2012

2012-3 Additive Plausibility Characterizes the Supports of Consistent Assessments

Peter A. Streufert

Additive Plausibility Characterizes the Supports of Consistent Assessments

by

Peter A. Streufert

Research Report # 2012-3 **November 2012**

Department of Economics Research Report Series

Department of Economics
Social Science Centre
The University of Western Ontario
London, Ontario, N6A 5C2
Canada

ADDITIVE PLAUSIBILITY CHARACTERIZES THE SUPPORTS OF CONSISTENT ASSESSMENTS

Peter A. Streufert
Department of Economics
University of Western Ontario

Abstract. We introduce three definitions. First, we let a "basement" be a set of nodes and actions that supports at least one assessment. Second, we derive from an arbitrary basement its implied "plausibility" (i.e. infinite-relative-likelihood) relation among the game's nodes. Third, we say that this plausibility relation is "additive" if it has a completion represented by the nodal sums of a mass function defined over the game's actions. This last construction is built upon Streufert (2012)'s result that nodes can be specified as sets of actions.

Our central result is that a basement has additive plausibility if and only if it supports at least one consistent assessment. The result's proof parallels the early foundations of probability theory and requires only Farkas' Lemma. The result leads to related characterizations, to an easily tested necessary condition for consistency, and to the repair of a nontrivial gap in a proof of Kreps and Wilson (1982).

1. Introduction

Recall that an assessment for an extensive-form game lists both a strategy and a belief for each agent (i.e. information set). The strategy specifies a probability distribution over the agent's actions, and the belief specifies a probability distribution over the agent's nodes

Date: November 16, 2012. *Email:* pstreuf@uwo.ca. *JEL classification:* C72.
Keywords: plausibility relation, additive representation, plausibility mass function, infinite relative likelihood.

I thank Walter Bossert, Audra Bowlus, Jean Guillaume Forand, Ian King, Val Lambson, Bart Lipman, Andrew McLennan, Andrés Perea, Al Slivinski, Bernhard von Stengel, Bob Wilson, and Charles Zheng; seminar participants at Waterloo and Western Ontario; and conference participants at the Canadian Economics Association annual meeting in Calgary, the Midwest Economic Theory conference in Bloomington, the game theory conference in Stony Brook, and the NBER/NSF/CEME conference in Bloomington.

(i.e. the information set's nodes). In many equilibrium concepts, the agent chooses her[1] strategy optimally, given her belief, the strategies of subsequent agents, and her own payoffs at the game's terminal nodes.

The assessment's strategies determine the probability of reaching any node in the game tree. If every strategy has full support (i.e. plays every action with positive probability), then every node in the tree is reached with positive probability. In this case, it is natural to assume that every agent (i.e. every information set) calculates its belief over its own nodes by applying the conditional-probability law (sometimes known as Bayes Rule).

However, if some strategies do not have full support, some nodes are reached with zero probability. In such a case, there may be agents that consist entirely of zero-probability nodes. Then, since the conditional-probability law cannot be applied, it is not at all obvious how to model the reasoning by which a zero-probability agent calculates its belief. Rather, we must awkwardly grapple with how to model the reasoning of an agent who unexpectedly discovers that one of its unreachable nodes has actually been reached.

This difficult issue can be usefully divided into two sub-issues: (a) modeling how a zero-probability agent calculates the support of its belief, and (b) modeling how the agent calculates a probability distribution over this support. In this paper, we are exclusively concerned with sub-issue (a).

Economists commonly address both of these sub-issues through the concept of consistency that was introduced in the path-breaking work of Kreps and Wilson (1982). Essentially, this concept imagines that a zero-probability agent calculates its belief by a three-stage process: (1) for each of the other agents it posits some sequence of full-support strategies that converges to that agent's actual strategy, (2) for each strategy profile in this sequence, it calculates its own belief by means of the conditional-probability law, and (3) it takes the topological limit of this sequence of beliefs. This definition is very natural because it states that what we do not understand (namely, how a zero-probability agent calculates its belief) must be near to what we do understand (namely, the conditional-probability law).

[1]Since we identify an agent with an information set, future paragraphs will refer to an agent with the pronoun "it" rather than "he" or "she".

Although the topological definition of consistency is natural, it is both conceptually interesting and computationally useful to understand consistency without reference to a converging sequence of full-support assessments. Indeed, most of the papers cited below contribute to this agenda. Our paper contributes to the agenda of understanding consistency without sequences by providing a new alternative characterization of the supports of consistent assessments.

Our new characterization uses three new definitions. The first definition is a small preliminary step. Recall that an assessment lists a strategy and a belief for each agent. Since a strategy's support consists of actions and a belief's support consists of nodes, it must be that an assessment's support consists of both actions and nodes. We define a "basement" to be a set of actions and nodes that supports at least one assessment. Our task is to characterize those basements that support at least one consistent assessment.

Second, we derive from an arbitrary basement its implied "plausibility" (i.e. infinite-relative-likelihood) relation $\succcurlyeq$. This binary relation compares nodes, but does so only in two circumstances. (1) Suppose one node immediately precedes another. Then the first is more plausible (i.e. infinitely more likely) than the second if the intervening action is outside the basement (i.e. is played with zero probability). Further the two nodes are equally plausible (i.e. neither can be infinitely more likely than the other) if the intervening action is in the basement (i.e. is played with positive probability). (2) Suppose two nodes belong to the same agent. Then the two are equally plausible if both are in the basement (i.e. both are believed with positive probability), and the first is more plausible than the second if the first is in the basement while the second is not.

This second definition is a minor departure from the literature. In the relevant literature, infinite relative likelihoods are derived indirectly from the rich probability structures that are derived from consistent assessments. These rich probability structures include the conditional probability systems of Myerson (1986), the logarithmic likelihood ratios of McLennan (1989), the lexicographic probability systems of Blume, Brandenburger, and Dekel (1991), the nonstandard probability systems of Hammond (1994), and the relative probability systems of Kohlberg and Reny (1997). In contrast, our relation $\succcurlyeq$ is derived directly from an arbitrary basement, which may or may not support a consistent assessment. Accordingly, their infinite relative likelihoods are complete and transitive, while our relation $\succcurlyeq$ is neither complete nor transitive.

Third, we introduce the idea of additively representing an ordering of the game's nodes. Since a node can be understood as a set of actions by Streufert (2012, Theorem 1), one can construct a representation of an ordering of the nodes simply by (a) assigning a number to each action and then (b) summing these numbers over the actions in each node. The assignment of numbers to actions in step (a) is naturally called a "mass function", and the nodal sums in step (b) are naturally called an "additive" representation of an ordering of the game's nodes. Accordingly, a basement is said to have "additive plausibility" if its plausibility relation $\succcurlyeq$ has a completion with an additive representation.

This paper proves that a basement has additive plausibility if and only if it supports at least one consistent assessment. In other words, it shows that additive plausibility characterizes the supports of consistent assessments. This new characterization is our main result, as reflected in the title of our paper.

Although it involves a long chain of definitions, additive plausibility is easy to interpret. First, one naturally calls the additive representation's mass function a "plausibility" mass function. Then, representation and part (1) in the above definition of a plausibility relation $\succcurlyeq$ together require that the plausibility mass function gives each positive-probability action zero plausibility, and further, that it gives each zero-probability action negative plausibility. Accordingly, a node's plausibility (that is, the sum of the plausibility numbers of a node's actions) is a measure of how far the node lies below the so-called "equilibrium path". It is slightly more sophisticated than (the negative of) the number of zero-probability actions leading to the node (i.e. the number of zero-probability actions in the node) because the mass function can give each zero-probability action its own negative plausibility number.

Broadly speaking, the definition of consistency has been easier shown to hold than shown to fail. This difference has arisen because it is easier to find an assessment sequence with certain properties than to show that such a sequence cannot exist. Therefore, new necessary conditions for consistency are particularly helpful.

Accordingly, the necessity of additive plausibility is the interesting half of our characterization of the supports of consistent assessments. Further, our proof of its necessity is unexpectedly straightforward. It requires nothing more than Farkas' Lemma, and it parallels the classic foundations of ordinary probability theory over a finite state space.

To see this parallel, think of an action as a state. Then, a node, which is a set of actions, resembles an event, which is a set of states. Further, our plausibility (i.e. infinite-relative-likelihood) relation $\succcurlyeq$, which conveys when one node is at least as plausible as another, resembles a so-called "probability relation", which conveys when one event is at least as probable as another. Consequently, our showing that a certain relation $\succcurlyeq$ has a completion with an additive representation parallels classic results showing that a certain probability relation has a completion with an additive representation. These classic results appear in Kraft, Pratt, and Seidenberg (1959) and Scott (1964). In this fashion, the additivity of our plausibility representations parallels the additivity of classic probability representations (although, as mentioned earlier, our plausibility numbers are nonpositive, while probability numbers are nonnegative and restricted to the unit interval).

Two less central results can now be discussed naturally. First, if the relation $\succcurlyeq$ has a completion with an additive representation, then it must necessarily have a transitive completion. Accordingly, we show that the existence of a transitive completion for the plausibility relation is a necessary condition for consistency which is very easily tested (but weaker than additive plausibility). Second, classic probability theory shows that the so-called "cancellation law" is equivalent to having a completion with an additive representation. Similarly, we show that this cancellation law provides a second full characterization of the supports of consistent assessments.

Finally, we discuss three other characterizations of the supports of consistent assessments. These other characterizations do not refer to a plausibility relation $\succcurlyeq$, but instead, directly link a mass function with a basement. These characterizations (of merely the supports of consistent assessments) are related to the characterizations of consistency developed by Kreps and Wilson (1982, Appendix) and Perea y

Monsuwé, Jansen, and Peters (1997). We believe that their characterizations should be much better understood and appreciated for their insights.

Our introduction of plausibility relations and plausibility mass functions serves to correct, simplify, unify, and extend this important literature. In particular, we fill a critical gap in a proof of Kreps and Wilson (1982), we substantially simplify the proof of Perea y Monsuwé, Jansen, and Peters (1997), and we show how these two papers are related. In addition, we extend their analyses to accommodate arbitrary chance players, and to allow non-absent-minded strategic players who violate perfect recall.

The paper is organized as follows. Section 2 recapitulates old concepts, while Section 3 defines new concepts. Section 4 contains Theorem 1, which re-uses classic probability theory to show that additive plausibility is necessary for consistency. Section 5 contains Theorem 2, which incorporates both Theorem 1 and its converse. The section also discusses related characterizations of the supports of consistent assessments, and repairs the gap in Kreps and Wilson (1982). Section 6 concludes.

2. Old Definitions

2.1. Reviewing set-tree games

This subsection specifies a finite game via the set-tree formulation of Streufert (2012). We choose this formulation because it specifies nodes as sets of actions, and because we will use the nodal sums of a mass function defined over actions to define an additive representation of an ordering over nodes. We follow Kreps and Wilson (1982) in restricting ourselves to games with a finite number of actions, and thereby forgo the full generality of the arbitrary finite-horizon games considered by Streufert (2012).

By Streufert (2012, Theorem 1), the set-tree formulation is less general than the standard formulation of Osborne and Rubinstein (1994, page 200) in only one respect: it implicitly rules out agents that are absent-minded in the sense of Piccione and Rubinstein (1997). Streufert (2012) calls the absence of absent-minded agents "agent recall", and notes that agent recall is weaker than perfect recall. Perfect recall has been a standard assumption in the consistency literature ever since Kreps and Wilson (1982, pages 863 and 867).

We now recapitulate the formal definition of a set-tree game from Streufert (2012). That paper's Subsection 2.2 defines a *set tree* (A,T) to be a set A of *actions* a and a collection T of subsets of A such that (a) $A = \bigcup T$, (b) $|T| \geq 2$, and (c) every nonempty $t \in T$ contains exactly one action whose removal results in another element of T. An element t of T is called a *node*. Streufert (2012) assumes that each node t is a finite subset of A, and this paper imposes the additional restriction that A itself is finite.

A set tree (A,T) determines the feasibility correspondence $F{:}T \twoheadrightarrow A$ by $(\forall t)\ F(t) = \{a{\notin}t|t{\cup}\{a\}{\in}T\}$, and also determines the set of terminal nodes by $Z = \{t|F(t){=}\varnothing\}$. Note that $F(t)$ is the set of actions that are feasible (i.e available) from node t, and that we use the symbol $F(t)$ rather than the more common $A(t)$.

Further recall that a *set-tree game* (A,T,H,I,i^c,ρ,u) is a set tree (A,T) together with five additional objects. (1) H is a collection of *agents* (i.e. information sets) h which partition the set $T{\smallsetminus}Z$ of nonterminal nodes. (2) I is a collection of *players* i which prepartition[2] H. (3) $i^c \in I$ is a possibly empty *chance* player. (4) $\rho : \bigcup_{h\in i^c} F(h) \rightarrow (0,1]$ assigns a positive probability to each chance action $a{\in}\bigcup_{h\in i^c} F(h)$. (5) $u : (I{\smallsetminus}\{i^c\}){\times}Z \rightarrow \mathbb{R}$ specifies a payoff $u_i(t)$ to each nonchance player $i{\in}I{\smallsetminus}\{i^c\}$ at each terminal node $t{\in}Z$. The agents are assumed to satisfy

$$\text{(1a)} \qquad (\forall t^1,t^2)\ \ [(\exists h)\{t^1,t^2\}{\subseteq}h] \ \Rightarrow\ F(t^1){=}F(t^2) \ \text{ and}$$

$$\text{(1b)} \qquad (\forall t^1,t^2)\ \ [(\nexists h)\{t^1,t^2\}{\subseteq}h] \ \Rightarrow\ F(t^1){\cap}F(t^2){=}\varnothing\ .$$

The first assumption is standard: it requires that the same actions are feasible from any two nodes in an agent (i.e. information set). The second requires that actions are *agent-specific* in the sense that nodes from different agents have different actions. This entails no loss of generality because one can always introduce enough actions so that agents never share actions. Further, the chance probabilities are assumed to satisfy $(\forall h{\in}i^c)\ \Sigma_{a\in F(h)}\rho(a) = 1$ so that they specify a probability distribution at each chance agent $h{\in}i^c$. Finally, without loss of generality,

[2] A prepartition of H is a collection of sets whose nonempty members partition H. Thus $\varnothing$ can be a member of a prepartition. Admitting $\varnothing$ allows one to specify a game "without chance" by setting $i^c = \varnothing$.

every nonchance player is assumed to be nonempty. All of the above are discussed in Streufert (2012, Subsections 2.1 and 2.2).

2.2. Reviewing Kreps-Wilson consistency

This subsection reformulates Kreps-Wilson consistency in terms of a set-tree game. We must provide full details because this paper is the first to specify strategies and beliefs in terms of a set-tree game.

First, we introduce notation that divides the nodes and actions into those of the chance player and those of the *strategic* (i.e. nonchance) player(s). Assume that there is at least one strategic player. Then, since the set H of agents is prepartitioned by the set I of players, we can prepartition H into the possibly empty set i^c of chance agents and the necessarily nonempty set $\bigcup(I \smallsetminus \{i^c\})$ of strategic agents. Let H^s denote $\bigcup(I \smallsetminus \{i^c\})$. Then since the set $T \smallsetminus Z$ of nonterminal nodes is partitioned by H, we can prepartition $T \smallsetminus Z$ into the set T^c of chance nodes and the set T^s of strategic (i.e. decision) nodes:

$$T^c = \bigcup_{h \in i^c} h \text{ and } T^s = \bigcup_{h \in H^s} h \ .$$

Similarly, since the set A of actions has the indexed partition $\langle F(h) \rangle_h$ by Streufert (2012, Lemma A.2), we can prepartition A into the set A^c of chance actions and the set A^s of strategic actions:

$$A^c = \bigcup_{h \in i^c} F(h) \text{ and } A^s = \bigcup_{h \in H^s} F(h) \ .$$

Note that T^c, T^s, A^c, and A^s are derived from the given game. Further note that the definition of A^c allows us to write the chance probabilities as $\rho{:}A^c \rightarrow (0,1]$, rather than $\rho{:}\bigcup_{h \in i^c} F(h) \rightarrow (0,1]$ as was done in part (4) of the above definition of a game. As one would expect, it can be shown that $A^c = \bigcup_{t \in T^c} F(t)$ and $A^s = \bigcup_{t \in T^s} F(t)$ (Lemma A.1 in the appendix).

Second, we introduce notation for strategies, beliefs, and assessments. A (behavioural) *strategy profile* is a function $\sigma{:}A^s \rightarrow [0,1]$ such that $(\forall h \in H^s)\ \Sigma_{a \in F(h)} \sigma(a) = 1$. Thus a strategy profile specifies a probability distribution $\sigma|_{F(h)}$ over the feasible set $F(h)$ of each strategic agent h. This $\sigma|_{F(h)}$ is h's strategy. A *belief system* is a function $\beta{:}T^s \rightarrow [0,1]$ such that $(\forall h \in H^s)\ \Sigma_{t \in h} \beta(t) = 1$. Thus a belief system specifies a probability distribution $\beta|_h$ over each strategic agent h. This $\beta|_h$ is h's belief. Finally, an *assessment* (σ, β) consists of a strategy profile σ and a belief system β.

Third, an assessment (σ,β) is *full-support Bayesian* if σ assumes only positive values and

$$(2) \qquad (\forall h{\in}H^s)(\forall t{\in}h) \;\; \beta(t) = \frac{\Pi_{a\in t}(\rho{\cup}\sigma)(a)}{\Sigma_{t'\in h}\Pi_{a\in t'}(\rho{\cup}\sigma)(a)} \;.$$

This equation calculates the belief $\beta|_h$ over any strategic agent h by means of the conditional probability law. Here $\rho{\cup}\sigma$ is the union of the functions ρ and σ. In particular, $\rho{\cup}\sigma{:}A{\rightarrow}[0,1]$ since (a) $\rho{:}A^c{\rightarrow}[0,1]$, (b) $\sigma{:}A^s{\rightarrow}[0,1]$, and (c) $\{A^c, A^s\}$ prepartitions A. Thus

$$\Pi_{a\in t}(\rho{\cup}\sigma)(a) = \Pi_{a\in t\cap A^c}\rho(a) \times \Pi_{a\in t\cap A^s}\sigma(a) \;,$$

where the right-hand side multiplies the probabilities of the chance actions leading to t (i.e. the actions in $t{\cap}A^c$) together with the probabilities of the strategic actions leading to t (i.e. the actions in $t{\cap}A^s$). Hence $\Pi_{a\in t}(\rho{\cup}\sigma)(a)$ is the probability of reaching node t. The denominator in (2) is positive because ρ has positive values by the definition of a game, and because σ has positive values by the definition of a full-support Bayesian assessment.

Finally, an assessment is *(Kreps-Wilson) consistent* if it is the limit of a sequence of full-support Bayesian assessments.

2.3. Reviewing the Support of an Assessment

As in Kreps and Wilson (1982), let the *support* of an assessment (σ,β) be the union of the support of $\sigma{:}A^s{\rightarrow}[0,1]$ and the support of $\beta{:}T^s{\rightarrow}[0,1]$. Note that the support of an assessment is a subset of $A^s{\cup}T^s$, and accordingly, it consists of both actions and nodes.

Figure 1 provides an example. Its game tree is essentially that of Kreps and Ramey (1987, Figure 1). A casual interpretation of this game tree might be that you manage two workers, that each has a switch, and that a lamp turns on exactly when both switches are on. You can observe the lamp but not the switches, and then if the lamp is dark, you can choose to penalize either the first worker or the second worker.

The figure also specifies an assessment: the strategy profile σ is given by the numbers without boxes and the belief system β is given by the numbers within boxes. Casually, this assessment might describe an equilibrium-like situation in which both workers work because (a) they think that if the light is dark, you would place probability 0.2 on both workers dozing, probability 0.4 on only the first worker dozing, and

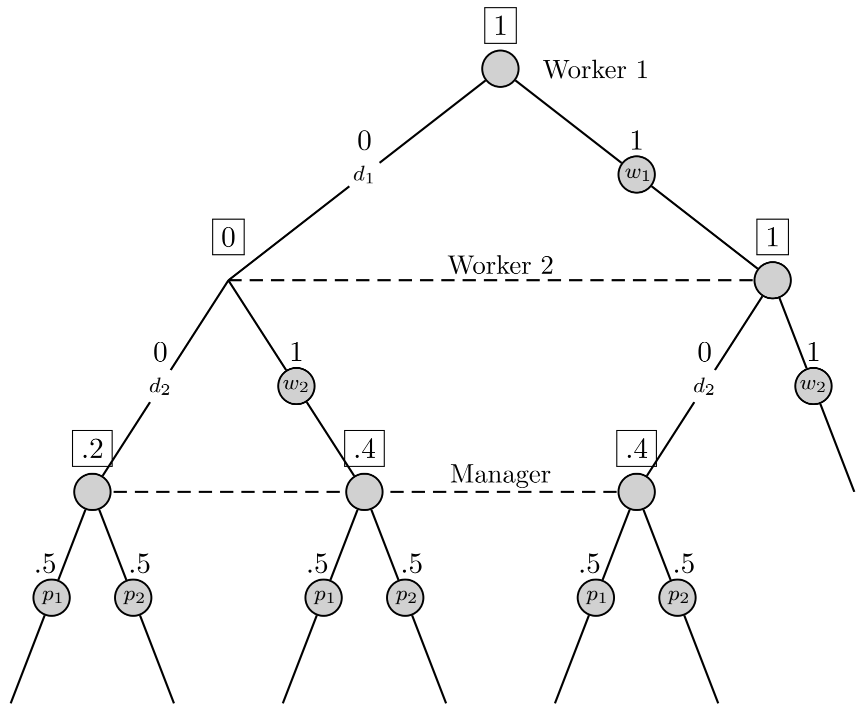


FIGURE 1. An assessment and its support.

probability 0.4 on only the second worker dozing, (b) they see that this belief would induce you to randomize between the two punishments, and (c) the threat of this randomized penalty motivates them both to work.

Finally, the figure shows this assessment's support: the actions and nodes in the support are encircled. In particular, the actions in the support are w_1, w_2, p_1, and p_2, and the nodes in the support are $\{\}$, $\{w_1\}$, $\{d_1, d_2\}$, $\{d_1, w_2\}$, and $\{w_1, d_2\}$.

This paper focuses on the support of an assessment. In particular, it focuses on what the consistency of a given assessment implies about the support of that assessment. Accordingly, it is unconcerned with the magnitudes of that assessment's positive probabilities.[3]

[3]To be clear, this paper considers the magnitudes of positive probabilities only when referring to the full-support assessments in the above definition of consistency. This occurs only in the proofs of Lemmas 4.2 and A.7 (the proof of Lemma 4.2 incorporates Lemma A.6).

3. New Definitions

3.1. Basements

As we have seen, the support of an assessment is a subset of $A^s \cup T^s$. Let a *basement*[4] b be a subset of $A^s \cup T^s$ such that

$$(\forall h \in H^s)\ F(h) \cap b \neq \varnothing \text{ and } h \cap b \neq \varnothing\ .$$

Lemma A.2 in the appendix shows that a subset of $A^s \cap T^s$ supports at least one assessment iff it is a basement. It does this by noting that an arbitrary subset $\bar{b}$ of $A^s \cap T^s$ supports (σ, β) iff, for every agent $h \in H^s$, $F(h) \cap \bar{b}$ supports the agent's strategy $\sigma|_{F(h)}$, and $h \cap \bar{b}$ supports the agent's belief $\beta|_h$.

For example, the encircled actions and nodes in Figure 2 constitute a basement. That basement supports many assessments, including the assessment of Figure 1.

3.2. The Plausibility Relation $\succcurlyeq$ of a Basement b

This paragraph mechanically defines the plausibility relation $\succcurlyeq$ of an arbitrary basement b. We will interpret $\succcurlyeq$ immediately after Lemma 3.1. Accordingly, for any basement b, define the five binary relations

$$\begin{aligned}
\succ_{A^s} &= \{\ (t, t \cup \{a\}) \mid a \in F(t) \text{ and } a \in A^s \smallsetminus b\ \}\ , \\
\approx_{A^s} &= \{\ (t, t \cup \{a\}) \mid a \in F(t) \text{ and } a \in A^s \cap b\ \} \\
&\quad \cup \{\ (t \cup \{a\}, t) \mid a \in F(t) \text{ and } a \in A^s \cap b\ \}\ , \\
\approx_{A^c} &= \{\ (t, t \cup \{a\}) \mid a \in F(t) \text{ and } a \in A^c\ \} \\
&\quad \cup \{\ (t \cup \{a\}, t) \mid a \in F(t) \text{ and } a \in A^c\ \}\ , \\
\succ_{T^s} &= \{\ (t^1, t^2) \mid (\exists h \in H^s)\ t^1 \in h \cap b \text{ and } t^2 \in h \smallsetminus b\ \}\ , \quad \text{and} \\
\approx_{T^s} &= \{\ (t^1, t^2) \mid (\exists h \in H^s)\ \{t^1, t^2\} \subseteq h \cap b \text{ and } t^1 \neq t^2\ \}\ .
\end{aligned}$$

Let $\succ$ be the union of $\succ_{A^s}$ and $\succ_{T^s}$. Let $\approx$ be the union of $\approx_{A^s}$, $\approx_{A^c}$, and $\approx_{T^s}$. And finally, let $\succcurlyeq$ be the union of $\succ$ and $\approx$. We call this $\succcurlyeq$ the *plausibility relation* of b.

Lemma 3.1. *Take any basement b, and derive $\succ_{A^s}$, $\approx_{A^s}$, $\approx_{A^c}$, $\succ_{T^s}$, $\approx_{T^s}$, $\succ$, $\approx$, and $\succcurlyeq$. Then $\succ$ is the asymmetric part of $\succcurlyeq$, and $\succ$ is*

[4]Closely related is the Kreps and Wilson (1982, page 880) concept of "basis", which is defined to be an arbitrary subset of $A^s \cup T^s$.

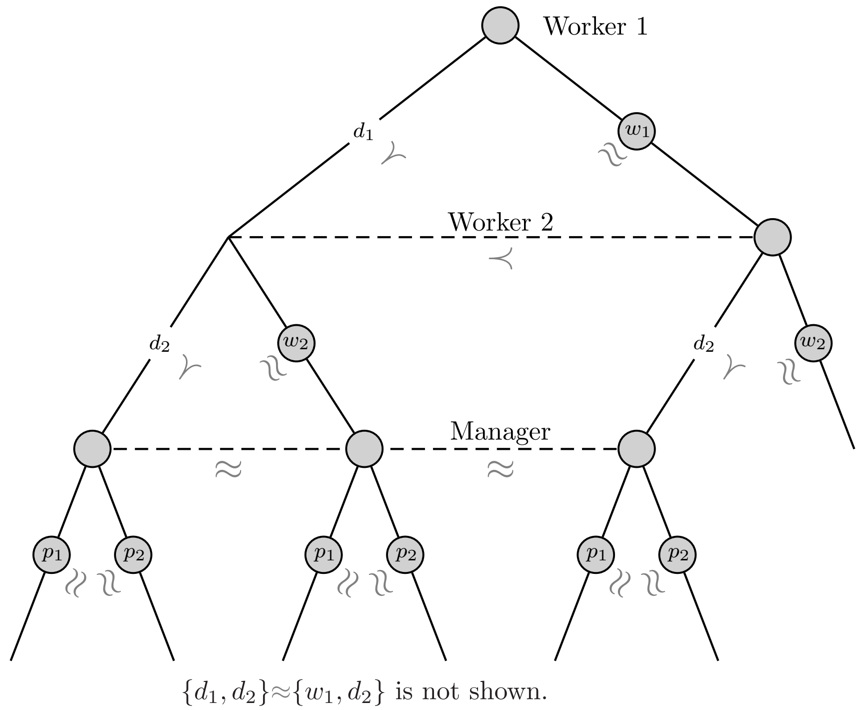


Figure 2. A basement b and its plausibility relation $\succcurlyeq$. This $\succcurlyeq$ does not have a transitive completion.

partitioned by $\{\ \succ_{A^s}, \succ_{T^s}\ \}$*. Similarly,* $\approx$ *is the symmetric part of* $\succcurlyeq$*, and* $\approx$ *is partitioned by* $\{\ \approx_{A^s}, \approx_{A^c}, \approx_{T^s}\ \}$*. (Proof A.3 in the appendix.)*

We now begin to interpret $\succcurlyeq$. This paragraph makes a few initial observations. Each of the above relations compares nodes. The definitions of $\succ_{A^s}$ and $\approx_{A^s}$ concern whether certain actions belong to the basement b. The definitions of $\succ_{T^s}$ and $\approx_{T^s}$ concern whether certain nodes belong to the basement b. The definition of $\approx_{A^c}$ is unconcerned with b.

The gist of $\succcurlyeq$'s interpretation is this. The asymmetric relation $\succ$ contains a given pair of nodes precisely when the basement dictates that the first node is more plausible (i.e. infinitely more likely) than the second. Meanwhile, the symmetric relation $\approx$ contains a given pair of nodes precisely when the basement dictates that neither of the nodes can be more plausible (i.e. infinitely more likely) than the other. We have introduced the word "plausibility" in lieu of the familiar phrase "infinite relative likelihood" only because it is grammatically more convenient.

This and the next four paragraphs discuss each of the five components of $\succcurlyeq$ in detail. First consider $\succ_{A^s}$. As with any relation, the notations $(t^1,t^2) \in \succ_{A^s}$ and $t^1 \succ_{A^s} t^2$ are equivalent. Thus $t^1 \succ_{A^s} t^2$ iff node t^1 immediately precedes node t^2 and the strategic action leading from t^1 to t^2 is not in b (i.e. is played by b with zero probability). For example, Figure 2 shows that $\{\} \succ_{A^s} \{d_1\}$, that is, that the origin $\{\}$ is more plausible than the node $\{d_1\}$ following action d_1. This holds because the intervening action d_1 is not in b (i.e. is played by b with zero probability). Similarly, $\{d_1\} \succ_{A^s} \{d_1, d_2\}$ and $\{w_1\} \succ_{A^s} \{w_1, d_2\}$ (as in any set tree, a node is identified with the set of actions leading to it).

Second, consider $\approx_{A^s}$. Both $t^1 \approx_{A^s} t^2$ and $t^2 \approx_{A^s} t^1$ hold if t^1 immediately precedes t^2 and the strategic action leading from t^1 to t^2 is in b (i.e. is played by b with positive probability). In such a case, we say that t^1 and t^2 are "equally plausible" in the sense that neither can be infinitely more likely than the other. For example, Figure 2 shows that $\{\} \approx_{A^s} \{w_1\}$, that $\{w_1\} \approx_{A^s} \{w_1, w_2\}$, that $\{d_1\} \approx_{A^s} \{d_1, w_2\}$, that $\{d_1, d_2\} \approx_{A^s} \{d_1, d_2, p_1\}$, and that $\approx_{A^s}$ contains five other pairs which end in terminal nodes like the last one listed. (The converses of these nine pairs are also in $\approx_{A^s}$ because $\approx_{A^s}$ was defined to be symmetric.)

Third, this notion of being "equally plausible" applies not only to strategic actions, but also to chance actions, which are played with positive probability by assumption. Accordingly, the definition of $\approx_{A^c}$ states that both $t^1 \approx_{A^c} t^2$ and $t^2 \approx_{A^c} t^1$ hold if a chance action leads from t^1 to t^2. Unlike the other four components of $\succcurlyeq$, this $\approx_{A^c}$ depends only on the game tree and not on the basement. (Figure 2's game tree has no chance actions, and hence $\approx_{A^c}$ is empty.)

Fourth, the definition of $\succ_{T^s}$ states that a node in the support of an agent's belief is more plausible than any node of the agent outside the support. For example, Figure 2 shows $\{w_1\} \succ_{T^s} \{d_1\}$.

Fifth, the definition of $\approx_{T^s}$ states that a node in the support of an agent's belief is tied in plausibility with any other node inside that support. For example, Figure 2 shows $\{d_1, d_2\} \approx_{T^s} \{d_1, w_2\}$ and $\{d_1, w_2\} \approx_{T^s} \{w_1, d_2\}$. ($\approx_{T^s}$ also contains $(\{d_1, d_2\}, \{w_1, d_2\})$, and because the relation is symmetric, the converses of the three pairs already mentioned.)

The typical $\succcurlyeq$ is pervasively incomplete in the sense that it fails to compare many pairs of nodes. For instance, in Figure 2's example, neither $\{\} \succcurlyeq \{d_1, d_2\}$ nor $\{d_1, d_2\} \succcurlyeq \{\}$. In general, if $|T|{\geq}3$, there must be two (distinct) nodes that cannot be compared by the plausibility relation $\succcurlyeq$ of any basement. To see this, note that if $|T|{\geq}3$, then there must be two nodes such that (a) neither is an immediate predecessor of the other, and (b) the two are not in the same agent. By (a) the two nodes cannot be ordered into a pair belonging to $\succ_{A^s}$, $\approx_{A^s}$, or $\approx_{A^c}$. By (b) the two nodes cannot be ordered into a pair belonging to $\succ_{T^s}$ or $\approx_{T^s}$. Hence the two cannot be ordered into a pair belonging to $\succcurlyeq$.

Further, the typical $\succcurlyeq$ is also intransitive. This accords with its incompleteness. For instance, in Figure 2's example, transitivity is violated by the lack of $\{\} \succcurlyeq \{d_1, d_2\}$. In general, such an intransitivity must occur whenever one agent follows another, regardless of the basement under consideration.

As mentioned in the introduction, our relation $\succcurlyeq$ is a minor departure from the literature. There, infinite relative likelihoods are derived *indirectly* from the rich probability structures that are derived from *consistent* assessments. Such rich probability structures include (1) the conditional probability systems of Myerson (1986), which are built on the mathematical foundations of Rényi (1955); (2) the logarithmic likelihood ratios of McLennan (1989); (3) the lexicographic probability systems of Blume, Brandenburger, and Dekel (1991); (4) the nonstandard probability systems of Hammond (1994) and Halpern (2010), which are built on the mathematical foundations of Robinson (1973); and (5) the relative probability systems of Kohlberg and Reny (1997).

In contrast, our $\succcurlyeq$ is derived *directly* from an *arbitrary* basement, which may or may not be the support of a consistent assessment. Accordingly, there are at least three differences. (a) Our $\succcurlyeq$ does not assume consistency while their constructions do. (b) Our $\succcurlyeq$ is easier to derive because its definition does not require their rich probability structures. (c) Our $\succcurlyeq$ is incomplete and intransitive while their infinite relative likelihoods are complete and transitive.[5]

[5] A further difference is that our (incomplete) $\succcurlyeq$ is uniquely determined by the assessment, while in their frameworks, one assessment can lead to several probability systems, each with its own (complete) infinite relative likelihoods. This multiplicity of probability systems appears to correspond with the possibility of our plausibility relation having multiple completions.

3.3. Introducing mass functions μ

As we have seen, a typical plausibility relation $\succcurlyeq$ is incomplete. A *completion* of $\succcurlyeq$ is a complete extension of $\succcurlyeq$. In other words, a complete $\succcurlyeq^*$ is a completion of $\succcurlyeq$ if for all t^1 and t^2

$$t^1 \succ t^2 \;\Rightarrow\; t^1 \succ^* t^2 \text{ and}$$
$$t^1 \approx t^2 \;\Rightarrow\; t^1 \approx^* t^2 \;,$$

where $\succ^*$ and $\approx^*$ are the asymmetric and symmetric parts of $\succcurlyeq^*$. Any $\succcurlyeq$ has at least one (possibly intransitive) completion.

Since T is finite, the existence of a transitive completion is equivalent to the existence of a function $\varphi{:}T{\to}\mathbb{R}$ which *represents*[6] *a completion of* $\succcurlyeq$ in the sense that for all t^1 and t^2

$$\begin{aligned} t^1 \succ t^2 \;&\Rightarrow\; \varphi(t^1) > \varphi(t^2) \text{ and} \\ t^1 \approx t^2 \;&\Rightarrow\; \varphi(t^1) = \varphi(t^2) \;. \end{aligned} \tag{3}$$

For example, Figure 2's plausibility relation does not have a transitive completion because $\{d_1\} \succ \{d_1, d_2\}$ and yet $\{d_1\} \approx \{d_1, w_2\} \approx \{d_1, d_2\}$. Accordingly, that figure's plausibility relation does not have a completion that can be represented by a φ.

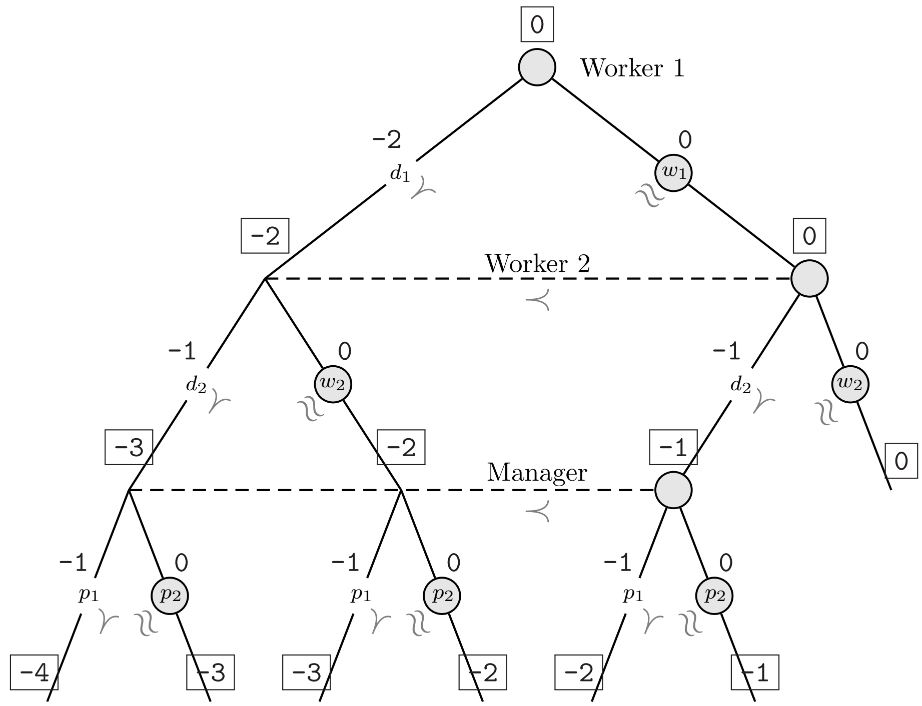
Stronger than the existence of a transitive completion would be the existence of a function $\mu{:}A{\to}\mathbb{R}$ whose nodal sums represent a completion of $\succcurlyeq$ in the sense that for all t^1 and t^2

$$\begin{aligned} t^1 \succ t^2 \;&\Rightarrow\; \Sigma_{a\in t^1}\mu(a) > \Sigma_{a\in t^2}\mu(a) \text{ and} \\ t^1 \approx t^2 \;&\Rightarrow\; \Sigma_{a\in t^1}\mu(a) = \Sigma_{a\in t^2}\mu(a) \;. \end{aligned} \tag{4}$$

This is stronger than the existence of a transitive completion because (4) implies that (3) holds with the special functional form $\varphi(t) = \Sigma_{a\in t}\mu(a)$. For brevity, we will often omit mentioning the nodal sums $\Sigma_{a\in t}\mu(a)$ and simply say that such a μ *additively represents a completion of* $\succcurlyeq$.

For example, Figure 3's plausibility relation $\succcurlyeq$ has a completion that is represented by the nodal sums of a function μ. Such a function $\mu{:}A{\to}\mathbb{R}$ is given by the numbers without boxes that appear over the

[6] We use the term "represent" as it is used in standard consumer theory. In contrast, much of the bibliography's non-economics literature would use "represent" to mean our "represent a completion of". If we were studying complete relations, these two meanings of "represent" would be equivalent since the only completion of a complete relation is the relation itself. Interestingly, a typical plausibility relation is incomplete, and thus our choice of terminology is substantial.



$\{d_1, d_2\} \prec \{w_1, d_2\}$ is not shown.

FIGURE 3. A basement b, its plausibility relation $\succcurlyeq$, and a mass function μ whose nodal sums (which appear in the boxes) represent a completion of $\succcurlyeq$.

actions, and its nodal sums $\Sigma_{a \in t} \mu(a)$ are given by the numbers within boxes that appear over the nodes. If brevity were important, as it often is, we would omit mentioning the nodal sums, and simply say that the figure's μ additively represents a completion of the figure's $\succcurlyeq$.

Because of its resemblance to a probability mass function, we call a function $\mu{:}A \to \mathbb{R}$ a *mass function*. To explore this resemblance in detail, suppose that $p{:}\Omega \to [0, 1]$ is a probability mass function defined on a finite set Ω of states ω (sometimes p is called a discrete probability "density" function). Then, as we all know, the probability of any event $e \subseteq \Omega$ can be calculated by the sum $\Sigma_{\omega \in e} p(\omega)$. Analogously, (1) an action $a \in A$ is like a state $\omega \in \Omega$, (2) a mass function $\mu{:}A \to \mathbb{R}$ is like a probability mass function $p{:}\Omega \to [0, 1]$, (3) a node $t \subseteq A$ is like an event $e \subseteq \Omega$, and (4) a nodal sum $\Sigma_{a \in t} \mu(a)$ is like a sum of the form $\Sigma_{\omega \in e} p(\omega)$. Although this analogy is quite useful, it is imperfect in the sense that we do not require that the values of a mass function μ be nonnegative or that they sum to one.

It is natural to call μ a *plausibility* mass function, to call $\mu(a)$ the *plausibility* of the action a, and to call $\Sigma_{a \in t}\mu(a)$ the *plausibility* of the node t. Further, it is natural to say that a basement has *additive plausibility* if its plausibility relation has a completion with an additive representation.

4. The Necessity of Additive Plausibility

4.1. Re-using classic probability theory

Theorem 1. *Suppose an assessment is consistent. Then its support's plausibility relation* $\succcurlyeq$ *has a completion represented additively by a mass function* μ*. Further, the mass function* μ *can be made to assume integer values. (Proof 4.3 below.)*

This theorem closely resembles a well-known result from classic probability theory. From an abstract perspective, $\succcurlyeq$ is a binary relation comparing sets t of actions $a \in A$. Similarly, Kraft, Pratt, and Seidenberg (1959) and Scott (1964) consider a binary relation $\succsim$ comparing sets e of states $\omega \in \Omega$. There, the statement $e^1 \succ e^2$ means that the event e^1 is regarded as "more probable" than e^2, and the statement $e^1 \approx e^2$ means that the events e^1 and e^2 are regarded as "equally probable". Kraft, Pratt, and Seidenberg (1959, Theorem 2) and Scott (1964, Theorem 4.1) then state conditions on $\succsim$ which imply the existence of a probability mass function $p{:}\Omega \to [0,1]$ such that for all e^1 and e^2

$$\begin{aligned} e^1 \succ e^2 \;&\Rightarrow\; \Sigma_{\omega \in e^1} p(\omega) > \Sigma_{\omega \in e^2} p(\omega) \text{ and} \\ e^1 \approx e^2 \;&\Rightarrow\; \Sigma_{\omega \in e^1} p(\omega) = \Sigma_{\omega \in e^2} p(\omega)\ . \end{aligned}$$

In the terminology of the previous subsection, they state conditions on $\succsim$ which imply that $\succsim$ has a completion represented additively by a probability mass function $p{:}\Omega \to [0,1]$. Theorem 1 is unexpectedly similar.

In accord with this similarity, the remainder of this subsection will derive Theorem 1 from a well-known result in the non-economics literature. To begin, consider an arbitrary finite set A and an arbitrary binary relation $\succsim$ comparing subsets of A, which are denoted here by

$s \subseteq A$ and $t \subseteq A$. In this abstract setting, $\succsim$ is said to have a completion represented additively by a mass function $\mu{:}A{\rightarrow}\mathbb{R}$ if for all s and t

$$s \succ t \ \Rightarrow \ \Sigma_{a\in s}\mu(a) > \Sigma_{a\in t}\mu(a) \text{ and}$$
$$s \approx t \ \Rightarrow \ \Sigma_{a\in s}\mu(a) = \Sigma_{a\in t}\mu(a) \ ,$$

where $\succ$ and $\approx$ are the asymmetric and symmetric parts of $\succcurlyeq$.

Now let a *cancelling sample* from $\succsim$ be a finite indexed collection $\langle (s^m,t^m)\rangle_{m=1}^M$ of pairs (s^m,t^m) taken from $\succsim$ such that

$$(\forall a) \ |\{m|a{\in}s^m\}| = |\{m|a{\in}t^m\}| \ .$$

Note that the sample is taken "with replacement" in the sense that a pair can appear more than once. Further, by the equation, every action appearing on the left side of some pair is "cancelled" by the identical action appearing on the right side of that or some other pair. For example, if $\{a,a'\} \succsim \{a,a'\}$, then a cancelling sample from $\succsim$ is given by $M{=}1$ and $(s^1,t^1) = (\{a,a'\},\{a,a'\})$. The relation $\succsim$ is said to satisfy the *cancellation law* if every cancelling sample from $\succsim$ must be taken from the symmetric part of $\succsim$.

The cancellation law is equivalent to the existence of a completion represented additively by a mass function.[7] This result undergirds the foundations for probability in Kraft, Pratt, and Seidenberg (1959, Theorem 2) and Scott (1964, Theorem 4.1). It also undergirds the abstract representation theory in Krantz, Luce, Suppes, and Tversky (1971, Sections 2.3 and 9.2) and Narens (1985, pages 263-265).[8]

The following lemma is a very minor adaptation of that well-known result. Its proof requires nothing more than Farkas' Lemma (which can be derived from undergraduate linear algebra as in Vohra (2005, pages 16-18)). The lemma obtains an integer-valued mass function by employing a version of Farkas' Lemma for rational matrices.

[7]This result can be regarded as a generalization of Suzumura (1976, Theorem 3) in the social-choice literature. There it is shown that a relation on a finite set has a transitive completion iff every cycle in the relation contains no pairs from the asymmetric part of the relation. That result follows from the general result here by identifying every $a{\in}A$ with the singleton $\{a\}$ that contains it. In the realm of such singletons, the cancellation law is equivalent to Suzumura's rule on cycles, and a mass function μ reduces to Subsection 3.2's representation φ (which is in turn equivalent to the existence of a transitive extension).

[8]These classic results over discrete spaces complement Debreu (1960)'s derivation of an additive representation over continuum product spaces. Debreu imposes weaker cancellation assumptions (e.g., Debreu (1960, Assumption 1.3)) and compensates with topological assumptions.

Lemma 4.1. *Let A be a finite set, and let $\succsim$ be a relation comparing subsets of A. Then $\succsim$ satisfies the cancellation law iff it has a completion represented additively by a mass function $\mu{:}A{\to}\mathbb{Z}$. (Proof A.5 in the appendix.)*

We now use this abstract framework to prove Theorem 1.

Lemma 4.2. *Suppose an assessment is consistent. Then its support's plausibility relation $\succcurlyeq$ satisfies the cancellation law.*

Proof. Let (σ, β) be a consistent assessment, let $\succcurlyeq$ be its plausibility relation, and let $\langle (s^m, t^m) \rangle_{m=1}^M$ be cancelling sample from $\succcurlyeq$. By the definition of a cancelling sample,

$$(\forall n)\ \Pi_{m=1}^M \Pi_{a \in s^m} (\rho \cup \sigma_n)(a) = \Pi_{m=1}^M \Pi_{a \in t^m} (\rho \cup \sigma_n)(a)\ .$$

Thus

$$(\forall n)\ \Pi_{m=1}^M \frac{\Pi_{a \in t^m} (\rho \cup \sigma_n)(a)}{\Pi_{a \in s^m} (\rho \cup \sigma_n)(a)} = 1\ ,$$

which obviously implies

$$\lim_n\ \Pi_{m=1}^M \frac{\Pi_{a \in t^m} (\rho \cup \sigma_n)(a)}{\Pi_{a \in s^m} (\rho \cup \sigma_n)(a)} = 1\ . \tag{5}$$

By consistency, there is a sequence $\langle (\sigma_n, \beta_n) \rangle_{n=1}^\infty$ of Bayesian full-support assessments that converge to (σ, β). Thus by applying Lemma A.6 at $(t^1, t^2) = (s^m, t^m)$, we have (note s^m is in the denominator)

$$\begin{aligned} &(\forall\, s^m \succ t^m)\ \ \lim_n\ \frac{\Pi_{a \in t^m} (\rho \cup \sigma_n)(a)}{\Pi_{a \in s^m} (\rho \cup \sigma_n)(a)} = 0 \ \text{ and} \\ &(\forall\, s^m \approx t^m)\ \ \lim_n\ \frac{\Pi_{a \in t^m} (\rho \cup \sigma_n)(a)}{\Pi_{a \in s^m} (\rho \cup \sigma_n)(a)} \in (0, \infty)\ . \end{aligned} \tag{6}$$

If $\langle (s^m, t^m) \rangle_{m=1}^M$ has a pair from $\succ$, equations (5) and (6) contradict the product rule for limits. Hence no such pair exists. □

Proof 4.3 (for Theorem 1)**.** Consider a consistent assessment. By Lemma 4.2, its support's plausibility relation satisfies the cancellation law. Thus by Lemma 4.1, its support's plausibility relation has a completion represented additively by an integer-valued mass function. □

4.2. A Useful Digression

As discussed in Subsection 3.3, having a completion with an additive representation implies having a transitive completion. Hence Corollary 1 follows easily from Theorem 1 (and does not use the full force of the theorem).

Corollary 1. *If an assessment is consistent, then its support's plausibility relation has a transitive completion.*

Corollary 1 provides a weak but easily-tested necessary condition for consistency. For example, consider Figure 2's basement. Subsection 3.3 showed that this basement's plausibility relation does not have a transitive completion. Thus Corollary 1 implies the inconsistency of any assessment that is supported by this basement. One of many such assessments is the assessment of Figure 1.

4.3. Plausibility numbers "count steps below path"

In probability theory, an event e is at least as probable as any of its subsets, and thus, a probability mass function is nonnegative-valued. In contrast, a node t is at *most* as plausible as any of its subnodes (i.e. predecessors), and thus, a plausibility mass function is *nonpositive*-valued. The following lemma provides details. Its proof is uncomplicated.

Lemma 4.4. *Suppose b's plausibility relation has a completion represented additively by μ. Then (a) $\mu(A^c) = \{0\}$, (b) $\mu(A^s{\cap}b) = \{0\}$, (c) $\mu(A^s{\smallsetminus}b) \subseteq \mathbb{R}_{++}$, and (d) μ is nonpositive-valued.*

Proof. (a) Take any $a{\in}A^c$. By Lemma A.1(a), there is some t such that $a{\in}F(t)$. Then $t \approx t{\cup}\{a\}$ by the definition of $\approx_{A^c}$, which implies $\Sigma_{a'\in t}\mu(a') = \Sigma_{a'\in t\cup\{a\}}\mu(a')$ by representation (4), which implies $\mu(a) = 0$ since $a{\notin}t$ by $a{\in}F(t)$ and the definition of F.

(b) Take any $a{\in}A^s{\cap}b$. By Lemma A.1(b), there is some t such that $a{\in}F(t)$. Then $t \approx t{\cup}\{a\}$ by the definition of $\approx_{A^s}$, which implies $\Sigma_{a'\in t}\mu(a') = \Sigma_{a'\in t\cup\{a\}}\mu(a')$ by representation (4), which implies $\mu(a){=}0$ since $a{\notin}t$ by $a{\in}F(t)$ and the definition of F.

(c) Take any $a{\in}A^s{\smallsetminus}b$. By Lemma A.1(b), there is some t such that $a{\in}F(t)$. Thus $t \succ t{\cup}\{a\}$ by the definition of $\succ_{A^s}$, which implies $\Sigma_{a'\in t}\mu(a') > \Sigma_{a'\in t\cup\{a\}}\mu(a')$ by representation (4), which implies $\mu(a){<}0$.

(d) Because the domain of μ is $A = A^c{\cup}A^s$, parts (a)–(c) imply that μ is nonpositive-valued. □

In brief, parts (a) and (b) show that zero plausibilities are assigned to actions played with positive probability, and part (c) shows that negative plausibilities are assigned to actions played with zero probability. Thus a node's plausibility $\Sigma_{a \in t}\mu(a)$ is a measure of how far the node t is below the so-called "equilibrium path". This measure is slightly more sophisticated than (the negative of) the number of zero-probability actions in the node (i.e. leading to the node) because each zero-probability action can be assigned its own negative plausibility number.

To sharpen these ideas, we could refer to the "equilibrium path" as the "basement's path" since we are considering a basement which may or may not be the support of an equilibrium of some sort. Further, Theorem 1 suggests that we can "count steps" rather than "measure distance" below the basement's path because a plausibility mass function can be made to assume integer values.

5. CHARACTERIZATIONS OF THE SUPPORTS OF CONSISTENT ASSESSMENTS

5.1. CHARACTERIZATIONS USING $\succcurlyeq$

Theorem 1 shows that a basement has additive plausibility if it supports at least one consistent assessment. The converse of this result can be proved by applying the Kreps-Wilson definition of consistency. Accordingly, Theorem 2(a$\Leftrightarrow$d) (below) states that a basement has additive plausibility iff it supports at least one consistent assessment. In other words, it shows that additive plausibility characterizes the supports of consistent assessments. This equivalence is our main result, as reflected in the title of the paper.

In addition, Theorem 2(a$\Leftrightarrow$b) (below) shows that the supports of consistent assessments can be characterized by the cancellation law. And finally, Theorem 2(c$\Leftrightarrow$d) (below) shows that the integer-valuedness of a mass function is inconsequential. All of the above characterizations are new.

5.2. CHARACTERIZATIONS WITHOUT $\succcurlyeq$

It can also be useful to characterize the supports of consistent assessments without reference to a plausibility relation $\succcurlyeq$. Toward that end, say that a mass function μ *indicates* a basement b if

$$\mu(A^c) = \{0\} \; , \tag{7a}$$

$$\mu(A^s \cap b) = \{0\} \; , \tag{7b}$$

$$\mu(A^s \smallsetminus b) \subseteq \mathbb{R}_{--} \; , \text{ and} \tag{7c}$$

$$(\forall h \in H^s) \quad h \cap b = \text{argmax}_{t \in h} \, \Sigma_{a \in t} \mu(a) \; . \tag{7d}$$

Lemma 5.1. *μ indicates b iff μ additively represents a completion of the plausibility relation of b. (Proof A.8 in the appendix.)*

Theorem 2(a$\Leftrightarrow$g) (below) shows that indication characterizes the supports of consistent assessments. This characterization is related to a portion of the characterizations of consistency in Kreps and Wilson (1982, Lemmas A1 and A2) and Perea y Monsuwé, Jansen, and Peters (1997, Theorem 3.1). Note that these two papers characterize not merely the supports of consistent assessments, as we do here, but more substantially, the consistent assessments themselves. We believe that their important results have been largely overlooked, and that they deserve much more recognition.

The remainder of this subsection (a) reviews the relevant concepts from these two papers, (b) states and proves Theorem 2, and (c) compares portions of Theorem 2 to these two papers.

First consider Kreps and Wilson (1982).[9] As in their equation (A.1) on page 887, say that a function $K{:}A^s \to \mathbb{Z}_+$ *labels* a basement b if

$$(\forall a \in A^s) \quad a \in b \;\text{ iff }\; K(a) = 0 \; , \text{ and} \tag{8a}$$

$$(\forall h \in H^s)(\forall t \in h) \quad t \in b \;\text{ iff }\; t \in \text{argmin}_{t' \in h} \, \Sigma_{a \in t'} K(a) \; . \tag{8b}$$

---

[9]Let A^c be their W, let T^c be $\{\{w\}|w \in W\}$, let A^s be their A, let $t \in T^s$ be their $x \in X$, and let (σ, β) be their (π, μ). Further, note that our concept of a basement is a restriction of their concept of a "basis", which they define on page 880 to be an arbitrary subset of $A^s \cup T^s$. Although this restriction is inconsequential by Lemma A.2, it conveniently allows us to omit one condition of their definition of labelling. In particular, we can omit their condition $(\forall h \in H^s)(\exists a \in F(h))\, K(a)=0$ because it is implied by (8a) and the definition of basement. To see this, take any $h \in H^s$. The definition of a basement b implies that there is an $a \in F(h) \cap b$, and thus (8a) implies that this a satisfies $K(a)=0$. Finally, in a separate matter, note that their $J_K(t)$ is equal to $\Sigma_{a \in t} K(a)$ in our set-tree formulation.

The following lemma relates a labelling K to a mass function μ.

Lemma 5.2. *The following are equivalent.*
(a) $\mu(A) \subseteq \mathbb{Z}$ *and* μ *indicates* b.
(b) $\mu(A^c) = \{0\}$ *and* $-\mu|_{A^s}$ *is a* K *that labels* b.
(Proof A.9 in the appendix.)

Second consider Perea y Monsuwé, Jansen, and Peters (1997).[10] For our purposes, say that a basement b is *justified* if there exists $\varepsilon{:}A^s\smallsetminus b\rightarrow(0,1)$ such that

$$(\forall h{\in}H^s)(\forall t{\in}h) \quad t \in b \quad \text{iff} \quad t \in \operatorname{argmax}_{t'\in h} \Pi_{a\in t'\cap(A^s\smallsetminus b)}\varepsilon(a) \ ,$$

where the product over the empty set is defined to be one. If b is the support of (σ,β), then b is justified iff there exists $\varepsilon{:}A^0(\sigma)\rightarrow(0,1)$ such that

$$(\forall h{\in}H^s)(\forall t{\in}h) \quad \beta(t) > 0 \quad \text{iff} \quad t \in \operatorname{argmax}_{t'\in h} \Pi_{a\in t'\cap A^0(\sigma)}\varepsilon(a) \ ,$$

where $A^0(\sigma) = \{a{\in}A^s|\sigma(a){=}0\}$ is the set of strategic actions that are played by σ with zero probability. This latter condition is equivalent to condition (1) of Perea y Monsuwé, Jansen, and Peters (1997, Theorem 3.1). The values of ε can usefully be called "error likelihoods", for as Perea (2001, page 75) points out, they measure the relative likelihoods of different errors (i.e., zero-probability actions).

The following lemma relates an error-likelihood function ε to a plausibility mass function μ. Note that $e^\mu|_{A^s\smallsetminus b}$ is the restriction of the composition $e^\mu = \exp\circ\mu$ to $A^s\smallsetminus b$. Thus, for all $a{\in}A^s\smallsetminus b$, a plausibility mass function μ is related to an error-likelihood function ε by $e^{\mu(a)} = \varepsilon(a)$, or in other words, by $\mu(a) = \ln(\varepsilon(a))$.

Lemma 5.3. μ *indicates* b *iff (a)* $\mu(A^c) = \{0\}$ *, (b)* $\mu(A^s{\cap}b) = \{0\}$ *, (c)* $\mu(A^s\smallsetminus b) \subseteq \mathbb{R}_{++}$ *, and (d)* $e^\mu|_{A^s\smallsetminus b}$ *is an* ε *that justifies* b. *(Proof A.10 in the appendix.)*

The following theorem subsumes Theorem 1 and incorporates all of the above concepts.

Theorem 2. *The following are equivalent for any basement* b.
(a) b *is the support of at least one consistent assessment.*

---

[10]Let A^s be their A, and let $t{\in}T^s$ be their $x{\in}X$. Note that their A_x (the set of actions leading to a node) is identical to the node itself in our set-tree formulation. Finally note that if b supports (σ,β), then their $A^0(\sigma)$ equals our $A^s\smallsetminus b$.

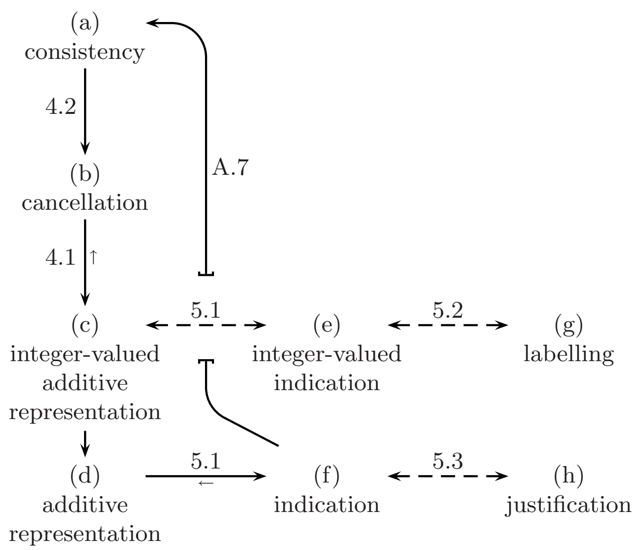


FIGURE 4. The solid arrows down the left side show Theorem 1's proof. Altogether the solid arrows show Theorem 2's proof's first paragraph. The dotted arrows show its second paragraph. (The tiny arrows show where converses can be proved directly.)

(b) b's plausibility relation obeys the cancellation law.

(c) b's plausibility relation has a completion represented additively by an integer-valued mass function.

(d) b's plausibility relation has a completion represented additively by a (real-valued) mass function.

(e) b is indicated by an integer-valued mass function.

(f) b is indicated by a (real-valued) mass function.

(g) b can be labelled (a la Kreps and Wilson (1982)).

(h) b can be justified (a la Perea y Monsuwé, Jansen, and Peters (1997)).

Proof. (See Figure 4.) (a) implies (b) by Lemma 4.2, (b) implies (c) by Lemma 4.1, and (c) implies (d) obviously. (Theorem 1 was proved by this chain of reasoning.) Conversely, (d) implies (f) by Lemma 5.1, and (f) implies (a) by Lemma A.7. Hence (a), (b), (c), (d), and (f) are equivalent.

(c) is equivalent to (e) and (g) by Lemmas 5.1 and 5.2. (f) is equivalent to (h) by Lemma 5.3. □

The equivalence of (a) and (g) extends Kreps and Wilson (1982, Lemma A1) to the extent that (1) it relaxes perfect recall to agent recall and (2) it accommodates arbitrary chance agents. In addition, we correct their proof: Streufert (2006, Subsection 3.2) shows by counterexample that their proof has a significant gap. This is important because their Lemma A1 undergirds all of their paper's theorems.

The equivalence of (a) and (h) extends a portion of Perea y Monsuwé, Jansen, and Peters (1997, Theorem 3.1): it is more general to the extent that (1) it relaxes perfect recall to agent recall, (2) it accommodates arbitrary chance agents, and (3) it derives a mass function with integer rather than real values. In addition, our mathematics is at a lower level: while their proof uses the separating hyperplane theorem, ours uses algebra alone.

Finally, the equivalence of (g) and (h) unifies Kreps and Wilson (1982) and Perea y Monsuwé, Jansen, and Peters (1997). Only a superficial similarity has been noted earlier.

In this manner the last two characterizations of Theorem 2 extend, correct, simplify, and unify an important literature which we believe deserves much greater recognition.

6. Conclusion

Theorem 1 shows that a consistent assessment's support must have additive plausibility. It does this by combining classic probability theory's use of Farkas' Lemma with Streufert (2012)'s result that a node can be identified with the set of actions leading to it. Corollary 1 then provides a weaker necessary condition for consistency: the plausibility relation of a consistent assessment's support must have a transitive extension.

Theorem 2 incorporates Theorem 1 and its converse. Thus additive plausibility characterizes the supports of consistent assessments. This is our main result. Theorem 2 also shows that additive plausibility is equivalent to classic probability theory's cancellation law, and that it is equivalent to an algebraic condition defined without reference to a plausibility relation. This last characterization allows us to fill a gap in a proof of Kreps and Wilson (1982).

Appendix

A.1. Preliminaries

Lemma A.1. *(a)* $A^c = \bigcup_{t\in T^c} F(t)$. *(b)* $A^s = \bigcup_{t\in T^s} F(t)$.

Proof. (a) First take any $a{\in}A^c$. By the definition of A^c there exists some $h{\in}i^c$ such that $a{\in}F(h)$. Take any $t{\in}h$. By $h{\in}i^c$ and the definition of T^c, we have $t{\in}T^c$. Hence $a \in F(h) = F(t) \subseteq \bigcup_{t'\in T^c} F(t')$. Second take any $a{\in}\bigcup_{t'\in T^c} F(t')$. Then let $t{\in}T^c$ be such that $a{\in}F(t)$. By the definition of T^c there exists $h{\in}i^c$ such that $t{\in}h$. Hence $a \in F(t) = F(h) \subseteq \bigcup_{h'\in i^c} F(h') = A^c$, where the last equality is the definition of A^c.

(b) A symmetric argument holds with A^s replacing A^c, T^s replacing T^c, and $h{\in}H^s$ replacing $h{\in}i^c$. □

Lemma A.2. *A subset of $A^s{\cap}T^s$ supports at least one assessment iff it is a basement.*

Proof. Let $\bar{b}$ be an arbitrary subset of $A^s{\cup}T^s$.

Consider any assessment (σ,β). Note that $\bar{b}$ supports (σ,β) iff $\bar{b}{\cap}A^s$ supports $\sigma{:}A^s{\to}[0,1]$ and $\bar{b}{\cap}T^s$ supports $\beta{:}T^s{\to}[0,1]$. Since $A^s = \bigcup_{h\in H^s} F(h)$, $\bar{b}{\cap}A^s$ supports $\sigma{:}A^s{\to}[0,1]$ iff $(\forall h{\in}H^s)$ $F(h){\cap}\bar{b}$ supports $\sigma|_{F(h)}$. Similarly, since $T^s = \bigcup_{h\in H^s} h$, $\bar{b}{\cap}T^s$ supports $\beta{:}T^s{\to}[0,1]$ iff $(\forall h{\in}H^s)$ $h{\cap}\bar{b}$ supports $\beta|_h$. Thus $\bar{b}$ supports (σ,β) iff, for every agent $h{\in}H^s$, $F(h){\cap}\bar{b}$ supports the agent's strategy $\sigma|_{F(h)}$ and $h{\cap}\bar{b}$ supports the agent's belief $\beta|_h$.

Therefore, if $\bar{b}$ supports an assessment, it must be that every $F(h){\cap}\bar{b}$ and every $h{\cap}\bar{b}$ is nonempty. Conversely, if every $F(h){\cap}\bar{b}$ and every $h{\cap}\bar{b}$ is nonempty, then $\bar{b}$ supports the assessment (σ,β) defined at each $h{\in}H^s$ by

$$(\forall a{\in}F(h))\ \sigma(a) = \begin{pmatrix} \frac{1}{|F(h)\cap\bar{b}|} & \text{if } a{\in}F(h){\cap}\bar{b} \\ 0 & \text{if } a{\in}F(h){\setminus}\bar{b} \end{pmatrix} \text{ and}$$

$$(\forall t{\in}h)\ \beta(t) = \begin{pmatrix} \frac{1}{|h\cap\bar{b}|} & \text{if } a{\in}h{\cap}\bar{b} \\ 0 & \text{if } a{\in}h{\setminus}\bar{b} \end{pmatrix} .$$

By the definition of basement, the last two sentences imply that $\bar{b}$ supports at least one assessment iff it is a basement. □

Proof A.3 (for Lemma 3.1)**.** Note $\approx_{A^c}$ is symmetric and equal to

$$\begin{aligned} &\{\ (t, t{\cup}\{a\}) \mid a{\in}F(t) \text{ and } a{\in}A^c\ \} \\ \cup\ &\{\ (t{\cup}\{a\}, t) \mid a{\in}F(t) \text{ and } a{\in}A^c\ \}\ . \end{aligned} \tag{9}$$

Further, $\approx_{A^s}$ is symmetric, $\succ_{A^s}$ is asymmetric, and the two are disjoint subsets of

$$\begin{aligned} &\{\ (t, t{\cup}\{a\}) \mid a{\in}F(t) \text{ and } a{\in}A^s\ \} \\ \cup\ &\{\ (t{\cup}\{a\}, t) \mid a{\in}F(t) \text{ and } a{\in}A^s\ \}\ . \end{aligned} \tag{10}$$

Similarly, $\approx_{T^s}$ is symmetric, $\succ_{T^s}$ is asymmetric, and the two are disjoint subsets of

$$\{\ (t_1, t_2) \mid (\exists h{\in}H^s)\ \{t_1, t_2\}{\in}h\ \}\ . \tag{11}$$

This paragraph observes that these three sets are pairwise disjoint. (9) and (10) are disjoint because A is partitioned by $\{A^c, A^s\}$. Further, the union of (9) and (10) is disjoint from (11). If this were not the case, there would be t, a, and h such that $a{\in}F(t)$ and $\{t, t{\cup}\{a\}\}{\in}h$. Since $a{\in}F(t)$ and t and $t{\cup}\{a\}$ share an agent, assumption (1a) would imply that $a{\in}F(t{\cup}\{a\})$. However, this would contradict the definition of F, which would require that $a{\notin}t{\cup}\{a\}$.

Since the sets (9), (10), and (11) are pairwise disjoint, the disjointedness observed in the first paragraph implies that $\succcurlyeq$ is partitioned by $\{\ \approx_{A^c}, \approx_{A^s}, \succ_{A^s}, \approx_{T^s}, \succ_{T^s}\ \}$. Thus the symmetries and asymmetries observed in the first paragraph imply that $\approx$ is partitioned by $\{\approx_{A^c}, \approx_{A^s}, \approx_{T^s}\ \}$ and that $\succ$ is partitioned by $\{\ \succ_{A^s}, \succ_{T^s}\ \}$. □

A.2. For Theorem 1's proof

Fact A.4 (Farkas Lemma for Rational Matrices)**.** *Let $D \in \mathbb{Q}^{dp}$ and $E \in \mathbb{Q}^{ep}$ be two rational matrices. Then the following are equivalent ($D\mu{\gg}0$ means every element of $D\mu$ is positive and δ^T means the transpose of δ).*

(a) $\qquad (\exists\mu{\in}\mathbb{Z}^p)\ D\mu \gg 0$ *and* $E\mu = 0$.

(b) $\qquad$ *Not* $(\exists\delta{\in}\mathbb{Z}^d_+{\smallsetminus}\{0\})(\exists\varepsilon{\in}\mathbb{Z}^e)\ \delta^T D + \varepsilon^T E = 0$.

(From Krantz, Luce, Suppes, and Tversky (1971, pages 62–63) with D replacing $[\alpha_i]_{i=1}^{m'}$ and E replacing $[\beta_i]_{i=1}^{m''}$.)

Proof A.5 (for Lemma 4.1)**.** Sufficiency of the Cancellation Law.

First take any relation $\succsim$. For any t, define the row vector $1^t \in \{0,1\}^{|A|}$ by $1^t_a = 1$ if $a \in t$ and $1^t_a = 0$ if $a \notin t$. Then define the matrices $D = [1^s{-}1^t]_{s\succ t}$ and $E = [1^s{-}1^t]_{s\approx t}$ whose rows are indexed by the pairs of the relations $\succ$ and $\approx$.

Now suppose $\succsim$ satisfies the cancellation law. This paragraph will argue that there cannot be column vectors $\delta \in \mathbb{Z}_+^{|\succ|} \smallsetminus \{0\}$ and $\varepsilon \in \mathbb{Z}^{|\approx|}$ such that $\delta^T D + \varepsilon^T E = 0$. To see this, suppose that there were such δ and ε. By the symmetry of $\approx$, we may define $\varepsilon_+ \in \mathbb{Z}_+^{|\approx|}$ by

$$(\forall s{\approx}t)\ (\varepsilon_+)_{(s,t)} = \begin{pmatrix} \varepsilon_{(s,t)}{-}\varepsilon_{(t,s)} & \text{if } \varepsilon_{(s,t)}{-}\varepsilon_{(t,s)} \geq 0 \\ 0 & \text{otherwise} \end{pmatrix}$$

so that $\varepsilon^T E = \varepsilon_+^T E$. Thus we have $\delta \in \mathbb{Z}_+^{|\succ|} \smallsetminus \{0\}$ and $\varepsilon_+ \in \mathbb{Z}_+^{|\approx|}$ such that $\delta^T D + \varepsilon_+^T E = 0$. Now define the sequence $\langle (s^m, t^m) \rangle_{m=1}^M$ of pairs from $\succsim$ in such a way that every pair from $\succ$ appears $\lambda_{(s,t)}$ times and every pair from $\approx$ appears $(\mu_+)_{(s,t)}$ times. The equality $\delta^T D + \varepsilon_+^T E = 0$ yields that this sequence satisfies $(\forall a)\ |\{m|a{\in}s^m\}| = |\{m|a{\in}t^m\}|$, and $\delta \in \mathbb{Z}_+^{|\prec|} \smallsetminus \{0\}$ yields that it contains at least one pair from $\succ$. By the cancellation law, this is impossible.

Since the result of the previous paragraph is equivalent to condition (b) of Lemma A.4, there is a vector $\mu \in \mathbb{Z}^{|A|}$ such that $D\mu \gg 0$ and $E\mu = 0$. By the definitions of D and E, this is equivalent to μ being a mass function representing a completion of $\succcurlyeq$.

Necessity of the Cancellation Law.[11] Suppose μ represents an completion of $\succsim$ and that $\langle (s^m, t^m) \rangle_{m=1}^M$ is a cancelling sample from $\succsim$. By the definition of a cancelling sample,

$$\Sigma_{m=1}^M \Sigma_{a \in s^m} \mu(a) = \Sigma_{m=1}^M \Sigma_{a \in t^m} \mu(a)\ .$$

Yet by representation,

$$\begin{aligned} &(\forall\, s^m \succ t^m)\ \Sigma_{a \in s^m} \mu(a) > \Sigma_{a \in t^m} \mu(a) \ \text{ and} \\ &(\forall\, s^m \approx t^m)\ \Sigma_{a \in s^m} \mu(a) = \Sigma_{a \in t^m} \mu(a)\ . \end{aligned}$$

The last two sentences contradict if $\langle (s^m, t^m) \rangle_{m=1}^M$ has a pair from $\succ$. Hence no such pair exists. □

Lemma A.6. *Suppose that $\langle (\sigma_n, \beta_n) \rangle_n$ is a sequence of full-support Bayesian assessments that converges to (σ, β), and that (σ, β)'s support's plausibility relation is $\succcurlyeq$. Then*

$$(\forall\, t^1 \succ t^2)\ \lim_n \frac{\Pi_{a \in t^2} (\rho{\cup}\sigma_n)(a)}{\Pi_{a \in t^1} (\rho{\cup}\sigma_n)(a)} = 0 \quad \textit{and}$$

[11] This half is easy and is included only to round out the picture. It is a good place to begin if the cancellation law is unfamiliar.

$$(\forall\, t^1 \approx t^2)\ \lim_n \frac{\Pi_{a\in t^2}(\rho\cup\sigma_n)(a)}{\Pi_{a\in t^1}(\rho\cup\sigma_n)(a)} \in (0,\infty)$$

(where $\Pi_{a\in\{\}}(\rho\cup\sigma_n)(a)$ *is defined to be one).*

Proof. Take any such $\langle(\sigma_n,\beta_n)\rangle_n$ and (σ,β), let (σ,β)'s support be b, and let b's plausibility relation be $\succcurlyeq$.

This paragraph shows

$$(12) \qquad (\forall\, t^1 \succ_{A^s} t^2)\ \lim_n \frac{\Pi_{a'\in t^2}(\rho\cup\sigma_n)(a')}{\Pi_{a'\in t^1}(\rho\cup\sigma_n)(a')} = 0\ .$$

Suppose $t^1 \succ_{A^s} t^2$. By the definition of $\succ_{A^s}$, there exists $a\in A^s\smallsetminus b$ such that $a\in F(t^1)$ and $t^1\cup\{a\} = t^2$. Thus,

$$\lim_n \frac{\Pi_{a'\in t^2}(\rho\cup\sigma_n)(a')}{\Pi_{a'\in t^1}(\rho\cup\sigma_n)(a')} = \lim_n \sigma_n(a) = \sigma(a) = 0\ .$$

where the first equality holds since $a\notin t^1$ by the definition of F, and the third equality holds because $a\in A^s\smallsetminus b$.

This paragraph shows

$$(13) \qquad (\forall\, t^1 \approx_{A^s} t^2)\ \lim_n \frac{\Pi_{a'\in t^2}(\rho\cup\sigma_n)(a')}{\Pi_{a'\in t^1}(\rho\cup\sigma_n)(a')} \in (0,\infty)\ .$$

Suppose $t^1 \approx_{A^s} t^2$. By the definition of $\approx_{A^s}$, there exists $a\in A^s\cap b$ such that either [a] $a\in F(t^1)$ and $t^1\cup\{a\}{=}t^2$, or symmetrically [b] $a\in F(t^2)$ and $t^2\cup\{a\}{=}t^1$. In case [a],

$$\lim_n \frac{\Pi_{a'\in t^2}(\rho\cup\sigma_n)(a')}{\Pi_{a'\in t^1}(\rho\cup\sigma_n)(a')} = \lim_n \sigma_n(a) = \sigma(a) \in (0,1]\ ,$$

where the first equality holds since $a\notin t^1$ by the definition of F, and the set inclusion holds since $a\in A^s\cap b$. In case [b],

$$\lim_n \frac{\Pi_{a'\in t^2}(\rho\cup\sigma_n)(a')}{\Pi_{a'\in t^1}(\rho\cup\sigma_n)(a')} = \lim_n \frac{1}{\sigma_n(a)} = \frac{1}{\sigma(a)} \in [1,\infty)\ .$$

where the first equality holds since $a\notin t^2$ by the definition of F, and the set inclusion holds since $a\in A^s\cap b$.

In a similar fashion, this paragraph shows

$$(14) \qquad (\forall\, t^1 \approx_{A^c} t^2)\ \lim_n \frac{\Pi_{a'\in t^2}(\rho\cup\sigma_n)(a')}{\Pi_{a'\in t^1}(\rho\cup\sigma_n)(a')} \in (0,\infty)\ .$$

Suppose $t^1 \approx_{A^c} t^2$. By the definition of $\approx_{A^c}$, there exists $a\in A^c$ such that either, [a] $a\in F(t^1)$ and $t^1\cup\{a\}{=}t^2$, or symmetrically [b] $a\in F(t^2)$ and

$t^2{\cup}\{a\}{=}t^1$. In case [a], $a{\notin}t^1$ and thus

$$\lim_n \frac{\Pi_{a'\in t^2}(\rho{\cup}\sigma_n)(a')}{\Pi_{a'\in t^1}(\rho{\cup}\sigma_n)(a')} = \lim_n \rho(a) \in (0,1]\ .$$

In case [b], $a{\notin}t^2$ and thus

$$\lim_n \frac{\Pi_{a'\in t^2}(\rho{\cup}\sigma_n)(a')}{\Pi_{a'\in t^1}(\rho{\cup}\sigma_n)(a')} = \lim_n \frac{1}{\rho(a)} \in [1,\infty)\ .$$

This paragraph notes that if t^1 and t^2 belong to the same $h{\in}H^s$, and if $\beta(t^1) > 0$, then

$$\begin{aligned} &\lim_n \frac{\Pi_{a\in t^2}(\rho{\cup}\sigma_n)(a)}{\Pi_{a\in t^1}(\rho{\cup}\sigma_n)(a)} \\ = \ &\lim_n \frac{\Pi_{a\in t^2}(\rho{\cup}\sigma_n)(a)/\Sigma_{t\in h}\Pi_{a\in t}(\rho{\cup}\sigma_n)(a)}{\Pi_{a\in t^1}(\rho{\cup}\sigma_n)(a)/\Sigma_{t\in h}\Pi_{a\in t}(\rho{\cup}\sigma_n)(a)} \\ = \ &\lim_n \frac{\beta_n(t^2)}{\beta_n(t^1)} = \frac{\beta(t^2)}{\beta(t^1)}\ . \end{aligned}$$

The second equality follows from the conditional-probability law (2), which applies over h because each (σ_n, β_n) was assumed to be Bayesian. The third follows from this paragraph's assumption that $\beta(t^1) > 0$.

This paragraph shows

$$(15) \qquad (\forall\, t^1 \succ_{T^s} t^2)\ \lim_n \frac{\Pi_{a\in t^2}(\rho{\cup}\sigma_n)(a)}{\Pi_{a\in t^1}(\rho{\cup}\sigma_n)(a)} = \frac{\beta(t^2)}{\beta(t^1)} = 0\ .$$

Suppose $t^1 \succ_{T^s} t^2$. By the definition of $\succ_{T^s}$, we have that (a) t^1 and t^2 belong to the same $h{\in}H^s$, (b) $t^1{\in}T^s{\cap}b$, and (c) $t^2{\in}T^2{\smallsetminus}b$. The first equality in (15) holds by the last paragraph, (a), and (b), because (b) implies $\beta(t^1) > 0$. The second equality follows from (c), because (c) implies $\beta(t^2) = 0$.

Similarly, this paragraph shows

$$(16) \qquad (\forall\, t^1 \approx_{T^s} t^2)\ \lim_n \frac{\Pi_{a\in t^2}(\rho{\cup}\sigma_n)(a)}{\Pi_{a\in t^1}(\rho{\cup}\sigma_n)(a)} = \frac{\beta(t^2)}{\beta(t^1)} \in (0,\infty)\ .$$

Suppose $t^1 \approx_{T^s} t^2$. By the definition of $\approx_{T^s}$, we have that (a) t^1 and t^2 belong to the same $h{\in}H^s$, and (b) $\{t^1, t^2\}{\subseteq}T^2{\cap}b$. The first equality in (16) holds by the next-to-last paragraph, (a), and (b), because (b) implies $\beta(t^1) > 0$. The inequality also follows from (b), because (b) also implies $\beta(t^2) > 0$.

The lemma's conclusion follows from (12)–(16) and from the definitions of $\succ$ and $\approx$. $\square$

A.3. For Theorem 2's Proof

Lemma A.7. *If a basement is indicated by a mass function, then it supports a consistent assessment.*

Proof. Suppose that μ indicates b.

This and the next paragraph define an assessment (σ,β). Since $A^s = \bigcup_{h\in H^s} F(h)$, define $\sigma{:}A^s\to[0,1]$ at each $h\in H^s$ and each $a\in F(h)$ by

$$\sigma(a) = \begin{pmatrix} \frac{1}{|F(h)\cap b|} & \text{if } a\in F(h)\cap b \\ 0 & \text{if } a\in F(h)\smallsetminus b \end{pmatrix} .$$

Every $\sigma(a)$ is a well-defined number in $[0,1]$ because the definition of a basement states that $|F(h)\cap b|\geq 1$ at each $h\in H^s$. Further, at each $h\in H^s$, the restriction $\sigma|_{F(h)}{:}F(h)\to[0,1]$ is a probability distribution over $F(h)$ because $\Sigma_{a\in F(h)}\sigma(a) = \Sigma_{a\in F(h)\cap b}\,\frac{1}{|F(h)\cap b|} = 1$.

The definition of β requires two steps. First, since $A^s = \bigcup_{h\in H^s} F(h)$, define $\lambda{:}A^s\to(0,1]$ at each $h\in H^s$ and each $a\in F(h)$ by

$$\lambda(a) = \begin{pmatrix} \sigma(a) & \text{if } a\in F(h)\cap b \\ 1 & \text{if } a\in F(h)\smallsetminus b \end{pmatrix} .$$

Then, since $T^s = \bigcup_{h\in H^s} h$, define $\beta{:}T^s\to[0,1]$ at each $h\in H^s$ and each $t\in h$ by

$$\beta(t) = \begin{pmatrix} \dfrac{\Pi_{a\in t}\,(\rho\cup\lambda)(a)}{\Sigma_{t'\in T^\mu(h)}\,\Pi_{a\in t'}\,(\rho\cup\lambda)(a)} & \text{if } t\in T^\mu(h) \\ 0 & \text{if } t\in h\smallsetminus T^\mu(h) \end{pmatrix} ,$$

where $T^\mu(h) = \text{argmax}_{t'\in h}\,\Sigma_{a\in t'\cap A^s}\mu(a)$, and where $\rho\cup\lambda{:}A\to(0,1]$ is the union of $\rho{:}A^c\to(0,1)$ and $\lambda{:}A^s\to(0,1]$. Every $\beta(t)$ is a well-defined number in $[0,1]$ because $\rho\cup\lambda$ is positive-valued. Further, at each $h\in H^s$, the restriction $\beta|_h{:}h\to[0,1]$ is a probability distribution because $\Sigma_{t\in h}\beta(t) = \Sigma_{t\in T^\mu(h)}\beta(t) = 1$.

This paragraph shows that b supports (σ,β). This is equivalent to showing that (a) the support of each $\sigma|_h$ is $F(h)\cap b$ and (b) the support of each $\beta|_h$ is $h\cap b$. Statement (a) holds by the definition of σ. Statement (b) holds since the support of each β_h is defined to be $T^\mu(h) = \text{argmax}_{t'\in h}\Sigma_{a\in t'\cap A^s}\mu(a)$, and since this equals $h\cap b$ by the assumption that μ indicates b.

It remains to show that (σ,β) is consistent. Accordingly, this paragraph defines a sequence $\langle(\sigma_n,\beta_n)\rangle_n$ of full-support Bayesian assessments. Take any n. First define σ_n by

$$(\forall a{\in}A^s)\ \ \sigma_n(a) = \frac{\lambda(a)n^{\mu(a)}}{\Sigma_{a'\in F(F^{-1}(a))}\,\lambda(a')n^{\mu(a')}}\ ,$$

where $F^{-1}(a)$ denotes the agent that plays a (this $F^{-1}(a)$ is well-defined because actions are assumed to be agent-specific by assumption (1b)). Note that σ_n is positive-valued because λ is positive-valued. Then derive β_n by the conditional-probability rule. In other words, define β_n by

$$(\forall h{\in}H^s)(\forall t{\in}h)\ \ \beta_n(t) = \frac{\Pi_{a\in t}(\rho{\cup}\sigma_n)(a)}{\Sigma_{t'\in h}\Pi_{a\in t'}(\rho{\cup}\sigma_n)(a)}\ .$$

This β_n is well-defined because σ_n is positive-valued.

It remains to show that $\langle(\sigma_n,\beta_n)\rangle_n$ converges to (σ,β). As a first step, this paragraph shows

$$(\forall a{\in}A^s)\ \lim_n\ \lambda(a)n^{\mu(a)} = \sigma(a)\ . \tag{17}$$

On the one hand, consider $a{\in}A^s{\cap}b$. Here we have

$$\lim_n\ \lambda(a)n^{\mu(a)} = \lim_n\ \sigma(a)n^0 = \sigma(a)$$

since $\lambda(a) = \sigma(a)$ by the definition of λ, and since $\mu(A^s{\cap}b) = \{0\}$ by the assumption that μ indicates b. On the other hand, consider any $a{\in}A^s{\smallsetminus}b$. Here we have

$$\lim_n\ \lambda(a)n^{\mu(a)} = \lim_n\ n^{\mu(a)} = 0 = \sigma(a)\ ,$$

where the first equality holds since $\lambda(a) = 1$ by the definition of λ, the second holds since $\mu(A^s{\smallsetminus}b){\subseteq}\mathbb{R}_{--}$ by the assumption that μ indicates b, and the third holds because $a{\notin}b$ by assumption.

This paragraph merely notes that

$$(\forall h{\in}H^s)\ \lim_n\ \Sigma_{a\in F(h)}\lambda(a)n^{\mu(a)} = \Sigma_{a\in F(h)}\sigma(a) = 1\ . \tag{18}$$

The first equality follows from (17). The second follows from $\sigma|_h$ being a probability distribution, as shown in the paragraph defining σ.

We now show that $\langle\sigma_n\rangle_n$ converges to σ. In particular, we argue that for any a,

$$\lim_n\ \sigma_n(a) = \lim_n\ \frac{\lambda(a)n^{\mu(a)}}{\Sigma_{a'\in F(F^{-1}(a))}\,\lambda(a')n^{\mu(a')}} = \sigma(a)\ .$$

The first equality holds by the definition of σ_n. The second equality is an application of the quotient rule for limits: the limit of the numerator is $\sigma(a)$ by (17) and the limit of the denominator is one by (18) applied at $h = F^{-1}(a)$.

It remains to show that $\langle\beta_n\rangle_n$ converges to β. To do this, fix $h{\in}H^s$ and define $m = \max_{t\in h} \Sigma_{a\in t\cap A^s}\mu(a)$.[12] As a first step, the remainder of this paragraph argues that

$$
\begin{aligned}
&(\forall t{\in}h) \quad \lim_n\, n^{-m} \times \Pi_{a\in t\cap A^s}\sigma_n(a) \\
\text{(19a)} \quad &= \lim_n\, n^{-m} \times \Pi_{a\in t\cap A^s} \frac{\lambda(a)n^{\mu(a)}}{\Sigma_{a'\in F(F^{-1}(a))}\,\lambda(a')n^{\mu(a')}} \\
\text{(19b)} \quad &= \lim_n\, \frac{n^{-m} \times \Pi_{a\in t\cap A^s}\lambda(a)n^{\mu(a)}}{\Pi_{a\in t\cap A^s}\Sigma_{a'\in F(F^{-1}(a))}\lambda(a')n^{\mu(a')}} \\
\text{(19c)} \quad &= \lim_n\, \frac{\Pi_{a\in t\cap A^s}\lambda(a) \times n^{-m+\Sigma_{a\in t\cap A^s}\mu(a)}}{\Pi_{a\in t\cap A^s}\Sigma_{a'\in F(F^{-1}(a))}\lambda(a')n^{\mu(a')}} \\
\text{(19d)} \quad &= \Pi_{a\in t\cap A^s}\lambda(a) \times \lim_n\, n^{-m+\Sigma_{a\in t\cap A^s}\mu(a)}\ .
\end{aligned}
$$

(19a) holds by the definition of each σ_n. (19b) and (19c) hold by algebraic manipulation. (19d) is an application of the quotient rule for limits: the limit of the numerator is finite by the constancy of $\Pi_{a\in t\cap A^s}\lambda(a)$ and by definition of m, and in the denominator, at each $a{\in}t{\cap}A^s$, the limit of the sum is one by (18) applied at the agent $F^{-1}(a)$.

Finally we argue that

$$
\begin{aligned}
&(\forall t{\in}h) \quad \lim_n\, \beta_n(t) \\
\text{(20a)} \quad &= \lim_n\, \frac{\Pi_{a\in t}(\rho{\cup}\sigma_n)(a)}{\Sigma_{t'\in h}\Pi_{a\in t'}(\rho{\cup}\sigma_n)(a)} \\
\text{(20b)} \quad &= \lim_n\, \frac{\Pi_{a\in t\cap A^c}\rho(a) \times \Pi_{a\in t\cap A^s}\sigma_n(a)}{\Sigma_{t'\in h}\ \Pi_{a\in t'\cap A^c}\rho(a) \times \Pi_{a\in t'\cap A^s}\sigma_n(a)} \\
\text{(20c)} \quad &= \lim_n\, \frac{\Pi_{a\in t\cap A^c}\rho(a) \times n^{-m} \times \Pi_{a\in t\cap A^s}\sigma_n(a)}{\Sigma_{t'\in h}\ \Pi_{a\in t'\cap A^c}\rho(a) \times n^{-m} \times \Pi_{a\in t'\cap A^s}\sigma_n(a)} \\
\text{(20d)} \quad &= \frac{\Pi_{a\in t\cap A^c}\rho(a) \times \Pi_{a\in t\cap A^s}\lambda(a) \times \lim_n\ n^{-m+\Sigma_{a\in t\cap A^s}\mu(a)}}{\Sigma_{t'\in h}\ \Pi_{a\in t'\cap A^c}\rho(a) \times \Pi_{a\in t'\cap A^s}\lambda(a) \times \lim_n\ n^{-m+\Sigma_{a\in t'\cap A^s}\mu(a)}}
\end{aligned}
$$

[12]This m is the maximized value of the maximization problem defining $T^\mu(h)$. Although it would have been natural to denote it by $m^\mu(h)$, we chose the shorter notation because both μ and h are fixed for the remainder of this proof.

$$
\begin{aligned}
(20\text{e}) \quad &= \frac{\Pi_{a\in t}(\rho\cup\lambda)(a) \times \lim_n \, n^{-m+\Sigma_{a\in t\cap A^s}\mu(a)}}{\Sigma_{t'\in h}\, \Pi_{a\in t'}(\rho\cup\lambda)(a) \times \lim_n \, n^{-m+\Sigma_{a\in t'\cap A^s}\mu(a)}} \\
(20\text{f}) \quad &= \frac{\Pi_{a\in t}(\rho\cup\lambda)(a) \times \lim_n \, n^{-m+\Sigma_{a\in t\cap A^s}\mu(a)}}{\Sigma_{t'\in T^\mu(h)}\, \Pi_{a\in t'}(\rho\cup\lambda)(a)} \\
(20\text{g}) \quad &= \begin{pmatrix} \dfrac{\Pi_{a\in t}\,(\rho\cup\lambda)(a)}{\Sigma_{t'\in T^\mu(h)}\,\Pi_{a\in t'}\,(\rho\cup\lambda)(a)} & \text{if } t \in T^\mu(h) \\ 0 & \text{if } t \notin T^\mu(h) \end{pmatrix} \\
(20\text{h}) \quad &= \beta(t)\ .
\end{aligned}
$$

(20a) holds by the definition of σ_n. (20b) and (20c) hold by algebraic manipulation. (20d) is an application of the quotient rule for limits. In particular, the limit of the numerator is

$$\Pi_{a\in t\cap A^c}\rho(a) \times \Pi_{a\in t\cap A^s}\lambda(a) \times \lim_n \, n^{-m+\Sigma_{a\in t\cap A^s}\mu(a)}$$

by the constancy of $\Pi_{a\in t\cap A^c}\rho(a)$ and by (19). This limit is finite by the definition of m. Similarly, the limit of each of the $|h|$ terms in the denominator is

$$\Pi_{a\in t'\cap A^c}\rho(a) \times \Pi_{a\in t'\cap A^s}\lambda(a) \times \lim_n \, n^{-m+\Sigma_{a\in t'\cap A^s}\mu(a)}$$

by the constancy of $\Pi_{a\in t\cap A^c}\rho(a)$ and by (19) applied at $t = t'$. By the definition of m, this limit is finite for every $t'\in h$ and positive for at least one $t'\in h$. (20e) holds by algebraic manipulation. (20f) holds because $-m+\Sigma_{a\in t'\cap A^s}\mu(a)$ is zero for every $t'\in T^\mu(h)$ and negative for every $t'\in h\smallsetminus T^\mu(h)$. Similarly, (20g) holds because $-m+\Sigma_{a\in t\cap A^s}\mu(a)$ is zero if $t\in T^\mu(h)$ and negative if $t\in h\smallsetminus T^\mu(h)$. (20h) is the definition of β. □

Proof A.8 (for Lemma 5.1)**.** Take any b and μ, and let $\succcurlyeq$ be the plausibility relation of b. We will show that representation (4) is equivalent to indication (7).

Necessity of (7). (7a), (7b), and (7c) follow from Lemma 4.4. To see (7d), take any $h\in H^s$ and any $t^o\in h$. On the one hand, suppose $t^o\in h\smallsetminus b$. Since b is a basement, there is some $t^*\in h\cap b$. This implies $t^* \succ t^o$ by the definition of $\succ_{T^s}$, which implies $\Sigma_{a\in t^*}\mu(a) > \Sigma_{a\in t^o}\mu(a)$ by representation (4), which implies $t^o \notin \text{argmax}_{t'\in h}\, \Sigma_{a\in t'}\mu(a)$. On the other hand, suppose $t^o\in h\cap b$. Then consider any other $t\in h$. By the definitions of $\succ_{T^s}$ and $\approx_{T^s}$, either $t^o \succ_{T^s} t$ or $t^o \approx_{T^s} t$, and thus in either event we have $t^o \succcurlyeq t$. Hence by representation (4), $\Sigma_{a\in t^o}\mu(a) \geq$

$\Sigma_{a\in t}\mu(a)$. Since this has been demonstrated for any $t\in h$, we have that $t^o \in \text{argmax}_{t\in h}\,\Sigma_{a\in t}\mu(a)$.

Sufficiency of (7). (a) Take any t^1 and t^2 such that $t^1 \succ t^2$. By the definition of $\succ$, we have $t^1 \succ_{A^s} t^2$ or $t^1 \succ_{T^s} t^2$.

In the first case, there is an $a\in F(t^1)$ such that $t^1\cup\{a\}=t^2$ and $a\in A^s\smallsetminus b$. Note that $\{a\} = t^2\smallsetminus t^1$ because $t^1\cap\{a\}=t^2$ and because $a\notin t^1$ by $a\in F(t^1)$ and the definition of F. Therefore $\Sigma_{a'\in t^1}\mu(a') > \Sigma_{a'\in t^2}\mu(a')$ because $\mu(a)<0$ by $a\in A^s\smallsetminus b$ and (7c).

In the second case, there is an $h\in H^s$ such that $t^1\in h\cap b$ and $t^2\in h\smallsetminus b$. Thus $\Sigma_{a\in t^1}\mu(a) = \max_{t'\in h}\Sigma_{a\in t'}\mu(a) > \Sigma_{a\in t^2}\mu(a)$ by two applications of (7d).

(b) Take any t^1 and t^2 such that $t^1 \approx t^2$. By the definition of $\approx$, we have $t^1 \approx_{A^s} t^2$, or $t^1 \approx_{A^c} t^2$, or $t^1 \approx_{T^s} t^2$.

In the first case, there is an $a\in A^s\cap b$ such that either [a] $a\in F(t^1)$ and $t^1\cap\{a\}=t^2$, or symmetrically [b] $a\in F(t^2)$ and $t^2\cap\{a\}=t^1$. In subcase [a], $\Sigma_{a\in t^1}\mu(a) = \Sigma_{a\in t^1\cup\{a\}}\mu(a) = \Sigma_{a\in t^2}\mu(a)$, where the first equality holds because $\mu(a)=0$ by $a\in A^s\cap b$ and (7b), and where the second equality holds by $t^1\cup\{a\}=t^2$. Subcase [b] can be treated symmetrically.

In the second case, there is an $a\in A^c$ such that either [a] $a\in F(t^1)$ and $t^1\cap\{a\}=t^2$, or symmetrically [b] $a\in F(t^2)$ and $t^2\cap\{a\}=t^1$. In subcase [a], $\Sigma_{a\in t^1}\mu(a) = \Sigma_{a\in t^1\cup\{a\}}\mu(a) = \Sigma_{a\in t^2}\mu(a)$, where the first equality holds because $\mu(a)=0$ by $a\in A^c$ and (7a), and where the first equality holds by $t^1\cup\{a\}=t^2$. Subcase [b] can be treated symmetrically.

In the third case, there is an $h\in H^s$ such that $t^1\neq t^2$ and $\{t^1,t^2\}\subseteq h\cap b$. Thus $\Sigma_{a\in t^1}\mu(a) = \max_{t'\in h}\Sigma_{a\in t'}\mu(a) = \Sigma_{a\in t^2}\mu(a)$ by two applications of (7d). □

Proof A.9 (for Lemma 5.2)**.** Take any b and μ. By the definitions of indication and labelling, we are to show that the combination of $\mu(A)\subseteq\mathbb{Z}$ and (7) is equivalent to the combination of

$$\mu(A^c) = \{0\}\ , \tag{21a}$$

$$-\mu|_{A^s}{:}A^s\to\mathbb{Z}_+\ , \tag{21b}$$

$$(\forall a\in A^s)\ \ a\in b\ \text{ iff }\ -\mu(a)=0\ ,\ \text{and} \tag{21c}$$

$$(\forall h\in H^s)(\forall t\in h)\ \ t\in b\ \text{ iff }\ t\in \text{argmin}_{t'\in h}\,\Sigma_{a\in t'}-\mu(a)\ . \tag{21d}$$

Necessity of (21). Assume $\mu(A)\subseteq\mathbb{Z}$ and (7). (21a) follows from (7a). (21b) follows from $\mu(A)\subseteq\mathbb{Z}$ and (7b&c). (21c) follows from (7b&c). (21d) follows from (7d).

Sufficiency of (21). Assume (1). $\mu(A) \subseteq \mathbb{Z}$ follows from (7a&b). (7a) follows from (21a). (7b) follows from (21c). (7c) follows from (21b&c). (7d) follows from (21d). □

Proof A.10 (for Lemma 5.3)**.** Take b and μ. By the definitions of indication and justification, we are to show that (7) is equivalent to the combination of

$$\mu(A^c) = \{0\} \ , \tag{22a}$$

$$\mu(A^s \cap b) = \{0\} \ , \tag{22b}$$

$$\mu(A^s \smallsetminus b) \subseteq \mathbb{R}_{++} \ , \text{ and} \tag{22c}$$

$$(\forall h \in H^s)(\forall t \in h) \ \ t \in b \ \text{ iff } \ t \in \mathrm{argmax}_{t' \in h} \, \Pi_{a \in t' \cap (A^s \smallsetminus b)} e^{\mu(a)} \ . \tag{22d}$$

(7a–c) is identical to (22a–c). Further, (7d) is equivalent to (22d) because for any $h \in H^s$

$$\begin{aligned}
& \mathrm{argmax}_{t' \in h} \, \Pi_{a \in t' \cap (A^s \smallsetminus b)} e^{\mu(a)} \\
= \ & \mathrm{argmax}_{t' \in h} \ln\big(\Pi_{a \in t' \cap (A^s \smallsetminus b)} e^{\mu(a)} \big) \\
= \ & \mathrm{argmax}_{t' \in h} \, \Sigma_{a \in t' \cap (A^s \smallsetminus b)} \mu(a) \\
= \ & \mathrm{argmax}_{t' \in h} \big(\Sigma_{a \in t' \cap A^c} \mu(a) + \Sigma_{a \in t' \cap (A^s \cap b)} \mu(a) + \Sigma_{a \in t' \cap (A^s \smallsetminus b)} \mu(a) \big) \\
= \ & \mathrm{argmax}_{t' \in h} \, \Sigma_{a \in t'} \mu(a) \ .
\end{aligned}$$

The first equality holds by monotonicity and the second by algebraic manipulation. The third holds because both $\mu(A^c)$ and $\mu(A^s \cap b)$ are $\{0\}$ by (22a&b) (which is identical to (7a&b) by the first sentence of this paragraph). Finally, the fourth holds because $\{A^c, A^s \cap b, A^s \smallsetminus b\}$ partitions A. □

REFERENCES

BLUME, L., A. BRANDENBURGER, AND E. DEKEL (1991): "Lexicographic Probabilities and Equilibrium Refinements," *Econometrica*, 59, 81–98.

DEBREU, G. (1960): "Topological Methods in Cardinal Utility Theory," in *Mathematical Methods in the Social Sciences, 1959*, ed. by K. J. Arrow, S. Karlin, and P. Suppes, pp. 16–26. Stanford University Press, reprinted in Debreu (1983).

——— (1983): *Mathematical Economics: Twenty Papers of Gerard Debreu*. Cambridge University Press.

FARKAS, J. (1902): "Über die Theorie der einfachen Ungleichungen," *J. Math.*, 124, 1–24.

HALPERN, J. Y. (2010): "Lexicographic probability, conditional probability, and nonstandard probability," *Games and Economic Behavior*, 68, 155–179.

Hammond, P. J. (1994): "Elementary Non-Archimedean Representations of Probability for Decision Theory and Games," in *Patrick Suppes: Scientific Philosopher, Volume 1*, ed. by P. Humphreys, pp. 25–61. Kluwer.

Kohlberg, E., and P. J. Reny (1997): "Independence on Relative Probability Spaces and Consistent Assessments in Game Trees," *Journal of Economic Theory*, 75, 280–313.

Kraft, C. H., J. W. Pratt, and A. Seidenberg (1959): "Intuitive Probability on Finite Sets," *Annals of Mathematical Statistics*, 30, 408–419.

Krantz, D. H., R. D. Luce, P. Suppes, and A. Tversky (1971): *Foundations of Measurement, Volume I: Additive and Polynomial Representations.* Academic Press.

Kreps, D. M., and G. Ramey (1987): "Structural Consistency, Consistency, and Sequential Rationality," *Econometrica*, 55, 1331–1348.

Kreps, D. M., and R. Wilson (1982): "Sequential Equilibria," *Econometrica*, 50, 863–894.

McLennan, A. (1989): "Consistent Conditional Systems in Noncooperative Game Theory," *International Journal of Game Theory*, 18, 141–174.

Myerson, R. B. (1986): "Multistage Games with Communication," *Econometrica*, 54, 323–358.

Narens, L. (1985): *Abstract Measurement Theory.* MIT Press.

Osborne, M. J., and A. Rubinstein (1994): *A Course in Game Theory.* MIT Press.

Perea, A. (2001): *Rationality in Extensive Form Games.* Kluwer.

Perea y Monsuwé, A., M. Jansen, and H. Peters (1997): "Characterization of Consistent Assessments in Extensive Form Games," *Games and Economic Behavior*, 21, 238–252.

Piccione, M., and A. Rubinstein (1997): "On the Interpretation of Decision Problems with Imperfect Recall," *Games and Economic Behavior*, 20, 3–24.

Rényi, A. (1955): "On a New Axiomatic Theory of Probability," *Acta Mathematica Academiae Scientiarum Hungaricae*, 6, 285–335.

Robinson, A. (1973): "Function Theory on Some Nonarchimedean Fields," *American Mathematical Monthly: Papers in the Foundation of Mathematics*, 80, S87–S109.

Scott, D. (1964): "Measurement Structures and Linear Inequalities," *Journal of Mathematical Psychology*, 1, 233–247.

Streufert, P. A. (2006): "A Comment on 'Sequential Equilibria'," University of Western Ontario, Economics Department Research Report 2006-03, 10 pages.

——— (2012): "Specifying Nodes as Sets of Actions," University of Western Ontario, Department of Economics Research Report 2012-2, October, 34 pages.

Suzumura, K. (1976): "Remarks on the Theory of Collective Choice," *Economica*, 43, 381–390.

Vohra, R. V. (2005): *Advanced Mathematical Economics.* Routledge.