$$

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, $\{1, 2, 4, 4, 5, 6\} \sqsubset \{1, 1, 2, 4, 4, 5, 5, 6\}$.

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

[10] As in Definition 8, let $S_0 = \{s \in S : s \precsim t, \ \ \forall t \in S\}$, and, for every integer $k \geq 1$, $S_k = \{h \in S \setminus S_0 \cup ... \cup S_{k-1} : s \precsim t, \ \ \forall t \in S \setminus S_0 \cup ... \cup S_{k-1}\}$. The canonical ordinal integer-valued representation of $\precsim$, $\rho : S \to \mathbb{N}$, is given by $\rho(s) = k$ if and only if $s \in S_k$.

[11] For example, $\doteq$ is the smallest reflexive, symmetric and transitive relation that contains the pairs given in (11).

every $k = 1, ..., m$) we associate with it the bag $B_P = B_{(i_1,j_1)} \uplus B_{(i_2,j_2)} \uplus ... \uplus B_{(i_m,j_m)}$. 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 13 *For every element of $\doteq$, expressed (using the canonical representation ρ) as $(i,j) \doteq (k,\ell)$ (with $i \leq j$ and $k \leq \ell$), the equation corresponding to it is $x_{i+1} + x_{i+2} + ... + x_j = x_{k+1} + x_{k+2} + ... + x_\ell$. 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 (10) 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:

$$\begin{aligned} x_1 + x_2 + x_3 &= x_3 + x_4 \\ x_3 + x_4 &= x_1 + x_2 + x_3 \\ x_3 + x_4 &= x_4 + x_5 \\ x_4 + x_5 &= x_3 + x_4 \\ x_1 + x_2 + x_3 &= x_4 + x_5 \\ x_4 + x_5 &= x_1 + x_2 + x_3 \end{aligned} \tag{12}$$

which can also be expressed as

$$Ax = 0 \text{ where } x = (x_1, ..., x_5) \text{ and } A = \begin{pmatrix} 1 & 1 & 0 & -1 & 0 \\ -1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 0 & 1 \\ 1 & 1 & 1 & -1 & -1 \\ -1 & -1 & -1 & 1 & 1 \end{pmatrix} \tag{13}$$

Note that, because of symmetry of $\doteq$, for every row a_k of the matrix A there is a row a_j such that $a_j = -a_k$.

We are now ready to state the solution to Problem 12. The proof is given in Appendix B.

Proposition 14 *Given a pair $(\precsim, \doteq)$, where $\precsim$ is a total pre-order on a finite set S and $\doteq$ is an equivalence relation on the set of equivalence classes of $\precsim$, the following are equivalent.*

(A) There is a function $F : S \to \mathbb{N}$ such that, for all $s, t, x, y \in S$, (1) $F(s) \leq F(t)$ if and only if $s \precsim t$ and (2) if $([s],[t]) \doteq ([x],[y])$, with $s \prec t$ and $x \prec y$, then $F(t) - F(s) = F(y) - F(x)$,[12]

(B) The system of equations corresponding to $\doteq$ (Definition 13) has a solution consisting of positive integers.

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

[12] Using the canonical ordinal representation of $\precsim$, condition (2) can also be expressed as follows: if $(i,j) \doteq (k,\ell)$ (with $i < j$, and $k < \ell$) and $\rho(s) = i$, $\rho(t) = j$, $\rho(x) = k$ and $\rho(y) = \ell$, then $F(t) - F(s) = F(y) - F(x)$.

As an application of Proposition 14 consider the total pre-order given in (5).[13] Two elements of $\doteq$ are $(M, Mr) \doteq (R, Rr)$ and $(L, Lm) \doteq (M, Mm)$, which - expressed in terms of the canonical ordinal representation ρ - can be written as

$$(0,3) \doteq (1,6)$$
$$(4,5) \doteq (0,2)$$

Then $B_{left} = \{1,2,3\} \uplus \{5\} = \{1,2,3,5\}$ and $B_{right} = \{2,3,4,5,6) \uplus \{1,2\} = \{1,2,2,3,4,5,6\}$. Thus, since $B_{left} \sqsubset B_{right}$, by Part C of the above proposition $\precsim$ is not choice measurable (as we had determined above with a different method).

As a further application of Proposition 14 consider the total pre-order $\precsim$ given in (10) together with the subset of the equivalence relation $\doteq$ given in (11). Then there is no cardinal representation of $\precsim$ that satisfies the constraints expressed by $\doteq$, because of Part C of the above proposition and the following sequence:[14]

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

where $B_{left} = \{0,1\} \uplus \{1,3\} \uplus \{3,4\} \uplus \{4,6\} = \{1,2,3,4,5,6\} \sqsubset B_{right} = \{3,4\} \uplus \{4,6\} \uplus \{1,3\} \uplus \{0,2\} = \{1,2,2,3,4,5,6\}$.
In fact, the above sequence corresponds to the following system of equations:

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

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 (in particular with $x_2 > 0$).

Remark 15 *In [6] 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 Perfect Bayesian equilibrium and sequential equilibrium.

We now show that choice measurability of the underlying plausibility order is what essentially is needed to go from perfect Bayesian equilibrium to 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 $\langle \sigma^1, ..., \sigma^m, ... \rangle$ of completely mixed strategy profiles such that, letting μ^m be the unique system of beliefs obtained from σ^m by applying Bayes' rule,[15]

[13] A more complete, but tedious, argument would make use of the full total pre-order given in Footnote 8.

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

[15] That is, for every $h \in D$, $\mu^m(h) = \dfrac{\prod_{a \in h} \sigma^m(a)}{\sum_{x \in I(h)} \prod_{a \in x} \sigma^m(a)}$, where $a \in h$ means that action a occurs 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 $h \in D$.

$\lim_{m\to\infty}(\sigma^m,\mu^m) = (\sigma,\mu)$. An assessment (σ,μ) is a *sequential equilibrium* if it is KW-consistent and sequentially rational.

It is shown in [5] that if (σ,μ) is a KW-consistent assessment then there is a choice-measurable plausibility order that rationalizes it and that the notion of sequential equilibrium is a strict refinement of perfect Bayesian equilibrium. In this section we show that choice measurability together with a strengthening of Definition 4 is *necessary and sufficient* for a perfect Bayesian equilibrium to be a sequential equilibrium.

Given a system of beliefs μ we denote by D^+_μ the set of decision histories that are assigned positive probability by μ, that is, $D^+_\mu = \{h \in D : \mu(h) > 0\}$.

Definition 16 *Fix an extensive form. Let* $\precsim$ *be a plausibility order that rationalizes the assessment* (σ,μ). *We say that* (σ,μ) *is* uniformly Bayesian relative to $\precsim$ *if there exists a function* $\nu : D \to (0,1]$ *that satisfies the following properties:*

UB1. $\forall h,h' \in D$, $\forall a \in A(h)$ *if* $h' \in I(h)$ *then* $\frac{\nu(ha)}{\nu(h)} = \frac{\nu(h'a)}{\nu(h')}$.

UB2. $\forall h \in D$, $\forall a \in A(h)$ *if* $h \sim ha$ *then* $\nu(ha) = \nu(h) \times \sigma(a)$.

UB3. If $E \subseteq H$ *is an equivalence class of* $\precsim$ *such that* $E \cap D^+_\mu \neq \varnothing$ *then* $\nu_E : H \to [0,1]$ *defined by* $\nu_E(h) = \begin{cases} 0 & \text{if } h \notin E \cap D^+_\mu \\ \frac{\nu(h)}{\sum_{h' \in E \cap D^+_\mu} \nu(h')} & \text{if } h \in E \cap D^+_\mu \end{cases}$ *satisfies Property B3 of Definition 4.*

The following lemma, proved in Appendix B, shows that Definition 16 is a strengthening of Definition 4.

Lemma 17 *Fix an extensive form. Let* $\precsim$ *be a plausibility order that rationalizes the assessment* (σ,μ). *If* (σ,μ) *is uniformly Bayesian relative to* $\precsim$ *then it is Bayesian relative to* $\precsim$.

The following proposition, which is proved in Appendix B, makes use of a result proved in [19].

Proposition 18 *Fix an extensive form and an assessment* (σ,μ). *The following are equivalent:*

(A) (σ,μ) *is a perfect Bayesian equilibrium which is rationalized by a choice measurable plausibility order* $\precsim$ *and is uniformly Bayesian relative to* $\precsim$.

(B) (σ,μ) *is a sequential equilibrium.*

As an application of Proposition 18 consider the extensive game of Figure 4. Let (σ,μ) be an assessment with $\sigma = ((a,f),T,L)$, $\mu(b) > 0$ and $\mu(c) > 0$. Then (σ,μ) can be rationalized by a choice-measurable plausibility order only if μ is such that[16]

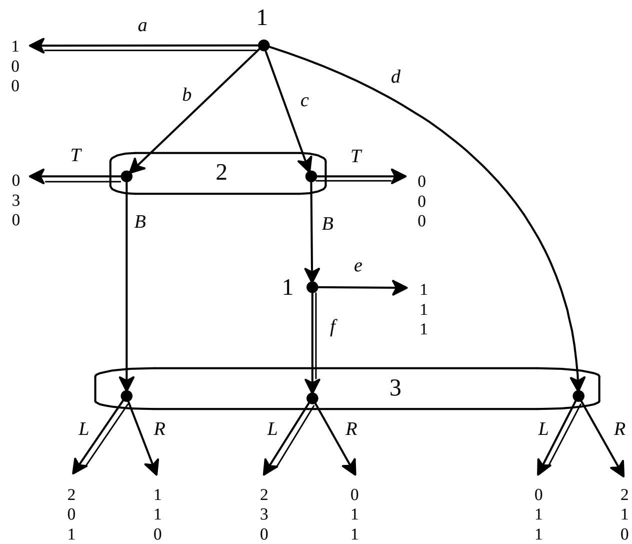
$$\text{either } \mu(bB) = \mu(cBf) = 0, \text{ or } \mu(bB) > 0 \text{ and } \mu(cBf) > 0. \tag{14}$$

[16]Proof. Let $\precsim$ be a choice measurable plausibility order that rationalizes (σ,μ) and let F be a cardinal representation of it. Since $\mu(b) > 0$ and $\mu(c) > 0$, $b \sim c$ and thus $F(b) = F(c)$. By choice measurability, $F(bB) - F(cB) = F(b) - F(c)$ and thus $F(bB) = F(cB)$, so that $bB \sim cB$. Since f is plausibility preserving, $cB \sim cBf$ and therefore, by transitivity of $\precsim$, $bB \sim cBf$. Hence if $\mu(bB) > 0$ then $bB \in Min_{\precsim}\{bB, cBf, d\}$ and thus $cBf \in Min_{\precsim}\{bB, cBf, d\}$ so that $\mu(cBf) > 0$. The proof that if $\mu(cBf) > 0$ then $\mu(bB) > 0$ is similar.

If, besides from being rationalized by a choice-measurable plausibility order $\precsim$, (σ, μ) is also uniformly Bayesian relative to $\precsim$ (Definition 16), then[17]

$$\mu(bB) > 0 \; \Rightarrow \frac{\mu(cBf)}{\mu(bB)} = \frac{\mu(c)}{\mu(b)}. \tag{15}$$

Thus for example, continuing to assume that $\sigma = ((a, f), T, L)$, the assessment $(\sigma, \tilde{\mu})$ with $\tilde{\mu}(b) = \frac{7}{10}$, $\tilde{\mu}(c) = \frac{3}{10}$, $\tilde{\mu}(bB) = \frac{7}{18}$, $\tilde{\mu}(cBF) = \frac{3}{18}$ and $\tilde{\mu}(d) = \frac{8}{18}$ is a sequential equilibrium,[18] while the assessment $(\sigma, \hat{\mu})$ with $\hat{\mu}(b) = \frac{7}{10}$, $\hat{\mu}(c) = \frac{3}{10}$, $\hat{\mu}(bB) = \hat{\mu}(cBF) = \hat{\mu}(d) = \frac{1}{3}$ is a perfect Bayesian equilibrium but not a sequential equilibrium.[19]



Implications of choice measurability and the uniform Bayesian property
for assessments of the form $\sigma = ((a, f), T, L)$, $\mu(b) > 0$, $\mu(c) > 0$.
Figure 4

[17] Proof. Suppose that $\mu(b) > 0$, $\mu(c) > 0$ and $\mu(bB) > 0$. Let ν be a function that satisfies the properties of Definition 16. Then, by $UB1$, $\frac{\nu(bB)}{\nu(b)} = \frac{\nu(cB)}{\nu(c)}$ so that $\frac{\nu(c)}{\nu(b)} = \frac{\nu(cB)}{\nu(bB)}$ and, since $\sigma(f) = 1 > 0$, $\nu(cBf) = \nu(cB) \times \sigma(f) = \nu(cB)$. Let E be the equivalence class that contains b. Then $\{b, c\} \subseteq E \cap D^+_\mu$. By $UB3$, by $\mu(b) = \frac{\nu(b)}{\sum_{h \in E \cap D^+_\mu}{}' \, v(h')}$ and $\mu(c) = \frac{\nu(c)}{\sum_{h \in E \cap D^+_\mu}{}' \, v(h')}$, so that $\frac{\mu(c)}{\mu(b)} = \frac{\nu(c)}{\nu(b)}$. Let G be the equivalence class that contains bB. Then, by (14), $\{bB, cBf\} \subseteq G \cap D^+_\mu$ and, by $UB3$, $\mu(bB) = \frac{\nu(bB)}{\sum_{h \in G \cap D^+_\mu}{}' \, v(h')}$ and $\mu(cBf) = \frac{\nu(cBf)}{\sum_{h \in G \cap D^+_\mu}{}' \, v(h')}$, so that $\frac{\mu(cBf)}{\mu(bB)} = \frac{\nu(cBf)}{\nu(bB)}$. Thus, since $\nu(cBf) = \nu(cB)$, $\frac{\mu(cBf)}{\mu(bB)} = \frac{\nu(cB)}{\nu(bB)}$ and, therefore, since - as shown above - $\frac{\nu(c)}{\nu(b)} = \frac{\nu(cB)}{\nu(bB)}$ we have that $\frac{\mu(cBf)}{\mu(bB)} = \frac{\mu(c)}{\mu(b)}$.

[18] It follows from Proposition 18 and the fact that (σ, μ) is rationalized by the following choice-measurable plausibility order: $\left(\begin{array}{ll} \precsim \;\; : & F: \\ \emptyset, a & 0 \\ b, c, bT, cT & 1 \\ d, dB, cB, dL, dBL, cBf, cBfL & 2 \\ bBR, cBe, cBfR, dR & 3 \end{array} \right)$ and we can take the function ν to be as follows: $\nu(\emptyset) = 1$, $\nu(b) = \nu(bB) = \frac{7}{18}$, $\nu(c) = \nu(cB) = \nu(cBf) = \frac{3}{18}$, $\nu(d) = \frac{8}{18}$.

[19] Both $(\sigma, \tilde{\mu})$ and $(\sigma, \hat{\mu})$ are rationalized by the choice measurable plausibility order given in Footnote 18.

One can argue that the core of the characterization of sequential equilibrium provided in Proposition 18 is the notion of choice measurability of the order that rationalizes the given assessment. As shown in the above example, choice measurability constrains the support of μ, while the additional condition that the assessment be uniformly Bayesian adds the requirement that the ratios of probabilities be preserved (whenever meaningful). Hence, Proposition 18, in conjunction with Proposition 14, provides an essentially qualitative characterization of sequential equilibrium.[20]

5 Related literature

The work which is most closely related to this paper is [19, 20]. In [20] the author shows the following (using our notation). A set of actions and histories $b \subseteq A \cup H$ is called a "basement" if it coincides with the support of at least one assessment, that is, if there is an assessment (σ, μ) such that $a \in b$ if and only if $\sigma(a) > 0$ and $h \in b$ if and only if $\mu(h) > 0$. Given a basement, one can construct a partial relation $\precsim$ on the set of histories H as follows: (1) if h and h' belong to the same information set then $h \sim h'$ if $h, h' \in b$ and $h \prec h'$ if $h \in b$ and $h' \notin b$, (2) if $h' = ha$ then $h \sim h'$ if $a \in b$ and $h \prec h'$ if $a \notin b$. Streufert calls this relation a "plausibility" relation. A total pre-order $\precsim^*$ on H is said to be "additively representable" if there exists a function $\lambda : A \to \mathbb{R}$, which Streufert calls a "mass function", such that $h \precsim^* h'$ if and only if $\sum_{a \in h} \lambda(a) \leq \sum_{a \in h'} \lambda(a)$. Streufert proves the following result.

Proposition 19 *[20, Theorem 1 and Corollary 1, pp. 17 and 20] If the plausibility relation $\precsim$ derived from a basement b can be extended to a total pre-order $\precsim^*$ that has an additive representation, then there exists a KW-consistent assessment (σ, μ) whose support coincides with b. Conversely, given a KW-consistent assessment (σ, μ), the plausibility relation $\precsim$ derived from its basement can be extended to a total pre-order $\precsim^*$ that has an additive representation.*

Streufert also proves [20, Lemma 4.1, p. 19] a result which is strictly related to Proposition 11 above, namely that if $\precsim$ is a binary relation on the set of subsets of a finite set A then it has a completion represented additively by a mass function $\lambda : A \to \mathbb{Z}$ if and only if it there is no strict canceling sequence.

There is a clear connection between Proposition 19 and Proposition 18. However, while Proposition 19 characterizes basements that are supported by a KW-consistent assessment, Proposition 18 focuses on a particular PBE (σ, μ) and on the conditions that are necessary and sufficient for (σ, μ) to be a sequential equilibrium. One of these conditions is choice measurability of the plausibility order that rationalizes (σ, μ), which coincides with the existence of an additive representation of a total pre-order that extends the plausibility relation obtained from the support of (σ, μ). The other condition is that (σ, μ) be uniformly Bayesian relative to $\precsim$. As detailed in the proof of Proposition 18, this condition is related to the existence of a pair of functions $c : A \to (0, 1]$ and

[20]Kreps and Wilson themselves [11, p. 876] express dissatisfaction with their definition of sequential equilibrium: "We shall proceed here to develop the properties of sequential equilibrium as defined above; however, we do so with some doubts of our own concerning what 'ought' to be the definition of a consistent assessment that, with sequential rationality, will give the 'proper' definition of a sequential equilibrium." In a similar vein, Osborne and Rubinstein [12, p. 225] write "we do not find the consistency requirement to be natural, since it is stated in terms of limits; it appears to be a rather opaque technical assumption". In these quotations "consistency" corresponds to what we called "KW-consistency".

$e : A \to \mathbb{Z}^- \cup \{0\}$ that Streufert proves to be necessary and sufficient for KW-consistency (see Proposition 23 in Appendix B).[21]

On the other hand, the characterization of choice measurability of a total pre-order provided in Proposition 14 is new and more general than Proposition 11 because it applies to arbitrary sets, whose elements do not necessarily have the structure of sets of atoms (such as actions).

Propositions 18, 19 and 23 are all related to two results provided in [11, Lemma A1, p. 887 and Lemma A2, p. 888]. For a detailed discussion of those results (and how to correct a flaw in the original proof) see [19].

The characterization of sequential equilibrium provided in Proposition 18 does not make any use of sequences and limits. This is true also of the characterizations of KW-consistency provided in [9] and [13], where it is shown how to derive a finite system of algebraic equations and inequalities on behavioral strategies and beliefs that characterizes the set of consistent assessments ([3] provides an indirect proof of the fact that consistent assessments are determined by finitely many algebraic equations and inequalities). [9] makes use of relative probability spaces (Ω, ρ) (which express the notion of an event being infinitely less likely than another) and random variables $s_i : \Omega \to S_i$ (where S_i is the set of pure strategies of player i), representing the beliefs of an external observer (who can assess the relative probabilities of any two strategy profiles, even those that have zero probability). The authors provide a characterization of KW-consistency in terms of the notion of strong independence for relative probability spaces (and, in turn, a characterization of strong independence in terms of weak independence and exchangeability).[22]

The algebraic characterization of KW-consistency provided in [13] is based on a system of equations and inequalities obtained from two functions: one defined on actions and the other on histories. The function defined on actions implicitly defines an "order of likelihood" on actions, with some actions being "infinitely less likely" than others. Together the two functions provide an "extended" behavioral conjecture profile. The characterization is essentially equivalent to the one reported in Proposition 23 in Appendix B and for details the reader is referred to Footnote 28.

6 Conclusion

Besides sequential rationality, the notion of perfect Bayesian equilibrium introduced in [5] is based on two elements: (1) the qualitative notions of plausibility order and AGM-consistency and (2) the notion of Bayesian updating relative to the plausibility order (which captures the requirement that Bayes' rule be used whenever possible, even after revision prompted by the observation of a zero probability event). In this paper we showed (Proposition 18) that by strengthening these two conditions one obtains a 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 distance between histories in a way that is preserved by the addition of a common action (condition IND_1C). Proposition 14 provides a qualitative characterization of choice measurability, which is very general, in that it applies to arbitrary sets (thus not only to sets consisting of sequences of actions); for instance it can be applied to the problem of determining if

[21] That proposition, in turn, is essentially equivalent to a result in [13]: see Footnote 28 in Appendix B. At the 13^{th} SAET conference in July 2013 Streufert presented a characterization of the support of KW-consistent assessments in terms of additive plausibility and a condition that he called "pseudo-Bayesian" which is essentially a reformulation of one of the conditions given in Proposition 23 (see Appendix B).

[22] [2] shows that in games with observable deviators weak independence suffices for KW-consistency.

there exists a cardinal utility function on a set of alternatives S under a set of constraints of the form $([a],[b]) \doteq ([c],[d])$, interpreted as "the change from a to b is just as good (or just as bad) as the change from c to d.

The strengthening of the second condition is called "uniform Bayesian updating" and amounts to the requirement that the relative ratios of the probabilities of two histories be preserved (whenever possible and meaningful: see the discussion of the example that follows Proposition 14).

A Appendix: history-based definition of extensive game

If A is a set, we denote by A^* the set of finite sequences in A. If $h = \langle a_1, ..., a_k \rangle \in A^*$ and $1 \leq j \leq k$, the sequence $h' = \langle a_1, ..., a_j \rangle$ is called a *prefix* of h. If $h = \langle a_1, ..., a_k \rangle \in A^*$ and $a \in A$, we denote the sequence $\langle a_1, ..., a_k, a \rangle$ by ha.

A *finite extensive form* is a tuple $\left\langle A, H, N, \iota, \{\approx_i\}_{i \in N} \right\rangle$ whose elements are:

- A finite set of actions A.
- A finite set of histories $H \subseteq A^*$ which is closed under prefixes (that is, if $h \in H$ and $h' \in A^*$ is a prefix of h, then $h' \in H$). The null history $\langle \rangle$, denoted by $\emptyset$, is an element of H and is a prefix of every history. A history $h \in H$ such that, for every $a \in A$, $ha \notin H$, is called a *terminal history*. The set of terminal histories is denoted by Z. $D = H \backslash Z$ denotes the set of non-terminal or *decision* histories. For every history $h \in H$, we denote by $A(h)$ the set of actions available at h, that is, $A(h) = \{a \in A : ha \in H\}$. Thus $A(h) \neq \varnothing$ if and only if $h \in D$. We assume that $A = \bigcup_{h \in D} A(h)$ (that is, we restrict attention to actions that are available at some decision history).
- A finite set $N = \{1, ..., n\}$ of players. In some cases there is also an additional, fictitious, player called *chance*.
- A function $\iota : D \to N \cup \{chance\}$ that assigns a player to each decision history. Thus $\iota(h)$ is the player who moves at history h. A game is said to be *without chance moves* if $\iota(h) \in N$ for every $h \in D$. For every $i \in N \cup \{chance\}$, let $D_i = \iota^{-1}(i)$ be the histories assigned to player i. Thus $\{D_{chance}, D_1, ..., D_n\}$ is a partition of D. If history h is assigned to chance, then a probability distribution over $A(h)$ is given that assigns positive probability to every $a \in A(h)$.
- For every player $i \in N$, $\approx_i$ is an equivalence relation on D_i. The interpretation of $h \approx_i h'$ is that, when choosing an action at history $h \in D_i$, player i does not know whether she is moving at h or at h'. The equivalence class of $h \in D_i$ is denoted by $I_i(h)$ and is called an *information set of player* i; thus $I_i(h) = \{h' \in D_i : h \approx_i h'\}$. The following restriction applies: if $h' \in I_i(h)$ then $A(h') = A(h)$, that is, the set of actions available to a player is the same at any two histories that belong to the same information set of that player.
- The following property, known as *perfect recall*, is assumed: for every player $i \in N$, if $h_1, h_2 \in D_i$, $a \in A(h_1)$ and $h_1 a$ is a prefix of h_2 then for every $h' \in I_i(h_2)$ there exists an $h \in I_i(h_1)$ such that ha is a prefix of h'. Intuitively, perfect recall requires a player to remember what she knew in the past and what actions she took previously.

Given an extensive form, one obtains an *extensive game* by adding, for every player $i \in N$, a *utility (*or *payoff) function* $U_i : Z \to \mathbb{R}$ (where $\mathbb{R}$ denotes the set of real numbers; recall that Z is the set of terminal histories).

Notation 20 *If h and h' are decision histories not assigned to chance, we often write $h' \in I(h)$ as a short-hand for $h' \in I_{\iota(h)}(h)$. Thus $h' \in I(h)$ means that h and h' belong to the same information set (of the player who moves at h). If h is a history assigned to chance, we use the convention that $I(h) = \{h\}$.*

Remark 21 *In order to simplify the notation we assume that no action is available at more than one information set: $\forall h, h' \in H, \forall a \in A$, if $a \in A(h) \cap A(h')$ then $h' \in I(h)$.*

Given an extensive form, a *pure strategy* of player $i \in N$ is a function that associates with every information set of player i an action at that information set, that is, a function $s_i : D_i \to A$ such that (1) $s_i(h) \in A(h)$ and (2) if $h' \in I_i(h)$ then $s_i(h') = s_i(h)$. A *behavior strategy* of player i is a collection of probability distributions, one for each information set, over the actions available at that information set; that is, a function $\sigma_i : D_i \to \Delta(A)$ (where $\Delta(A)$ denotes the set of probability distributions over A) such that (1) $\sigma_i(h)$ is a probability distribution over $A(h)$ and (2) if $h' \in I_i(h)$ then $\sigma_i(h') = \sigma_i(h)$. If the game does not have chance moves, we define a behavior strategy *profile* as an n-tuple $\sigma = (\sigma_1, ..., \sigma_n)$ where, for every $i \in N$, σ_i is a behavior strategy of player i. If the game has chance moves then we use the convention that a behavior strategy profile is an $(n+1)$-tuple $\sigma = (\sigma_1, ..., \sigma_n, \sigma_{chance})$ where, if h is a history assigned to chance and $a \in A(h)$ then $\sigma_{chance}(h)(a)$ is the probability associated with a. Given our assumption that no action is available at more than one information set, without risking ambiguity we denote by $\sigma(a)$ the probability assigned to action a by the relevant component of the strategy profile σ. Note that a pure strategy is a special case of a behavior strategy where each probability distribution is degenerate. A behavior strategy is *completely mixed* at history $h \in D$ if, for every $a \in A(h)$, $\sigma(a) > 0$.

An *assessment* is a pair (σ, μ) where σ is a behavior strategy profile and μ is a system of beliefs, that is, a collection of probability distributions, one for every information set, over the elements of that information set. Thus $\mu : D \to \Delta(H)$ (where $\Delta(H)$ is the set of probability distributions over the set of histories H) such that if $h \in D$ then $\mu(h)$ is a probability distribution over $I(h)$ (the information set that constains history h) and if $h' \in I(h)$ then $\mu(h) = \mu(h')$. If the game has chance moves, then we use the convention that $\mu(h) = 1$ for every history h assigned to chance. With slight abose of notation, we denote by $\mu(h)$ the probability assigned to history h by the system of beliefs μ [23]

B Appendix: proofs

Proof of Lemma 9. An example of a plausibility order that satisfies IND_2 but violates IND_1 is given in (1) for the assessment illustrated in Figure 2 (IND_1 is violated since $a \prec b$ but $be \prec ae$).

On the other hand, the following example shows a plausibility order that satisfies IND_1 but not IND_2 (the underlying game is a simultaneous game where Player 1 chooses between a and b and Player 2 chooses among c, d and e; the plausibility-preserving actions are a and c; note that $I(a) = \{a, b\}$ and $ad \prec ae$ but $be \prec bd$]:

[23] A more precise notation would be $\mu(h)(h)$: if $h \in D$ then $\mu(h)$ is a probability distribution over $I(h)$ and, for every $h' \in I(h)$, $\mu(h) = \mu(h')$ so that $\mu(h)(h) = \mu(h')(h)$. We denote this common probability by $\mu(h)$).

$$\begin{pmatrix} \emptyset, a, ac \\ b, bc \\ ad \\ ae \\ be \\ bd \end{pmatrix}$$

Next we show that $IND_1C \Rightarrow IND_2C$. Let $\precsim$ be a plausibility order on the set of histories H and let $F : H \to \mathbb{N}$ be an integer-valued representation that satisfies IND_1C. Without loss of generality (see Definition 10), we can assume that $F(\emptyset) = 0$. For every decision history h and action $a \in A(h)$, define

$$\lambda(a) = F(ha) - F(h). \tag{16}$$

The function $\lambda : A \to \mathbb{N}$ is well defined, since, by IND_1C, $h' \in I(h)$ implies that $F(h'a) - F(h') = F(ha) - F(h)$ (recall also, see Remark 21 in Appendix A, the assumption that no action belongs to two different information sets). Then, for every history $h = \langle a_1, a_2, ..., a_m \rangle$, $F(h) = \sum_{i=1}^{m} \lambda(a_i)$. In fact,

$\lambda(a_1) + \lambda(a_2) + ... + \lambda(a_m) =$
$= (F(a_1) - F(\emptyset)) + (F(a_1a_2) - F(a_1)) + ... + (F(a_1a_2...a_m) - F(a_1a_2...a_{m-1})) =$
$= F(a_1a_2...a_m) = F(h)$ (recall that $F(\emptyset) = 0$).

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, $F(hb) - F(ha) = F(h'b) - F(h'a)$.

Finally, we show that $IND_2C \Rightarrow IND_1C$. Let $\precsim$ be a plausibility order on the set of histories H and let $F : H \to \mathbb{N}$ be an integer-valued representation that satisfies IND_2C. Fix 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 $\precsim$, $F(hb) = F(h)$ and $F(h'b) = F(h')$ and thus

$$F(h') - F(h) = F(h'b) - F(hb). \tag{17}$$

By IND_2C, $F(h'b) - F(hb) = F(h'a) - F(ha)$. From this and (17) it follows that $F(h') - F(h) = F(h'a) - F(ha)$. ■

Proof of Proposition 14. $(A) \Rightarrow (B)$. Let $F' : S \to \mathbb{N}$ satisfy the properties of Part A. Fix an arbitrary $s_0 \in S_0 = \{s \in S : s \precsim t, \forall t \in S\}$ and define $F : S \to \mathbb{N}$ 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 $K = \{k \in \mathbb{N} : k = \rho(s)$ for some $s \in S\}$ (where ρ is the canonical ordinal representation of $\precsim$: see Footnote 10). For every $k \in K$, define

$$\begin{array}{l} \hat{x}_0 = 0 \\ \text{and, for } k > 0, \\ \hat{x}_k = F(t) - F(s) \quad \text{for some } s, t \in S \text{ such that } \rho(t) = k \text{ and } \rho(s) = k - 1. \end{array} \tag{18}$$

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 $k \in K \backslash \{0\}$, $\hat{x}_k$ is a positive integer, since $\rho(t) = k$ and $\rho(s) = k - 1$ imply that $s \prec t$ 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 (18) provide a solution to the system of equations corresponding to $\doteq$ (Deinition 13). Fix an arbitrary element of $\doteq$, $([s_1],[s_2]) \doteq ([t_1],[t_2])$ (with $s_1 \prec s_2$ and $t_1 \prec t_2$) and express it, using the canonical ordinal representation ρ (see Footnote 10), 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 13) is: $x_{i_1+1} + x_{i_1+2} + ... + x_{i_2} = x_{j_1+1} + x_{j_1+2} + ... + x_{j_2}$. By (2) of Property A,

$$F(s_2) - F(s_1) = F(t_2) - F(t_1) \tag{19}$$

Using (18), $F(s_2)-F(s_1) = \hat{x}_{i_1+1}+\hat{x}_{i_1+2}+...+\hat{x}_{i_2}$. To see this, for every $k \in \{i_1+1, i_1+2, ..., i_2-1\}$, fix an arbitrary $r_k \in S$ such that $\rho(r_k) = k$; then, by (18),

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

Similarly, $F(t_2) - F(t_1) = \hat{x}_{j_1+1} + \hat{x}_{j_1+2} + ... + \hat{x}_{j_2}$. Thus, by (19), $\hat{x}_{i_1+1} + \hat{x}_{i_1+2} + ... + \hat{x}_{i_2} = \hat{x}_{j_1+1} + \hat{x}_{j_1+2} + ... + \hat{x}_{j_2}$.

$(B) \Rightarrow (A)$. assume that the system of equations corresponding to $\doteq$ has a solution consisting of positive integers $\hat{x}_1, ..., \hat{x}_m$. Define $F : S \to \mathbb{N}$ 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 $F(s) = \hat{x}_1 + \hat{x}_2 + ... + \hat{x}_k$ (where ρ and the sets S_k are as defined in Footnote 10). We need to show that F satisfies the properties of Part A. Fix arbitrary $s, t \in S$ with $s \precsim t$. Then $\rho(s) \le \rho(t)$ and thus $F(s) = \hat{x}_1 + \hat{x}_2 + ... + \hat{x}_{\rho(s)} \le F(t) = \hat{x}_1 + \hat{x}_2 + ... + \hat{x}_{\rho(s)} + \hat{x}_{\rho(s)+1} + ... + \hat{x}_{\rho(t)}$. Conversely, suppose that $s, t \in S$ are such that $F(s) \le F(t)$. Then $\hat{x}_1 + \hat{x}_2 + ... + \hat{x}_{\rho(s)} \le \hat{x}_1 + \hat{x}_2 + ... + \hat{x}_{\rho(t)}$ and thus $\rho(s) \le \rho(t)$, so that $s \precsim t$. Thus Property (1) of Part A is satisfied. Now let $s, t, x, y \in S$ be such that $s \prec t$, $x \prec y$ 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 (18), $F(t) - F(s) = \hat{x}_{i+1} + \hat{x}_{i+2} + ... + \hat{x}_j$ and $F(y) - F(x) = \hat{x}_{k+1} + \hat{x}_{k+2} + ... + \hat{x}_\ell$. Since $x_{i+1} + x_{i+2} + ... + x_j = x_{k+1} + x_{k+2} + ... + x_\ell$ 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 $\hat{x}_{i+1} + \hat{x}_{i+2} + ... + \hat{x}_j = \hat{x}_{k+1} + \hat{x}_{k+2} + ... + \hat{x}_\ell$ 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 $\precsim$) $\langle ((i_1,j_1) \doteq (k_1,\ell_1)), ..., ((i_m,j_m) \doteq (k_m,\ell_m))\rangle$ such that

$$B_{left} \sqsubset B_{right} \tag{20}$$

where $B_{left} = B_{(i_1,j_1)} \uplus ... \uplus B_{(i_m,j_m)}$ and $B_{right} = B_{(k_1,\ell_1)} \uplus ... \uplus B_{(k_m,\ell_m)}$. Let $\mathbb{E} = \{E_1, ..., E_m\}$ be the system of equations corresponding to the above sequence (for example, E_1 is the equation $x_{i_1+1} + x_{i_1+2} + ... + x_{j_1} = x_{k_1+1} + x_{k_1+2} + ... + x_{\ell_1}$). Let L be the sum of the left-hand-side and R be the sum of the right-hand-side of the equations $E_1, ..., E_m$. 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 (20), if nx_i is a summand of L then mx_i is a summand of R with $m \ge n$ and, furthermore, $L \ne R$. 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 $i = 1, ..., m$, (2) $x < y$ if $x \leq y$ and $x \neq y$ and (3) $x \ll y$ if $x_i < y_i$, for every $i = 1, ..., m$.

Lemma 22 *Let A be an $m \times n$ matrix such that the system of equations corresponding to $\doteq$ (Definition 13) can be expressed as $Ax = 0$ (recall that, by symmetry of $\doteq$, for each row a_i of A there is another row a_k such that $a_k = -a_i$; an example of such a matrix is given in (13); note also that each entry of A is either -1, 0 or 1). If the system of equations $Ax = 0$ does not have a positive integer solution then there exist r rows of A, $a_{i_1}, ..., a_{i_r}$ with $1 < r \leq \frac{m}{2}$ and r positive integers $\alpha_1, ..., \alpha_r \in \mathbb{N}\backslash\{0\}$ such that if B is the submatrix of A consisting of the r rows $a_{i_1}, ..., a_{i_r}$ (thus for every $k = 1, ..., r$, $b_k = a_{i_k}$, where b_k is the k^{th} row of B) then $\sum_{k=1}^{r} \alpha_k b_k < 0$.*

Proof. By Stiemke's theorem[24] 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 $r \leq \frac{m}{2}$.[25] Furthermore, again without loss of generality, we can assume that for every $k \in K$, $y_k > 0$.[26] 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 $i = 1, ..., r$ let $\alpha = (\alpha_1, ..., \alpha_r)$ be the vector corresponding to $(y_k)_{k \in K}$.[27] Then $\alpha B = \sum_{j=1}^{r} \alpha_j b_j = yA < 0$ and $\alpha_i \in \mathbb{N}\backslash\{0\}$ for all $i = 1, ..., r$. ■

Completion of proof of Proposition 14. 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 $\alpha = (\alpha_1, ..., \alpha_r)$ the vector of positive integers of Lemma 22 such that $\alpha B = \sum_{j=1}^{r} \alpha_j b_j < 0$. Define two $r \times n$ matrices $C = (c_{ij})_{i=1,...,r;\ j=1,...,n}$ and $D = (d_{ij})_{i=1,...,r;\ j=1,...,n}$ as follows (recall that each entry of B is either -1, 0 or 1):

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

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

$$\sum_{i=1}^{r} \alpha_i c_i < \sum_{i=1}^{r} \alpha_i d_i. \tag{21}$$

[24] See, for example, [16, p. 216] or [7, Theorem 1.1, p. 65].

[25] Proof. 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

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

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.

[26] Proof. 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 $\tilde{y}_i = -y_k$.

[27] For 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 $\alpha_3 = 3$.

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 (21) that

$$\sum_{i=1}^{s} c'_i < \sum_{i=1}^{s} d'_i. \tag{22}$$

Consider the system of equations

$$C'x = D'x. \tag{23}$$

For every $j = 1, ..., n$, 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 (23) 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 (23). Hence, by (22), for every $j = 1, ..., n$, the number of times that the variable x_j appears on the left-hand-side of (23) is less than or equal to the number of times that it appears on the right-hand-side of (23) and for at least one j it is less. Thus, letting $\langle ((i_1, j_1) \doteq (k_1, \ell_1)), ..., ((i_s, j_s) \doteq (k_s, \ell_s)) \rangle$ be the sequence of elements of $\doteq$ corresponding to the equations in (23), we have that $B_{left} \sqsubset B_{right}$ where $B_{left} = B_{(i_1,j_1)} \uplus ... \uplus B_{(i_m,j_m)}$ and $B_{right} = B_{(k_1,\ell_1)} \uplus ... \uplus B_{(k_m,\ell_m)}$. ■

Proof of Lemma 17. Let (σ, μ) be rationalized by the plausibility order $\precsim$ and be uniformly Bayesian relative to $\precsim$. We need to show that (σ, μ) is Bayesian relative to $\precsim$. Fix an arbitrary equivalence class E of $\precsim$ that contains a history h such that $\mu(h) > 0$. Let $\nu_E : H \to [0, 1]$ be as defined in $UB3$ of Definition 16. Then, by construction, ν_E satisfies Property $B1$ of Definition 4 and, by $UB3$, it also satisfies Property $B3$. Thus it only remains to show that ν_E satisfies Property $B2$ of Definition 4. Fix arbitrary $h, h' \in E \cap D_\mu^+$ such that $h' = ha_1...a_m$. Since $h, h' \in E$, $h' \sim h$. By Property $PL1$ of definition of plausibility order (Definition 1), $h \precsim ha_1 \precsim ha_1a_2 \precsim ... \precsim ha_1...a_m = h'$. Thus, since $h' \sim h$, by transitivity of $\precsim$, $h \sim ha_1 \sim ha_1a_2 \sim ... \sim ha_1...a_m = h'$ (that is, $ha_1...a_j \in E$ for all $j = 1, ..., m$) and therefore each action a_j ($j = 1, ..., m$) is plausibility preserving; hence, by Property $P1$ of Definition 2, $\sigma(a_j) > 0$ for all $j = 1, ..., m$. It follows from Property $UB2$ of Definition 16 that $\nu(ha_1) = \nu(h) \times \sigma(a_1)$, $\nu(ha_1a_2) = \nu(ha_1) \times \sigma(a_2) = \nu(h) \times \sigma(a_1) \times \sigma(a_2)$, and so on. Hence

$$\nu(h') = v(h) \times \sigma(a_1) \times ... \times \sigma(a_m). \tag{24}$$

Since, by hypothesis, $h, h' \in E \cap D_\mu^+$, $\nu_E(h) = \frac{\nu(h)}{\sum_{h' \in E \cap D_\mu^+} \nu(h')}$ and $\nu_E(h') = \frac{\nu(h')}{\sum_{h' \in E \cap D_\mu^+} \nu(h')}$ and thus, using (24) we get that $\nu_E(h') = \frac{\nu(h) \times \sigma(a_1) \times ... \times \sigma(a_m)}{\sum_{h' \in E \cap D_\mu^+} \nu(h')} = \frac{\nu(h)}{\sum_{h' \in E \cap D_\mu^+} \nu(h')} \times \sigma(a_1) \times ... \times \sigma(a_m) = \nu_E(h) \times \sigma(a_1) \times ... \times \sigma(a_m)$. ■

In order to prove Proposition 18 we will make use of a result proved in [19]. First some notation.

If $c : A \to (0, 1]$ and $h = a_1...a_m$ is a history, we denote the product $c(a_1) \times ... \times c(a_m)$ by $\prod_{a \in h} c(a)$ ($a \in h$ means that a is an action that appears in history h). If $e : A \to \mathbb{Z}^- \cup \{0\}$ (where $\mathbb{Z}^-$ denotes the set of negative integers) and $h = a_1...a_m$ is a history, we define $e : H \to \mathbb{Z}^- \cup \{0\}$ by $e(h) = \sum_{a \in h} e(a)$ and if $I(h)$ is the information set containing h, we denote $I_e(h) = \arg\max\{e(h') : h' \in I(h)\} = \{h' \in I(h) : e(h') \geq e(h''), \forall h'' \in I(h)\}$.

The following proposition is proved in [19, Theorem 2.1, p. 11]. A very similar result is provided in [13] and [14, p. 74].[28]

Proposition 23 *Fix an extensive game and let (σ, μ) be an assessment. Then the following are equivalent.*

(A) *There exist two functions $c : A \to (0,1]$ and $e : A \to \mathbb{Z}^- \cup \{0\}$ such that*

(1) for every action a, $e(a) = 0$ if and only if $\sigma(a) > 0$,

(2) for every action a, $\sigma(a) = \begin{cases} c(a) & \text{if } e(a) = 0 \\ 0 & \text{if } e(a) < 0 \end{cases}$, and

(3) for every decision history h, $\mu(h) = \begin{cases} \frac{\prod_{a \in h} c(a)}{\sum_{h' \in I_e(h)} \prod_{a \in h'} c(a)} & \text{if } h \in I_e(h) \\ 0 & \text{if } h \notin I_e(h) \end{cases}$.

(B) *(σ, μ) is KW-consistent.*

We can now proceed to the proof of Proposition 18.

Proof of Proposition 18. $(A) \Rightarrow (B)$. Let (σ, μ) be a perfect Bayesian equilibrium which is rationalized by a choice measurable plausibility order $\precsim$ and is uniformly Bayesian relative to $\precsim$. We need to show that (σ, μ) is a sequential equilibrium. Since (σ, μ) is a perfect Bayesian equilibrium, it is sequentially rational and thus we only need to show that (σ, μ) is KW-consistent. Let $v : D \to (0,1]$ be the function given in Definition 16 and define $c : A \to (0,1]$ as follows:

$$c(a) = \begin{cases} \nu(a) & \text{if } a \in A(\emptyset) \\ \frac{\nu(ha)}{\nu(h)} \text{ for some } h \in H \text{ such that } a \in A(h) & \text{otherwise.} \end{cases} \tag{25}$$

By Property $UB1$, $c(a)$ is well defined for every $a \in A$. Next we show that, for every $a \in A$,

$$\text{if } \sigma(a) > 0 \text{ then } c(a) = \sigma(a). \tag{26}$$

Fix an arbitrary $a \in A$ and let I_a be the information set at which a is available, that is, $I_a = \{h \in D : a \in A(h)\}$ (recall the assumption that no action is available at more than one information set: see Remark 21). Assume that $\sigma(a) > 0$; then, by $P1$ of the definition of AGM-consistency (Definition 2), a is plausibility preserving, that is, $h \sim ha$ for every $h \in I_a$. Hence, by Property $UB2$ of Definition 16, $\nu(ha) = \nu(h) \times \sigma(a)$. Thus $c(a) = \frac{\nu(ha)}{\nu(h)} = \sigma(a)$. Next we show that,

$$\text{for every history } h = a_1 ... a_m, \;\; \nu(h) = \prod\nolimits_{i=1}^{m} c(a_i). \tag{27}$$

We prove this by induction. First note that $\nu(a_1 a_2) = \nu(a_1) \times \frac{\nu(a_1 a_2)}{\nu(a_1)} = c(a_1) \times c(a_2)$ (recall that, by (25), since $a_1 \in A(\emptyset)$, $c(a_1) = \nu(a_1)$). Now suppose that, for $2 \leq k < m$, $v(a_1 a_2 ... a_k) = \prod_{i=1}^{k} c(a_i)$. Then $v(a_1 a_2 ... a_k a_{k+1}) = \frac{v(a_1 a_2 ... a_k a_{k+1})}{v(a_1 a_2 ... a_k)} \times v(a_1 a_2 ... a_k) = c(a_{k+1}) \times \prod_{i=1}^{k} c(a_i) = \prod_{i=1}^{k+1} c(a_i)$.

[28] The "completely mixed pseudo behavioral strategy profile" $\hat{\sigma}$ defined in [13] is essentially equivalent to the function $c : A \to (0,1]$ of the following proposition; furthermore, the logarithms of the mistake probabilities ε in [13] play the same role as the values of the function $e : A \to \mathbb{Z}^- \cup \{0\}$ (the latter are used as exponents of monomials). However, the values of the fuction e are integers, while [13] work with real numbers. It it also worth noting that the proof of Proposition 23 makes use of linear algebra, while [13] uses the separating hyperplane theorem.

Given a history h and the information set $I(h)$ that contains it, let $Min_{\precsim} I(h) = \{h' \in I(h) : h' \precsim h'', \forall h'' \in I(h)\}$. Next we show that, for every decision history h

$$\mu(h) = \begin{cases} \frac{\prod_{a \in h} c(a)}{\sum_{h' \in Min_{\precsim} I(h)} \prod_{a \in h'} c(a)} & \text{if } h \in Min_{\precsim} I(h) \\ 0 & \text{if } h \notin Min_{\precsim} I(h) \end{cases}. \tag{28}$$

Fix an arbitrary decision history h. By $P2$ of the definition of AGM-consistency (Definition 2), $\mu(h) > 0$ (that is, $h \in D_\mu^+$) if and only if $h \in Min_{\precsim} I(h)$, that is, $Min_{\precsim} I(h) = I(h) \cap D_\mu^+$. Thus we only need to show that if $h \in Min_{\precsim} I(h)$ then $\mu(h) = \frac{\prod_{a \in h} c(a)}{\sum_{h' \in Min_{\precsim} I(h)} \prod_{a \in h'} c(a)}$. Suppose that $h \in Min_{\precsim} I(h) = I(h) \cap D_\mu^+$. Let E be the equivalence class of $\precsim$ that contains h; then $Min_{\precsim} I(h) \subseteq E \cap D_\mu^+$. Let $\nu_E(\cdot)$ be the function given in Definition 16; then, for every $h' \in Min_{\precsim} I(h)$,

$$\nu_E(h') = \frac{\nu(h')}{\sum_{h'' \in E \cap D_\mu^+} \nu(h'')}. \tag{29}$$

By Property $UB3$ of Definition 16, for every $h' \in E \cap D_\mu^+$, $\mu(h') = \frac{\nu_E(h')}{\nu_E(I(h))} = \frac{\nu_E(h')}{\nu_E(Min_{\precsim} I(h))}$ and thus, by (29),

$$\forall h' \in Min_{\precsim} I(h), \quad \mu(h') = \frac{\nu(h')}{\nu(Min_{\precsim} I(h))} = \frac{\nu(h')}{\sum_{h' \in Min_{\precsim} I(h)} \nu(h'')}. \tag{30}$$

This, together with (27), yields (28).
By hypothesis $\precsim$ is choice measurable. Let F be a cardinal integer-valued representation of $\precsim$ (see Definition 10) and, for every action a, define $e(a) = F(h) - F(ha)$ for some h such that $a \in A(h)$. Then $e(a) \leq 0$ for all a and $e(a) = 0$ if and only if a is plausibility preserving. Then, letting $e(h) = \sum_{a \in h} e(a)$, it follows that $e(h) = -F(h)$ and thus, given an arbitrary decision history h, $Min_{\precsim} I(h) = \arg\min\{F(h') : h' \in I(h)\} = \arg\max\{e(h') : h' \in I(h)\} = I_e(h)$. It follows from this and (28) that

$$\text{for every decision history } h, \quad \mu(h) = \begin{cases} \frac{\prod_{a \in h} c(a)}{\sum_{h' \in I_e(h)} \prod_{a \in h'} c(a)} & \text{if } h \in I_e(h) \\ 0 & \text{if } h \notin I_e(h) \end{cases}. \tag{31}$$

Thus, by (26) and (31), it follows from Proposition 23 that (σ, μ) is KW-consistent.

$(B) \Rightarrow (A)$. Let (σ, μ) be a sequential equilibrium. Then, by Proposition 14 in [5], (σ, μ) is a perfect Bayesian equilibrium. Thus we only need to show that (σ, μ) is uniformly Bayesian relative to a plausibility order $\precsim$ that rationalizes (σ, μ). By Proposition 23 there exist $c : A \to (0, 1]$ and $e : A \to \mathbb{Z}^- \cup \{0\}$ such that

$$e(a) = 0 \text{ if and only if } \sigma(a) > 0, \tag{32}$$

$$\forall a \in A, \sigma(a) = \begin{cases} c(a) & \text{if } e(a) = 0 \\ 0 & \text{if } e(a) < 0 \end{cases}, \tag{33}$$

and

$$\forall h \in D,\ \mu(h) = \begin{cases} \frac{\prod_{a\in h} c(a)}{\sum_{h'\in I_e(h)} \prod_{a\in h'} c(a)} & \text{if } h \in I_e(h) \\ 0 & \text{if } h \notin I_e(h) \end{cases} \tag{34}$$

where $I_e(h) = \arg\max\{e(h') : h' \in I(h)\} = \{h' \in I(h) : e(h') \geq e(h''), \forall h'' \in I(h)\}$ and $e(h) = \sum_{a\in h} e(a)$. Define the following total pre-order on H:

$$h \precsim h' \text{ if and only if } -e(h) \leq -e(h'). \tag{35}$$

First we show that $\precsim$ rationalizes (σ, μ) (Definition 2). Fix an arbitrary action a and an arbitrary h such that $a \in A(h)$. If $\sigma(a) > 0$ then, by (32), $e(a) = 0$ and thus $e(ha) = e(h) + e(a) = e(h)$ and hence $h \sim ha$; conversely, if $h \sim ha$ then $e(ha) = e(h) + e(a) = e(h)$ and thus $e(a) = 0$, so that, by (32), $\sigma(a) > 0$. Thus $P1$ of Definition 2 is satisfied. Now fix an arbitrary decision history h; then, by (34), $\mu(h) > 0$ if and only if $h \in I_e(h)$ if and only if $-e(h) \leq -e(h')$ for all $h' \in I(h)$ if and only if $h \precsim h'$ for all $h' \in I(h)$. Thus also $P2$ of Definition 2 is satisfied. For every $h \in H$ define $\nu(h) = \prod_{a\in h} c(a)$. We need to show that $\nu(\cdot)$ satisfies the properties of Definition 16. If $h' \in I(h)$ and $a \in A(h)$ then $\frac{\nu(ha)}{v(h)} = \frac{\left(\prod_{a'\in h} c(a')\right)\times c(a)}{\prod_{a'\in h} c(a')} = c(a)$ and, similarly, $\frac{\nu(h'a)}{v(h')} = c(a)$; thus $UB1$ is satisfied. If $h \sim ha$ then $e(a) = 0$ and thus, by (33), $c(a) = \sigma(a)$ so that $\nu(ha) = \nu(h) \times c(a) = \nu(h) \times \sigma(a)$ and hence $UB2$ is also satisfied. Now fix an arbitrary decision history h and let E be the equivalence class to which h belongs; define $\nu_E : H \to [0, 1]$ as follows:

$$\nu_E(h) = \begin{cases} \frac{\nu(h)}{\sum_{h'\in E\cap D_\mu^+} v(h')} & \text{if } h \in E \text{ and } \mu(h) > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{36}$$

By (34)

$$\mu(h) > 0 \text{ if and only if } h \in I_e(h), \text{ that is, if and only if } h \precsim h', \forall h' \in I(h) \tag{37}$$

and if $\mu(h) > 0$ then $\mu(h) = \frac{\prod_{a\in h} c(a)}{\sum_{h'\in I_e(h)} \prod_{a\in h'} c(a)}$, that is, dividing numerator and denominator by $\sum_{h'\in E\cap D_\mu^+} \prod_{a\in h'} c(a)$, $\mu(h) = \frac{\nu_E(h)}{\sum_{h'\in I_e(h)} \nu_E(h')}$. Thus Property $B3$ of Definition 4 holds if $\sum_{h'\in I_e(h)} \nu_E(h') = \sum_{h'\in I(h)} \nu_E(h')$ but this is an immediate consequence of (37), since if $h' \in I(h)\backslash I_e(h)$ then, by (37), $\mu(h') = 0$ and thus, by (36), $\nu_E(h') = 0$. ■

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] Blume, L. and W. R. Zame, The algebraic geometry of perfect and sequential equilibrium, *Econometrica*, 1994, 62: 783-794.

[4] 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.

[5] Bonanno, Giacomo, AGM-consistency and perfect Bayesian equilibrium. Part I: definition and properties, *International Journal of Game Theory*, 2013, 42: 567-592.

[6] Dines, L., On positive solutions of a system of linear equations, *Annals of Mathematics*, 1926-1927, 28: 386-392.

[7] Fishburn, P. C., Finite linear qualitative probability, *Journal of Mathematical Psychology*, 1996, 40: 64-77.

[8] Hendon, Ebbe, Jørgen Jacobsen and Brigitte Sloth, The one-shot-deviation principle for sequential rationality, *Games and Economic Behavior*, 1996, 12: 274-282.

[9] Kohlberg, Elon and Philip Reny, Independence on relative probability spaces and consistent assessments in game trees, *Journal of Economic Theory*, 1997, 75: 280-313.

[10] Kraft, C. H., J. W. Pratt and A. Seidenberg, Intuitive probability on finite sets, *Annals of Mathematical Statistics*, 1959, 30: 408-419.

[11] Kreps, David and Robert Wilson, Sequential equilibrium, *Econometrica,* 1982, 50: 863-894.

[12] Osborne, Martin and Ariel Rubinstein, *A course in game theory,* MIT Press, 1994.

[13] Perea, Andrés, Mathijs Jansen and Hans Peters, Characterization of consistent assessments in extensive-form games, *Games and Economic Behavior*, 1997, 21: 238-252.

[14] Perea, Andrés, *Rationality in extensive form games*, Kluwer Academic Publishers, 2001.

[15] Perea, Andrés, A note on the one-deviation property in extensive form games, *Games and Economic Behavior*, 2002, 40: 322-338.

[16] Schrijver, A., *Theory of linear and integer programming*, John Wiley & Sons, 1986.

[17] Scott, D., Measurement structures and linear inequalities, *Journal of Mathematical Psychology*, 1964, 1: 233-247.

[18] Selten, Reinhard, Re-examination of the perfectness concept for equilibrium points in extensive games, *International Journal of Game Theory*, 1975, 4: 25-55.

[19] Streufert, Peter, Characterizing consistency with monomials, working paper, University of Western Ontario, 2007

[20] Streufert, Peter, Additive plausibility characterizes the supports of consistent assessments, Research Report # 2012-3, University of Western Ontario, 2012.