CONSISTENCY IMPLIES THAT PLAUSIBILITY HAS A MASS FUNCTION

Peter A. Streufert
Department of Economics
University of Western Ontario

Abstract. We formulate and prove a new fundamental theorem about Kreps-Wilson consistency. First, we derive from an arbitrary assessment its implied plausibility (i.e. infinite relative likelihood) relation among the game's nodes. Typically such a plausibility relation is incomplete (i.e. fails to compare all nodes). Second, since nodes can be specified as sets of actions via Streufert (2012a), we introduce the concept of representing a completion of a plausibility relation by the nodal sums of a mass (i.e. density) function assigning plausibility numbers to the game's actions. Finally, we discover that the consistency of an assessment implies that its plausibility relation has a completion represented by a mass function. We prove this by re-using math from the early foundations of probability theory.

This theorem leads to a number of corollaries. First, we are able to formalize in two new ways that consistency specifies that zero-probability agents reason that past zero-probability actions were played independently. Second, we identify and repair a nontrivial gap in a Kreps-Wilson proof and then clarify two algebraic (i.e. non-topological) characterizations of consistency from the literature. Third, we discover a particularly simple characterization of consistency for degenerate-support (i.e. pure-strategy and sure-belief) assessments. All the paper's proofs are accessible in that they require nothing more than linear algebra.

Date: August 8, 2012.

I thank Walter Bossert, Audra Bowlus, Jean Guillaume Forand, Ian King, Val Lambson, Bart Lipman, Andrew McLennan, Al Slivinski, Bernhard von Stengel, Bob Wilson, and Charles Zheng; seminar participants at Auckland, Brigham Young, the London School of Economics, New South Wales, Sydney, and Western Ontario; and conference participants at the Econometric Society meeting in Sydney, the game theory conference in Stony Brook, and the Midwest mathematical economics meeting in Ann Arbor.

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 (i.e. the information set's nodes). In many equilibrium concepts, the agent chooses its strategy optimally, given its belief, the strategies of subsequent agents, and its 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 difficult to model the reasoning by which a zero-probability agent calculates its belief. Rather, we must awkwardly grapple with what an agent would think after discovering that one of its unreachable nodes was actually reached.

This problem is important. In order for there to be a zero-probability agent, prior agents must have chosen against actions that led to the nodes in the zero-probability agent. If those prior agents were optimizing, they must have contemplated what the zero-probability agent would do if it were actually reached. In other words, the prior agents must have known the strategy of the zero-probability agent. Alternatively and more deeply, if the prior agents knew the zero-probability agent's belief, they might have calculated its strategy. Or, even more deeply, if the prior agents knew how the zero-probability agent would reason, they might have calculated both its belief and its strategy.

Although the terminology of game theory is arcane, and the logic of zero-probability events is subtle, the issue itself is very familiar. For example, if a child is deterred from swiping a cookie from the cookie jar, then he must be thinking about what his parent would think and do in the zero-probability event that a cookie disappears.

Game theorists can choose to address this issue with various degrees of sophistication. The least sophisticated approach is to arbitrarily posit a strategy for the zero-probability agent and to be unconcerned about whether that agent is optimizing. This is implicitly done by the Nash equilibria of the strategic form derived from the extensive-form game. A more sophisticated approach is to arbitrarily posit a belief for the zero-probability agent and then to require that that agent is optimizing. This is the approach of weak perfect Bayesian equilibria.

A still more sophisticated approach is to posit 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. That, essentially, is the topological definition of consistency in Kreps and Wilson (1982). 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).

Although the topological definition of consistency is natural, one might seek to understand consistency without reference to a converging sequence of full-support strategies. Indeed, the literature discussed below has found this to be both conceptually interesting and computationally useful. This paper contributes to that endeavour by studying consistency from a new perspective which resembles the early foundations of ordinary probability theory.

The first step toward our new perspective is to define the "plausibility" (i.e., infinite relative likelihood) relation $\succcurlyeq$ of an arbitrary assessment. This binary relation compares nodes, but does so only in two circumstances. (1) When two nodes belong to the same agent, the two are equally plausible if both are in the support of the agent's belief, and the first is more plausible than the second if the first is in the support while the second is not. (2) When one node immediately precedes another, the two are equally plausible if the intervening action is played with positive probability, and the first is more plausible than the second if the intervening action is played with zero probability. This construction's novelty is modest but nontrivial. It differs from the infinite relative likelihoods of a conditional probability system in that (a) it is derived from an arbitrary assessment rather than a consistent assessment, (b) it is easier to derive because it does not require deriving a conditional probability system, and (c) it is incomplete and intransitive because it only compares nodes in a limited number of circumstances.

Next, we notice that Streufert (2012a, Theorem 1) allows one to specify each node in a game tree as the *set* of actions leading to it rather than the sequence of actions leading to it. Accordingly, the plausibility relation $\succcurlyeq$ can be understood to compare sets of actions.

At this point, we are surprised by a deep analogy with ordinary probability over a finite state space. An action resembles a state. A node, which is a set of actions, resembles an event, which is a set of states. Further, a plausibility (i.e. infinite relative likelihood) relation $\succcurlyeq$, which conveys when one node is least as plausible as another, resembles a so-called "probability relation", which conveys when one event is at least as probable as another.

The fundamental work of Kraft, Pratt, and Seidenberg (1959) and Scott (1964) showed that a well-behaved probability relation has a completion that is represented by a mass (i.e. density) function. In detail, a mass function assigns probability numbers to states. Then the probability of an event can be calculated as the sum of the probability numbers assigned to its states. Finally, the probabilities of all the events represent a completion of the original probability relation, which might or might not have been complete. In hindsight, the mathematics is easy: only Farkas' Lemma for rational numbers is required (and Farkas' Lemma itself can be derived in two pages from undergraduate linear algebra).

We re-use this math. In particular, Theorem 1 shows that the consistency of an assessment implies that its plausibility relation has a completion that is represented by a "mass function" which resembles a probability mass function. In detail, a mass function π assigns plausibility numbers to actions. Then the plausibility of a node can be calculated as the sum of the plausibility numbers assigned to its actions. Finally, the plausibilities of all the nodes represent a completion of the assessment's plausibility relation.

However, the resemblance between a probability mass function and a plausibility mass function is imperfect. In probability theory, an event is at least as probable as any of its subsets, and thus, a probability mass function is nonnegative-valued. In contrast, a node is at *most* as plausible as any of its subnodes (i.e. predecessors), and thus, a plausibility mass function is *nonpositive*-valued. More precisely, representation and part (2) in the above definition of $\succcurlyeq$ together require that each positive-probability action be given zero plausibility and that each zero-probability action be given 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 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 each zero-probability action can be assigned its own negative plausibility number.

The existence of a plausibility mass function has both conceptual and computational implications. We discuss each of these two categories in turn.

Conceptually, we saw in the opening paragraphs that consistency specifies the reasoning by which a zero-probability agent calculates its belief. In particular, consistency supposes that a zero-probability agent (1) posits some converging sequence of full-support strategies for each of the other agents, (2) calculates its own belief at each point in this sequence by the conditional-probability law, and (3) takes the limit of this sequence of beliefs. Implicit in steps (1) and (2) is the zero-probability agent's assumption that the posited full-support strategies for the other players are stochastically independent of one another in the usual sense. Thus step (3) suggests that the limiting belief should inherit something similar.

One naturally hopes to understand this limiting sort of independence directly, that is, without reference to a converging sequence of full-support strategies. This is subtle. If a consistent assessment specifies pure strategies, it is difficult to even define the sense in which these degenerate distributions could be independent without relying on converging full-support distributions. Nonetheless, much progress has been made. Blume, Brandenburger, and Dekel (1991a, Section 7), Hammond (1994, Section 6.5), and Halpern (2010, Section 6) all formulate independence in terms of non-Archimedean probability numbers.

Battigalli (1996, Section 2) formulates independence in terms of conditional probability systems. Kohlberg and Reny (1997, Section 2) formulate independence in terms of relative probability systems.

This paper's Theorem 1 offers a complementary alternative whose mathematics is more accessible. In particular, the additive functional form of mass-function representation can be seen to specify that a zero-probability agent reasons that past zero-probability actions were played independently of other past actions. In somewhat more detail, representation and part (1) in the above definition of $\succcurlyeq$ together require that the support of a zero-probability agent's belief must be the set consisting of the agent's most plausible node(s). The plausibility of each node is the sum of the plausibilities of the node's actions, and the mass function π implicitly requires that the plausibility of a zero-probability action cannot vary with anything but the action itself. In particular, it cannot vary with other (past) actions. That invariance is essentially how a mass function formalizes independence. The text will illustrate this formalization with three examples.

Further, this paper provides a second new formalization of independence. Essentially, the independence within a mass function π represents an underlying independence within the plausibility relation $\succcurlyeq$. Corollary 2 shows that this underlying independence can be made analogous to the concept of additive separability from preference theory. The familiar independence across consumption goods becomes independence across agents.

Computationally, the simplicity of mass functions leads to simple tests for the consistency of an assessment. In this regard, the contribution of Theorem 1 is to show that the existence of a plausibility mass function is a necessary condition for consistency. More precisely, it is a necessary condition on the supports of the strategies and beliefs of a consistent assessment.

Accordingly, numbers resembling plausibility numbers have appeared in the algebraic (i.e. non-topological) characterizations of consistency developed by Kreps and Wilson (1982) and Perea y Monsuwé, Jansen, and Peters (1997). We believe that these insightful characterizations should be much better understood and appreciated. Our introduction of plausibility relations and plausibility mass functions serves to clarify these characterizations and to substantially simplify their proofs. Further, the discussion following Corollary 3 reveals how Perea y Monsuwé, Jansen, and Peters (1997) is linked with Kreps and Wilson (1982), and our Corollary 4 fills a critical gap in a proof of Kreps and Wilson (1982).

Finally, Corollary 6 shows that the existence of a plausibility mass function is not only necessary but also sufficient for consistency when the assessment specifies strategies and beliefs with degenerate supports. In other words, Corollary 6 shows that the existence of a plausibility mass function characterizes consistency when every strategy is pure and every belief is sure. Although this new result does not begin to use the full strength of Theorem 1, it does provide a straightforward characterization of consistency that is applicable in many cases. Corollary 7 then employs Corollary 6 to provide an easy-to-use characterization of all sequential equilibria with pure strategies and sure beliefs.

This paper is organized as follows. Section 2 recapitulates old definitions from the literature, including that of a set-tree game and a consistent assessment. Section 3 provides new definitions, including that of a plausibility relation and a plausibility mass function. Section 4 states Theorem 1 and describes its close relation to the early foundations of probability theory.

Section 5 formalizes the independence of zero-probability actions via both mass functions and additive separability. Section 6 clarifies two earlier algebraic characterizations of consistency, and Section 7 shows that a degenerate-support assessment is consistent if and only if it admits a plausibility mass function. Section 8 concludes.

2. Old Definitions

2.1. Reviewing set-tree games

This subsection specifies a finite game via the set-tree formulation of Streufert (2012a). 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 represent a completion of a plausibility relation. 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 (2012a).

By Streufert (2012a, 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 (2012a) 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).

Accordingly, recall from Streufert (2012a, Subsection 2.2) that a *set tree* (A,T) is a set A of *actions* a and a collection T of subsets of A such that $|T|{\geq}2$, such that $A = \bigcup T$, and such that 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 (2012a) 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 $Z = \{t|F(t){=}\varnothing\}$.

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 $T{\sim}Z$ of nonterminal nodes. (2) I is a collection of *players* i which *prepartition* H in the sense that the nonempty players of I partition 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) $u : (I{\sim}\{i^c\}){\times}Z \rightarrow \mathbb{R}$ specifies a payoff $u_i(t)$ to each nonchance player $i{\in}I{\sim}\{i^c\}$ 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. 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, every nonchance player is assumed to be nonempty. All of the above are discussed in Streufert (2012a, 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. This reformulation is not immediate because this is the first paper to consider 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) players. 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 $\{h|h{\notin}i^c\}$ of strategic agents. Then since the set $T{\sim}Z$ of nonterminal nodes is partitioned by H, we can prepartition $T{\sim}Z$ into the possibly empty set T^c of chance nodes and the necessarily nonempty set T^s of strategic (i.e. decision) nodes:

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

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

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

Note that T^c, T^s, A^c, and A^s are derived from the given game, and that the definition of A^c allows us to write $\rho{:}A^c{\rightarrow}(0,1]$ 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 Appendix A).

Second, we introduce notation for strategies, beliefs, and assessments. A (behavioural) *strategy profile* is a function $\sigma{:}A^s{\rightarrow}[0,1]$ such that $(\forall h{\notin}i^c)\ \Sigma_{a\in F(h)}\sigma(a){=}1$. 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 $(\forall h{\notin}i^c)\ \Sigma_{t\in h}\beta(t){=}1$. 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. Note that

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

is the probability of reaching node t. 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. 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.

3. New Definitions

3.1. Defining an Assessment's Plausibility Relation $\succcurlyeq$

This subsection defines the plausibility relation $\succcurlyeq$ of an arbitrary assessment (σ, β). The assessment (σ, β) need not be consistent. The relation $\succcurlyeq$ compares nodes, and it is constructed from five components, in the five paragraphs that follow the next two. The novelty of this definition relative to the literature is modest but nontrivial, as this subsection's last paragraph will explain. We introduce the word "plausibility" in lieu of the familiar phrase "infinite relative likelihood" only because it is shorter.

We illustrate this construction by repeatedly referring to Figure 1. This figure's 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

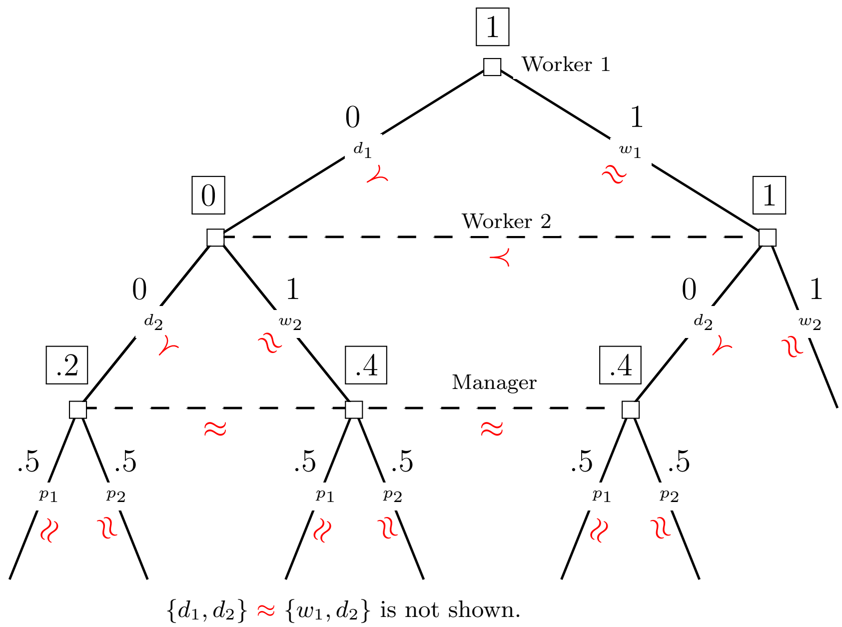

Figure 1. A plausibility relation $\succcurlyeq$ which does not have a transitive completion.

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 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.

As promised, the next five paragraphs define the plausibility relation of an arbitrary assessment. First, from the strategy profile σ derive the relation

$$\overset{\sigma}{\succ} \;=\; \{\ (t, t{\cup}\{a\}) \mid t{\in}T^s,\ a{\in}F(t),\ \text{and}\ \sigma(a){=}0\ \}\ .$$

As with any relation, the notations $(t^1, t^2) \in \overset{\sigma}{\succ}$ and $t^1 \overset{\sigma}{\succ} t^2$ are equivalent. Thus the definition of $\overset{\sigma}{\succ}$ says that $t^1 \overset{\sigma}{\succ} t^2$ iff t^1 immediately precedes t^2 and the action leading from t^1 to t^2 is played by σ with zero probability. In such a case, we say that t^1 is "more plausible" than t^2 in the sense that t^1 is infinitely more likely than t^2. For example, $\{\} \overset{\sigma}{\succ} \{d_1\}$, $\{d_1\} \overset{\sigma}{\succ} \{d_1,d_2\}$, and $\{w_1\} \overset{\sigma}{\succ} \{w_1,d_2\}$ in Figure 1 (as in any set tree, a node is identified with the set of actions leading to it).

Second, from the strategy profile σ derive the relation

$$\begin{aligned} \overset{\sigma}{\approx} \;=\; & \{\ (t, t{\cup}\{a\}) \mid t{\in}T^s,\ a{\in}F(t),\ \text{and}\ \sigma(a){>}0\ \} \\ & \cup \{\ (t{\cup}\{a\}, t) \mid t{\in}T^s,\ a{\in}F(t),\ \text{and}\ \sigma(a){>}0\ \}\ . \end{aligned}$$

The definition of $\overset{\sigma}{\approx}$ states that both $t^1 \overset{\sigma}{\approx} t^2$ and $t^2 \overset{\sigma}{\approx} t^1$ hold if t^1 immediately precedes t^2 and the action leading from t^1 to t^2 is played by σ with positive probability. In such a case, we say that t^1 and t^2 are "tied in plausibility" in the sense that neither can be infinitely more likely than the other. For example, Figure 1 shows $\{\} \overset{\sigma}{\approx} \{w_1\}$, $\{w_1\} \overset{\sigma}{\approx} \{w_1,w_2\}$, $\{d_1\} \overset{\sigma}{\approx} \{d_1,w_2\}$, $\{d_1,d_2\} \overset{\sigma}{\approx} \{d_1,d_2,p_1\}$, and five other pairs like the last one which also end in terminal nodes. (The converses of these nine pairs are also in $\overset{\sigma}{\approx}$ because $\overset{\sigma}{\approx}$ was defined to be symmetric.)

Third, this notion of tying in plausibility applies not only to strategic actions, but also to chance actions, which are played with positive probability by assumption. Accordingly, we define the relation

$$\begin{aligned} \overset{c}{\approx} \;=\; & \{\ (t, t{\cup}\{a\}) \mid t{\in}T^c\ \text{and}\ a{\in}F(t)\ \} \\ & \cup \{\ (t{\cup}\{a\}, t) \mid t{\in}T^c\ \text{and}\ a{\in}F(t)\ \}\ . \end{aligned}$$

Thus both $t^1 \overset{c}{\approx} t^2$ and $t^2 \overset{c}{\approx} t^1$ hold if t^1 is a chance node that immediately precedes t^2. Unlike the other components of $\succcurlyeq$, $\overset{c}{\approx}$ depends only on the game and not the assessment.

Fourth, from the belief system β derive the two relations

$$\begin{aligned} \overset{\beta}{\succ} \;&=\; \{\ (t^1,t^2) \mid (\exists h{\in}H^s)\ \{t^1,t^2\}{\subseteq}h,\ \beta(t^1){>}0,\ \text{and}\ \beta(t^2){=}0\ \}\ \ \text{and} \\ \overset{\beta}{\approx} \;&=\; \{\ (t^1,t^2) \mid (\exists h{\in}H^s)\ \{t^1,t^2\}{\subseteq}h,\ t^1{\neq}t^2,\ \beta(t^1){>}0,\ \text{and}\ \beta(t^2){>}0\ \}\ . \end{aligned}$$

Thus a node in the support of an agent's belief is more plausible than any node outside the support and is tied with any other node inside the support. For example, Figure 1 shows $\{w_1\} \overset{\beta}{\succ} \{d_1\}$, $\{w_1,d_2\} \overset{\beta}{\approx} \{d_1,w_2\}$, and $\{d_1,w_2\} \overset{\beta}{\approx} \{d_1,d_2\}$. (The relation $\overset{\beta}{\approx}$ also contains $(\{w_1,d_2\},\{d_1,d_2\})$, and because the relation is symmetric, the converses of the three pairs already mentioned.)

Fifth and finally, we define $\succ$, $\approx$, and $\succcurlyeq$. Let $\succ$ be the union of $\overset{\sigma}{\succ}$ and $\overset{\beta}{\succ}$. Let $\approx$ be the union of $\overset{\sigma}{\approx}$, $\overset{c}{\approx}$, and $\overset{\beta}{\approx}$. Let $\succcurlyeq$ be the union of $\succ$ and $\approx$. The following result is intuitive but not obvious.

Lemma 3.1. *Suppose that* $\overset{\sigma}{\succ}$, $\overset{\sigma}{\approx}$, $\overset{c}{\approx}$, $\overset{\beta}{\succ}$, $\overset{\beta}{\approx}$, $\succ$, $\approx$, *and* $\succcurlyeq$ *are derived from some assessment. Then* $\succ$ *is the asymmetric part of* $\succcurlyeq$, *and* $\succ$ *is partitioned by* $\{\overset{\sigma}{\succ}, \overset{\beta}{\succ}\}$. *Similarly,* $\approx$ *is the symmetric part of* $\succcurlyeq$, *and* $\approx$ *is partitioned by* $\{\overset{\sigma}{\approx}, \overset{c}{\approx}, \overset{\beta}{\approx}\}$. *(Proof A.2 in Appendix A.)*

The typical plausibility relation $\succcurlyeq$ is pervasively incomplete in the sense that it fails to compare many pairs of nodes. For instance, neither $\{\} \succcurlyeq \{d_1, d_2\}$ nor $\{d_1, d_2\} \succcurlyeq \{\}$ in Figure 1's example. In general, if $|T| \geq 3$, there must be at least one pair of nodes such that (a) the two are not in the same agent and (b) neither is an immediate predecessor of the other. No plausibility relation can compare such a pair of nodes.

Further, because of this pervasive incompleteness, the typical $\succcurlyeq$ is also intransitive. For instance, transitivity is violated by the lack of $\{\} \succcurlyeq \{d_1, d_2\}$ in Figure 1's example. Such intransitivities always occur when one agent follows another.

Our $\succcurlyeq$ differs from the infinite-relative-likelihood relations in the literature because it is derived *directly* from an *arbitrary* assessment. In contrast, most contributions in the literature have shown that a consistent assessment implies the existence of a rich probability structure which features infinite relative likelihoods. Such 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 (1991b); the nonstandard probability systems of Hammond (1994), Govindan and Klumpp (2002), and Halpern (2010); and the relative probability systems of Kohlberg and Reny (1997). 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 bypasses their rich probability structures. (c) Our $\succcurlyeq$ is incomplete and intransitive while their infinite relative probabilities are complete and transitive.[1]

[1] 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.

3.2. Introducing mass functions

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

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

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

Since T is finite, the existence of a transitive completion is equivalent to the existence of a function $\varphi{:}T{\rightarrow}\mathbb{R}$ which *represents*[2] *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 1's plausibility relation cannot be completed transitively 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 $\pi{:}A{\rightarrow}\mathbb{R}$ whose nodal sums represent a completion of $\succcurlyeq$ in the sense that for all t^1 and t^2

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

For brevity, we will often omit mentioning the nodal sums $\Sigma_{a\in t}\pi(a)$ and simply say that such a π *represents a completion of* $\succcurlyeq$. For example, Figure 2's plausibility relation $\succcurlyeq$ has a completion that is represented by the nodal sums of a function π. Such a function $\pi{:}A{\rightarrow}R$ is given by the numbers without boxes that appear over the actions, and its nodal sums $\Sigma_{a\in t}\pi(a)$ are given by the numbers within boxes that appear over the nodes. If brevity were important, as it often is, we would

[2] 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.

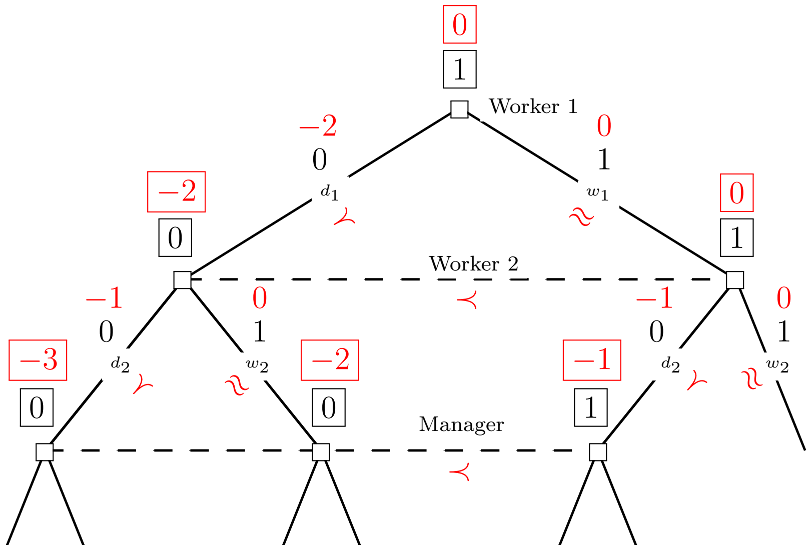

Figure 2. A plausibility relation $\succcurlyeq$ which has a completion represented by a mass function π.

omit mentioning the nodal sums, and simply say that the figure's π represents a completion of the figure's $\succcurlyeq$.

Comparing (4) with (3) shows that having a completion represented by a π is equivalent to having a completion represented by a $\varphi(t) = \Sigma_{a\in t}\pi(a)$ with a special functional form. Thus having a completion represented by a π implies having a completion represented by a φ, and this, as observed earlier, is equivalent to having a transitive completion. Therefore, having a completion represented by a π implies having a transitive completion. Figure 1's $\succcurlyeq$ has no transitive completion (and hence no completion represented by a π). Figure 2's $\succcurlyeq$ has a completion represented by a π (and hence a transitive completion). Later, Figure 3 will illustrate the last contingency. Its $\succcurlyeq$ has a transitive completion but not a completion represented by a π.

Because of its resemblance to a probability mass function, we call a function $\pi{:}A{\rightarrow}\mathbb{R}$ a *mass function*. To explore this resemblance in detail, suppose that $p{:}\varOmega{\rightarrow}[0,1]$ is a probability mass function defined on a finite set $\varOmega$ of states ω (some years ago p might have been called a discrete probability "density" function). Then, as we all know, the probability of any event $e \subseteq \varOmega$ 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 $\pi{:}A \to \mathbb{R}$ 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 $\Sigma_{a \in t}\pi(a)$ 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 π be nonnegative or that they sum to one.

Further, it is natural to call π a *plausibility* mass function, to call $\pi(a)$ the *plausibility* of the action a, and to call $\Sigma_{a \in t}\pi(a)$ the *plausibility* of the node t.

Finally, we digress to make a simplifying observation. A keen eye will notice that Figure 2 suppresses terminal nodes. This is done because terminal nodes have no bearing on the existence of a plausibility mass function. This observation is intuitive because a plausibility relation only compares terminal nodes to their immediate predecessors. Yet the observation is not obvious because an action, such as w_2 in Figure 2, may lead to both a terminal node and a nonterminal node. Accordingly, Lemma B.1 in Appendix B rigorously formulates and proves that terminal nodes are irrelevant to the existence of a plausibility mass function.

4. Theorem

4.1. Re-using the foundations of probability theory

Theorem 1. *Suppose an assessment is consistent. Then its plausibility relation has a completion represented by a mass function. Further, the mass function can be made to assume integer values. (Proof 4.2 below.)*

This theorem closely resembles a well-known result from the early foundations of ordinary 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 $\gtrsim$ 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 $\gtrsim$ which imply the existence of a probability mass function $p{:}\Omega \to [0,1]$ such that for all e^1 and e^2

$$e^1 \succ e^2 \ \Rightarrow\ \Sigma_{\omega\in e^1} p(\omega) > \Sigma_{\omega\in e^2} p(\omega) \text{ and}$$
$$e^1 \succsim e^2 \ \Rightarrow\ \Sigma_{\omega\in e^1} p(\omega) \geq \Sigma_{\omega\in e^2} p(\omega)\ .$$

In the terminology of the previous subsection, they state conditions on $\succsim$ which imply that $\succsim$ has a completion represented by a probability mass function $p{:}\Omega{\to}[0,1]$. Theorem 1 is surprisingly 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 by a mass function $\pi{:}A{\to}\mathbb{R}$ if for all s and t

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

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 by a mass function.[3] This result undergirds the foundations for probability in Kraft, Pratt, and Seidenberg (1959, Theorem 2) and

[3]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, a mass function π reduces to Subsection 3.2's representation φ (which is in turn equivalent to the existence of a transitive extension), and the cancellation law reduces to Suzumura's rule on cycles.

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).[4]

The following lemma is a very minor adaptation of that well-known result. Its proof requires nothing more than Farkas' Lemma (and that itself 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.

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 by a mass function $\pi{:}A{\to}\mathbb{Z}$. (Proof A.4 in Appendix A.)*

Proof 4.2 (for Theorem 1)**.** Let (σ,β) be a consistent assessment and let $\succcurlyeq$ be its plausibility relation.

This paragraph shows that $\succcurlyeq$ obeys the cancellation law. Accordingly, 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 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.5 at each (s^m,t^m), we have (note s^m is in the denominator)

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

[4] 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.

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.

Since $\succcurlyeq$ obeys the cancellation law, Lemma 4.1 implies that $\succcurlyeq$ has a completion represented by an integer-valued mass function. $\square$

4.2. A Digression

Theorem 1 shows that if an assessment is consistent, then its plausibility relation has a completion represented by a mass function. As discussed in Subsection 3.2, having a completion represented by a mass function implies having a transitive completion. Hence Corollary 1 follows easily from Theorem 1.

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

Corollary 1 provides a weak but easily tested necessary condition for consistency. For example, Subsection 3.2 showed that Figure 1's assessment does not have a transitive completion, and thus by Corollary 1 the assessment is inconsistent. (Corollary 1 does not use the full force of Theorem 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 spells this out in more detail. Although its proof must address a few technicalities, the essence is easily understood. Consider adding a new action a to a node t. This can be done exactly when $a \in F(t)$, and the resulting node is $t \cup \{a\}$. If a is a chance action, then (1) $t \approx t \cup \{a\}$ by the definition of $\overset{c}{\approx}$, which implies (2) $\Sigma_{a' \in t} \pi(a') = \Sigma_{a' \in t \cup \{a\}} \pi(a')$ by representation, which implies (3) that $\pi(a)$ must be zero. This is the lemma's part (a).

Alternatively, a might be a strategic action. On the one hand if $\sigma(a) > 0$, then (1) $t \approx t \cup \{a\}$ by the definition of $\overset{\sigma}{\approx}$, which implies (2) $\Sigma_{a' \in t} \pi(a') = \Sigma_{a' \in t \cup \{a\}} \pi(a')$ by representation, which implies (3) that $\pi(a)$ must be zero. This is the lemma's part (b). On the other hand if

$\sigma(a)=0$, then (1) $t \succ t \cup \{a\}$ by the definition of $\overset{\sigma}{\succ}$, which implies (2) $\Sigma_{a' \in t}\pi(a') > \Sigma_{a' \in t \cup \{a\}}\pi(a')$ by representation, which implies (3) that $\pi(a)$ must be negative. This is the lemma's part (c).

Lemma 4.3. *Suppose the plausibility relation $\succcurlyeq$ of (σ,β) has a completion represented by π. Then the following hold.*

(a) $$(\forall a \in A^c)\ \pi(a)=0\ .$$

(b) $$(\forall a \in A^s)\ \sigma(a)>0 \textit{ implies } \pi(a)=0\ .$$

(c) $$(\forall a \in A^s)\ \sigma(a)=0 \textit{ implies } \pi(a)<0\ .$$

These imply π is nonpositive-valued since $A = A^c \cup A^s$ and since σ is nonnegative-valued. (Proof A.6 in Appendix A.)

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 $\Sigma_{a \in t}\pi(a)$ is a measure of how far the node t is below the equilibrium path. This measure 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 each zero-probability action can be assigned its own negative plausibility number.

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

5. Independence

5.1. The additivity of mass-function representation

Here we argue that the mass function derived by Theorem 1 specifies two things about the reasoning of zero-probability agents. First, it specifies that zero-probability agents reason that the past agents who played zero-probability actions did so independently of the actions of the other past agents. This observation interprets the additive functional form of mass-function representation. Second, it specifies that the zero-probability agents have a common understanding of the plausibilities of zero-probability actions. This observation interprets the fact that the same mass function π applies to all agents. In summary, a mass function specifies that zero-probability agents share a common understanding of the independent plausibilities of past zero-probability actions.

The remainder of this subsection illustrates these observations with a number of examples.

First, consider the game tree of Figures 1 and 2. This game tree cannot illustrate commonality since there is at most one zero-probability agent. However, it can be used to illustrate independence. Toward that end, consider any assessment, such as that of Figure 1 or Figure 2, in which both workers are sure to work. In any such assessment, the manager is a zero-probability agent.

Now suppose that the assessment's plausibility relation $\succcurlyeq$ has a completion represented by a mass function π. By Lemma 4.3, both $\pi(w_1)$ and $\pi(w_2)$ are zero (since w_1 and w_2 are positive-probability actions) and both $\pi(d_1)$ and $\pi(d_2)$ are negative (since d_1 and d_2 are zero-probability actions). This readily implies that $\pi(d_1)+\pi(d_2) < \pi(d_1)+\pi(w_2)$ and hence, by the definition of representation (4), that $\{d_1, d_2\} \succcurlyeq \{d_1, w_2\}$ cannot be true. Thus, by the definition of the plausibility relation $\succcurlyeq$ (and of $\overset{\beta}{\succ}$ and $\overset{\beta}{\approx}$ in particular), we have that $\{d_1, d_2\}$ cannot be in the support of the manager's belief. (An symmetric alternative argument could have been constructed using w_1 rather than w_2.)

This is a substantial restriction on the reasoning of the manager. It precludes her from entertaining the possibility that the two workers may have coordinated their zero-probability actions. Rather, the additive functional form of mass-function representation specifies that she reasons that her workers play zero-probability actions independently of one another.

By Theorem 1, this is a consequence of consistency. Thus, since the manager included $\{d_1, d_2\}$ in the support of her belief in Figure 1's assessment, that assessment must be inconsistent (we also proved this earlier by Corollary 1 and the fact that this assessment's plausibility relation has no transitive completion). In contrast, the exclusion of

$\{d_1, d_2\}$ from the support of the manager's belief in Figure 2's assessment accords with, but does not imply, the consistency of Figure 2's assessment (its consistency will be proved later).

Second, consider Figure 3, which is closely related to Kohlberg and Reny (1997, Figure 7). To create a story for this figure's game tree, imagine that two sides in a dispute simultaneously choose left, middle, or right. If their choices agree, the game ends. If their choices disagree, a judge observes what choices have been made but not who made them. On the basis of this information, the judge chooses between two options (which can be left unnamed because Lemma B.1 shows that terminal nodes are irrelevant). Figure 3 provides symbols for the choices of the two sides and depicts the nonterminal nodes as empty squares. (As with any set tree, each node is identical to the set of actions taken to reach it.)

The numbers in Figure 3 specify an assessment. In particular, the numbers without boxes specify a strategy profile (the judges' strategies are left unspecified because terminal nodes are irrelevant by Lemma B.1), and the numbers within the boxes specify a belief system. The strategies specify that both sides choose middle. Thus play is sure to terminate at $\{m_1, m_2\}$, and all three judges are zero-probability agents.

The symbols $\succ$ and $\approx$ depict this assessment's plausibility relation. Every pair in $\succcurlyeq$ is shown, and thus the relation is pervasively incomplete, as is usual. Although the relation does have a transitive completion,[5] it does not have a completion represented by a mass function π. To see this formally, suppose π represented a completion of $\succcurlyeq$. Then

$$\pi(\ell_1) + \pi(m_2) > \pi(m_1) + \pi(\ell_2)$$

since $\{\ell_1, m_2\} \succ \{m_1, \ell_2\}$ from Judge LM's belief. Similarly,

$$\pi(m_1) + \pi(r_2) > \pi(r_1) + \pi(m_2)$$

since $\{m_1, r_2\} \succ \{r_1, m_2\}$ from Judge MR's belief. Adding these inequalities and cancelling like terms (which are zero anyway because m_1

[5]As discussed in Subsection 3.2, having a transitive completion is equivalent having a completion represented by a $\varphi{:}T \to \mathbb{R}$. One way is to construct such a φ is to assign -1 and -2 to Judge LM's nodes, and to assign -1 and -2 to Judge MR's nodes. These assignments imply that Side 2's nodes get -1, 0, and -2. Then arbitrarily assign -8 and -7 to Judge LR's nodes.

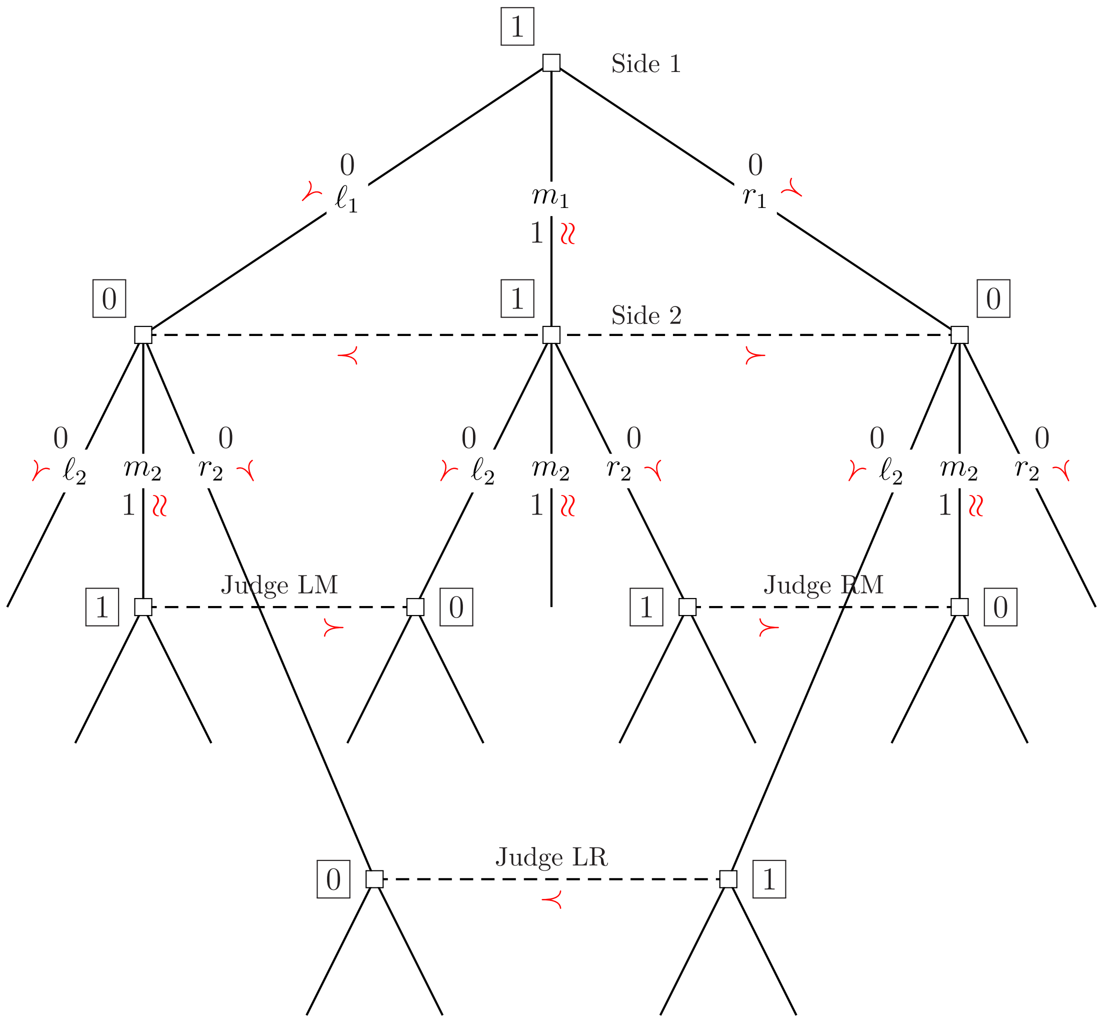

FIGURE 3. A plausibility relation $\succcurlyeq$ which has a transitive completion, but not a completion represented by a mass function π.

and m_2 are positive-probability actions) yields

$$\pi(\ell_1) + \pi(r_2) > \pi(r_1) + \pi(\ell_2) \ . \tag{7}$$

This contradicts $\{\ell_1, r_2\} \prec \{r_1, \ell_2\}$ from Judge LR's belief.

To see this intuitively, we proceed in a similar manner. First, Judge LM understands that neither side will actually choose left but nonetheless believes that left is more plausibly played by Side 1 rather than Side 2. Second, Judge MR understands that neither side will actually choose right but nonetheless believes that right is more plausibly played by Side 2 rather than Side 1. The beliefs of these two judges imply that the combination of left by 1 and right by 2 is more plausible the combination of left by 2 and right by 1. This contradicts the belief of Judge LR.

This example illustrates both independence and commonality. First consider the formal argument leading to (7). Commonality is implicit in the definition of mass-function representation (4). In particular, the same π is used to represent the different components of $\succcurlyeq$ that were derived from the beliefs of the different agents. Further, independence is also implicit in the definition of mass-function representation. In particular, (a) a mass function π assigns a plausibility number to each action and does not allow this number to be modified by the node at which the action is played, and (b) the plausibility numbers of the actions are added to derive the plausibility numbers of the nodes.

The intuitive argument two paragraphs above also illustrates this independence and commonality. Essentially, the intuitive argument combined Judge LM's belief that ℓ_1 is more plausible than ℓ_2, with Judge MR's belief that r_2 is more plausible than r_1, to conclude that Judge LR should believe that $\{\ell_1, r_2\}$ is more plausible than $\{\ell_2, r_1\} = \{r_1, \ell_2\}$. Clearly this relies upon the judges sharing a common understanding. To see independence, in very fine detail, note that this argument (a) uses Judge LM's belief about ℓ_2 played from $\{m_1\}$ to say something about ℓ_2 played from $\{r_1\}$ and (b) uses Judge MR's belief about r_2 played from $\{m_1\}$ to say something about r_2 played from $\{\ell_1\}$. This invariance to the node from which an action is played is a central feature of the independence embodied in mass-function representation.

5.2. Additive Separability

In some ways, the independence and commonality of π resemble the independence and commonality of $\rho \cup \sigma$. With regard to independence, note that the additive form $\Sigma_{a \in t} \pi(a)$ of mass-function representation (4) resembles the multiplicative form $\Pi_{a \in t}(\rho \cup \sigma)(a)$ of the conditional-probability law (2). With regard to commonality, note that both π and $\rho \cup \sigma$ are commonly understood by all the strategic agents.

However, π and $\rho \cup \sigma$ differ in that the independence within $\rho \cup \sigma$ can be firmly grounded in the concept of stochastic independence from standard probability theory, while the independence within π grows out of the plausibility relation $\succcurlyeq$. More specifically, the independence within π represents an underlying independence within the plausibility relation $\succcurlyeq$, and accordingly, this underlying independence should ultimately be expressed in terms of a concept for binary relations.

This foundation is provided by this subsection. In particular, Corollary 2 will use the concept of additive separability from standard consumer theory[6] to lay a theoretical foundation for the independence that is implied by consistency. This material is tangential to the remainder of the paper, and its notation is ungainly. Accordingly, its proofs are relegated to Appendix B, and its notation does not appear elsewhere in the paper.

To begin, we recall that in consumer theory, preferences are defined over vectors that each list a quantity for each consumption good. Analogously, this subsection will regard each node as a vector that lists an action for each agent. Accordingly, the next three paragraphs show how to embed the set T of nodes within a Cartesian product whose coordinates are indexed by the agents.

This embedding of T is defined in three steps. First, create a "null" action o, and then, for each agent h, let $\check{F}(h) = F(h) \cup \{o\}$. Thus $\check{F}(h)$ expands agent h's feasible set $F(h)$ to admit the possibility of inaction. Next, consider the Cartesian product $\Pi_h \check{F}(h)$ and let $\check{t} = \langle \check{t}_h \rangle_h$ denote an arbitrary vector in this product. Finally, let V be the function that maps each node $t \in T$ to the vector $V(t) \in \Pi_h \check{F}(h)$ that is defined at each h by

$$[V(t)]_h = \begin{pmatrix} o & \text{if } |t \cap F(h)| = 0 \\ \text{the element of } t \cap F(h) & \text{if } |t \cap F(h)| = 1 \end{pmatrix} .$$

The following lemma proves that V is well-defined by showing that $|t \cap F(h)|$ is 0 or 1 for any t and h. It also proves a few other basic facts about V. (The symbols "V" and "ˇ" are meant to suggest "Vector".)

Lemma 5.1. *V is a well-defined and invertible function from T onto $V(T) \subseteq \Pi_h \check{F}(h)$. Further, $V^{-1}(\check{t}) = \{\check{t}_h | \check{t}_h \in A\}$ for every $\check{t} \in V(T)$. (Proof B.3 in Appendix B.)*

For example, consider Figure 1. The three agents are

$$h^1 = \{\{\}\} \text{ with } \check{F}(h^1) = \{o, d_1, w_1\} ,$$

[6] For the additive separability of a preference relation, see Debreu (1960), Gorman (1968), and Blackorby, Primont, and Russell (1978, Section 4.4).

$$h^2 = \{\{d_1\},\{w_1\}\} \text{ with } \check{F}(h^2) = \{o, d_2, w_2\} \text{ , and}$$
$$h^M = \{\{d_1,d_2\},\{d_1,w_2\},\{w_1,d_2\}\} \text{ with } \check{F}(h^M) = \{o, p_1, p_2\} \ .$$

Thus the product $\Pi_h \check{F}(h)$ has 3^3=27 vectors. Meanwhile, the set $V(T) \subseteq \Pi_h \check{F}(h)$ has those 13 vectors which correspond to the 13 nodes in T, and two of these vectors are $V(\{\}) = (o,o,o)$ and $V(\{w_1,w_2\}) = (w_1,w_2,o)$. This figure's example is relatively simple because the agents' order of play is exogenously determined. When the order of play is endogenously determined, as it is in Streufert (2012a, Figure 6), it becomes even more apparent that an action's position in the vector is determined by its agent and not by the order of play.

Now consider a plausibility relation $\succcurlyeq$. Let the *embedding* of $\succcurlyeq$ in $\Pi_h \check{F}(h)$ be the binary relation

$$\check{\succcurlyeq} = \{ \ (V(t^1), V(t^2)) \mid (t^1,t^2) \in \succcurlyeq \ \} \ .$$

Since V is invertible by Lemma 5.1, there is a one-to-one correspondence between the pairs of $\check{\succcurlyeq}$ and the pairs of $\succcurlyeq$. For example, for any plausibility relation $\succcurlyeq$ over the set tree of Figure 1,

$$(o,o,o) \ \check{\succcurlyeq} \ (w_1,w_2,o) \ \Leftrightarrow \ \{\} \succcurlyeq \{w_1,w_2\} \ .$$

Finally, consider a vector $\langle \varphi_h \rangle_h$ which lists a function $\varphi_h{:}\check{F}(h) \rightarrow \mathbb{R}$ for each agent h. Such a $\langle \varphi_h \rangle_h$ is said to *represent a completion* of $\check{\succcurlyeq}$ if for all t^1 and t^2

$$\text{(8)} \qquad \begin{aligned} \check{t}^1 \ \check{\succ} \ \check{t}^2 \ &\Rightarrow \ \Sigma_h \varphi_h(\check{t}^1_h) > \Sigma_h \varphi_h(\check{t}^2_h) \text{ and} \\ \check{t}^1 \ \check{\approx} \ \check{t}^2 \ &\Rightarrow \ \Sigma_h \varphi_h(\check{t}^1_h) = \Sigma_h \varphi_h(\check{t}^2_h) \ . \end{aligned}$$

Lemma 5.2. *Let $\succcurlyeq$ be the plausibility relation of an assessment and let $\check{\succcurlyeq}$ be its embedding in $\Pi_h \check{F}(h)$. Then for any mass function π, π represents a completion of $\succcurlyeq$ iff $\langle \varphi_h \rangle_h$ represents a completion of $\check{\succcurlyeq}$, where $(\forall h)$ $\varphi_h = \pi|_{F(h)} \cup \{(o,0)\}$. (Proof B.4 in Appendix B.)*

Corollary 2. *Let $\succcurlyeq$ be the plausibility relation of an assessment and let $\check{\succcurlyeq}$ be its embedding in $\Pi_h \check{F}(h)$. If the assessment is consistent, then $\check{\succcurlyeq}$ has a completion represented by a $\langle \varphi_h \rangle_h$. (Proof: Theorem 1 and Lemma 5.2.)*

In standard consumer theory, an ordering $\succsim$ over a product space $\Pi_{i=1}^n X_i$ of consumption vectors is said to be additively separable if there exists a vector $\langle u_i{:}X_i \rightarrow \mathbb{R} \rangle_{i=1}^n$ of functions such that $\succsim$ is represented by $\Sigma_{i=1}^n u_i$. Additive separability is understood to define independence across consumption goods. Analogously, Corollary 2 shows that consistency implies independence across agents. Corollary 2 is equivalent to Theorem 1 by Lemma 5.2.

5.3. Other independence concepts

Although it is novel to express the independence of zero-probability actions either in terms of mass-function representation or in terms of additive separability, other papers have found that consistency implies other concepts of independence. Each alternative has its own advantages and disadvantages.

Earlier alternatives might be placed into three groups. First, Blume, Brandenburger, and Dekel (1991a, Section 7), Hammond (1994, Section 6.5), and Halpern (2010, Section 6) all formulate independence in terms of non-Archimedean probability numbers. Second, Battigalli (1996, Section 2) formulates independence in terms of conditional probability systems. Third, Kohlberg and Reny (1997, Section 2) formulate independence in terms of relative probability systems. All of these other results use the rich probability structures that are discussed at the end of Subsection 3.1, and accordingly, the two concepts here have the advantage of being relatively straightforward. For example, the concept of a mass function and fifteen years of hindsight make our discussion of Figure 3 more straightforward than a comparable discussion of Kohlberg and Reny (1997, Figure 7).

Finally, it should be mentioned that all these results on the independence of zero-probability actions are anticipated by the definition of consistency itself. There consistency is the limit of a sequence of full-support Bayesian assessments, each of which embodies the standard sort of stochastic independence among agents. All the above results have succeeded in expressing this independence among agents without reference to a converging sequence of assessments.

6. Plausibility-like Numbers Elsewhere

Numbers like the plausibility numbers of a plausibility density function have appeared in a variety of places and served a variety of roles in the literature. We leave aside the resemblance between plausibility numbers and the mistake probabilities of trembling-hand perfection. Rather, this section considers the instances where plausibility-like numbers have been used in connection with consistency.

6.1. Plausibility-like numbers are necessary for consistency

This and the next subsection divide results about plausibility-like numbers into two broad groups. This subsection discusses results that show plausibility-like numbers are necessary for consistency. Theorem 1 belongs here. Conversely, the next subsection will discuss results that use plausibility-like numbers as part of sufficient conditions for consistency. That direction is much more routine.

Only two other papers belong in this subsection: Perea y Monsuwé, Jansen, and Peters (1997) and Kreps and Wilson (1982). We refer to them as `PJP` and `KW`. We believe that `PJP` Theorem 3.1 and `KW` Lemma A1 have been largely overlooked, and that they deserve much more recognition.

First consider `PJP`. There, a strategy profile σ is used to partition the set of strategic actions (denoted here by A^s) into the set $A^+(\sigma) = \{a{\in}A^s|\sigma(s){>}0\}$ of positive-probability actions and the set $A^0(\sigma) = \{a{\in}A^s|\sigma(a){=}0\}$ of zero-probability actions. Then an assessment is *justified* by $\mu{:}A^0(\sigma){\rightarrow}(0,1)$ if

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

where a product over the empty set is defined to be one. This concept is condition (1) of `PJP` Theorem 3.1, and since the concept is not given a name there, we have taken the liberty of using the word "justify". The values of μ can be called "error likelihoods". As Perea (2001, page 75) points out, the values of μ are not probability numbers, but rather numbers that measure the relative likelihoods of different errors (i.e., zero-probability actions).

Lemma 6.1. *π represents a completion of the plausibility relation of (σ,β) iff (a) $\pi^{-1}(0){=}A^c{\cup}A^+(\sigma)$ and (b) $e^\pi|_{A^0(\sigma)}$ justifies (σ,β). (Proof A.8 in Appendix A.)*

Corollary 3. *If an assessment is consistent, it can be justified. (Proof: Theorem 1 and Lemma 6.1.)*

Corollary 3 is essentially identical to part of `PJP` Theorem 3.1 (Corollary 3 is more general to the extent that (a) our framework explicitly accommodates arbitrary chance players, (b) we assume agent recall rather than perfect recall, and (c) we derive a function with integer rather than real values.) While `PJP` proved this result by the separating hyperplane theorem, the combination of Theorem 1 and Lemma 6.1 proves it by algebra alone. Also, unlike `PJP`, we will connect this result to `KW` Lemma A1, in the fifth paragraph below this one.

Now consider `KW`.[7] As on `KW` page 880, let a *basis* b be an arbitrary subset of $A^s \cup T^s$. Then say that such a basis b *supports* an assessment (σ, β) if b is the the union of σ's support (which is a subset of A^s) and β's support (which is a subset of T^s). Finally, as on `KW` page 887, say that a function $K{:}A^s \to \mathbb{Z}_+$ *labels* a basis b if

(9a) $$(\forall h{\in}H^s)(\exists a{\in}F(h)) \quad K(a)=0 \ ,$$

(9b) $$(\forall a{\in}A^s) \quad a \in b \ \text{ iff } \ K(a)=0 \ , \text{ and}$$

(9c) $$(\forall h{\in}H^s)(\forall t{\in}h) \quad t \in b \ \text{ iff } \ t \in \operatorname{argmin}_{t' \in h} \Sigma_{a \in t'} K(a) \ .$$

Lemma 6.2. *π is an integer-valued representation of a completion of the plausibility relation of (σ, β) iff (a) $\pi|_{A^c} = 0$ and (b) $-\pi|_{A^s}$ labels the support of (σ, β). (Proof A.9 in Appendix A.)*

Corollary 4. *If an assessment is consistent, its support can be labelled. (Proof: Theorem 1 and Lemma 6.2.)*

Corollary 4 is essentially identical to half of `KW` Lemma A1 (Corollary 4 is more general to the extent that we (a) accommodate arbitrary chance players and (b) assume agent recall rather than perfect recall). Streufert (2006, Subsection 3.2) shows that the `KW` proof of this result has a significant gap.

Theorem 1, Corollary 3, and Corollary 4 are all equivalent by means of Lemmas 6.1 and 6.2. Thus we have clarified, unified, and repaired an important literature that we feel deserves much more recognition.

6.2. Plausibility-like numbers are part of sufficient conditions for consistency

We begin with an example. Figure 2's assessment[8] is consistent because it is the limit of the sequence of full-support Bayesian assessments depicted in Figure 4. The exponents on the index n in Figure 4 are

---

[7] 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).

[8] Figure 2 suppressed the terminal nodes because they are irrelevant to mass-function representation. In this context, we need to suppose that Figure 2's assessment has the manager playing p_2.

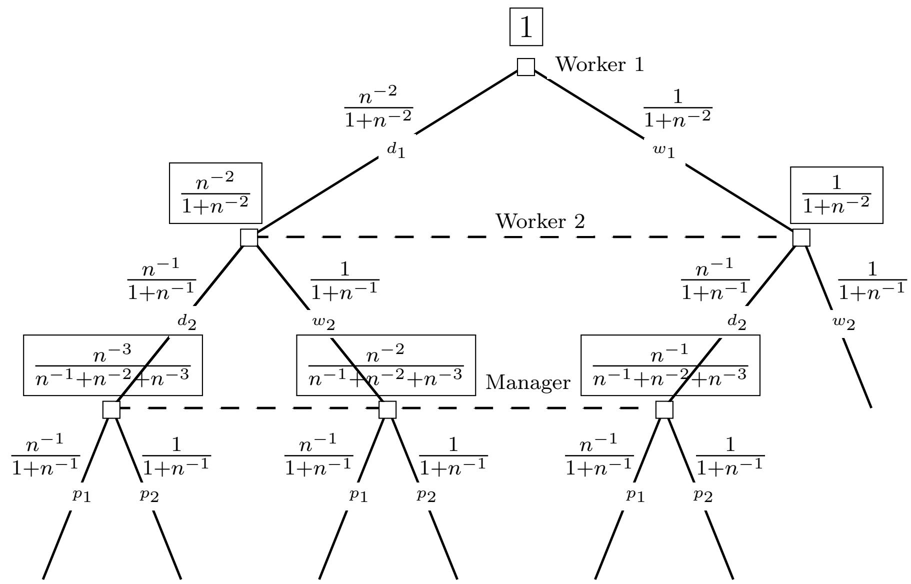

FIGURE 4. A sequence of full-support Bayesian assessments. The exponents are plausibility-like numbers.

plausibility-like numbers. This is particularly apparent if one compares Figure 4's numerators with Figure 2's plausibility numbers.

Similarly, many papers use plausibility-like numbers to derive consistency within interesting economic models. Some arbitrarily chosen examples are the "error likelihoods" in Anderlini, Gerardi, and Lagunoff (2008, page 359), the "orders of probability" in Kobayashi (2007, page 525), and counting the "steps off the equilibrium path" in Fudenberg and Levine (2006, Definition 3.2). Some authors, such as Anderlini, Gerardi, and Lagunoff (2008), replace n approaching infinity with $\varepsilon = 1/n$ approaching zero. In such cases, the exponent on ε is the negative of a plausibility-like number.

Such a customary application of the definition of consistency can be used to derive the converse of Corollary 4. The following equivalence results.

Corollary 5. *A basis supports at least one consistent assessment iff it can be labelled. (Proof: Corollary 4 and Lemma A.10.)*

Corollary 5 is virtually identical to `KW` Lemma A1 (whose proof we have repaired). At one time, all three `KW` theorems relied on this lemma. Since then, `KW` Theorems 2 and 3 have been superseded by Govindan and Wilson (2001, Theorem 2.2) and Blume and Zame (1994, Theorem 4). Both of those papers use abstract theorems about semi-algebraic sets.

However, there is some merit in setting the record straight for historical reasons. Further, we use Corollary 5 to prove Corollaries 6 and 7 below. And finally, KW Theorem 1 continues to provide an explicit partition of the set of sequential equilibria, and KW Lemma 2 continues to provide an explicit partition of the set of consistent assessments. Sometimes, these partitions are regarded as "long complicated construction[s]", as remarked by Govindan and Wilson (2001, page 765). However, both these partitions are indexed by the set of bases b that can be labelled, and this index set becomes more intuitive when Lemma 6.2 allows us to regard labellings as the negatives of mass functions. Accordingly, arguments using these relatively explicit KW partitions may yet complement arguments using relatively abstract theorems about semi-algebraic sets.

Finally, we note for completeness that plausibility-like numbers only concern the supports of an assessment's strategies and beliefs. Accordingly, they are only part (albeit the difficult part) of a full characterization of consistency. PJP accomplishes the remainder in part (2) of its Theorem 3.1, and KW does the same in its Lemma A2. Streufert (2012b)'s synthesis and reformulation endeavours to make the full characterizations of PJP and KW more easily accessible to a general audience.

Without considering matters other than the support of an assessment, one can only hope to characterize consistency in the relatively limited circumstances of the following section. Nonetheless, the results below appear to be both new and useful.

7. Special Case: Degenerate-Support Assessments

An assessment is said to have *degenerate support* if the support of each strategy and each belief is degenerate. In other words, a degenerate-support assessment is an assessment that specifies pure strategies and sure beliefs. In brief, it is all ones and zeros.

Corollary 6. *A degenerate-support assessment is consistent iff its plausibility relation has a completion represented by a mass function.*

Proof. Consider a degenerate-support assessment. Regardless of the degenerate-support assumption, the consistency of the assessment implies mass-function representation by Theorem 1. Conversely, suppose that the assessment's plausibility relation has a completion represented by a mass function. Then by Lemma 6.2, the assessment's support has a labelling. Thus by Corollary 5, the assessment's support supports at least one consistent assessment. By the degenerate-support assumption, the original assessment must equal the consistent assessment. □

For example, the assessments[9] of Figures 2 and 3 satisfy Corollary 6's degenerate-support assumption. In accord with the corollary, the assessment of Figure 2 is consistent, and the assessment of Figure 3 is inconsistent.

Finally, we recall that an assessment (σ, β) is a *sequential equilibrium* if (a) it is *sequentially rational* in the sense that

$$(\forall i)(\forall h{\in}i)(\forall a{\in}F(h))\ \sigma(a)>0 \text{ implies } a\in \operatorname{argmax}_{\hat{a}\in F(h)} \Sigma_{x\in h}\beta(x)\,\Sigma_{z\supseteq x\cup\{\hat{a}\}}\big[\Pi_{a'\in z\sim(x\cup\{\hat{a}\})}(\rho\cup\sigma)(a')\big]u_i(z)\,.$$

and (b) it is consistent. These definitions and Corollary 6 imply the following.

Corollary 7. *A degenerate-support assessment is a sequential equilibrium iff (a) it is sequentially rational and (b) its plausibility relation has a completion represented by a mass function.*

8. CONCLUSION

Theorem 1 shows that if an assessment is consistent, then its plausibility relation has a completion represented by a mass function. Its proof re-uses the early foundations of probability theory and requires nothing more than linear algebra. The key to using this straightforward framework is Streufert (2012a)'s result that a node can be identified with the set of actions leading to it.

The theorem has several conceptual and computational implications. First, the theorem and Corollary 2 provide two new ways to formalize that consistency specifies that zero-probability agents reason that

[9]Figures 2 and 3 suppress the terminal nodes because they are irrelevant to mass-function representation by Lemma B.1. Accordingly, this paragraph holds for any pure strategies chosen by the manager of Figure 2 and the three judges of Figure 3.

prior zero-probability actions were played independently of other past actions. Second, the theorem and Corollary 1 provide two necessary conditions for consistency, the theorem's being more discriminating and the corollary's easier to test. Third, Corollaries 3 and 4 clarify and unify two under-appreciated characterizations of consistency in the literature. Fourth, Corollary 5 repairs a useful characterization of the supports of consistent assessments. Fifth and finally, Corollaries 6 and 7 discover straightforward characterizations of consistency and sequential equilibrium in the special case of a degenerate-support assessment.

Appendix A. Main Arguments

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 $h \notin i^c$ replacing $h \in i^c$. □

Proof A.2 (for Lemma 3.1)**.** Note $\overset{c}{\approx}$ is symmetric and equal to

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

Further, $\overset{\sigma}{\approx}$ is symmetric, $\overset{\sigma}{\succ}$ is asymmetric, and the two are disjoint subsets of

$$(11)\qquad \begin{aligned} &\{\ (t, t \cup \{a\}) \mid a \in F(t) \text{ and } a \in A^s\ \} \\ \cup\ &\{\ (t \cup \{a\}, t) \mid a \in F(t) \text{ and } a \in A^s\ \}\ . \end{aligned}$$

Similarly, $\overset{\beta}{\approx}$ is symmetric, $\overset{\beta}{\succ}$ is asymmetric, and the two are disjoint subsets of

$$(12)\qquad \{\ (t_1, t_2) \mid (\exists h \in H^s)\ \{t_1, t_2\} \in h\ \}\ .$$

This paragraph observes that these three sets are pairwise disjoint. (10) and (11) are disjoint because A is partitioned by $\{A^c, A^s\}$. Further, the union of (10) and (11) is disjoint from (12). 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 (10), (11), and (12) are pairwise disjoint, the disjointedness observed in the first paragraph implies that $\succcurlyeq$ is partitioned by $\{\overset{c}{\approx}, \overset{\sigma}{\approx}, \overset{\sigma}{\succ}, \overset{\beta}{\approx}, \overset{\beta}{\succ}\}$. Thus the symmetries and asymmetries observed in the first paragraph imply that $\approx$ is partitioned by $\{\overset{c}{\approx}, \overset{\sigma}{\approx}, \overset{\beta}{\approx}\}$ and that $\succ$ is partitioned by $\{\overset{\sigma}{\succ}, \overset{\beta}{\succ}\}$. $\square$

A.2. For Theorem 1's proof

Fact A.3 (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 ($D\pi{\gg}0$ means every element of $D\pi$ is positive and δ^T means the transpose of δ).*

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

(b) *Not* $(\exists\delta{\in}\mathbb{Z}^d_+{\sim}\{0\})(\exists\varepsilon{\in}\mathbb{Z}^e)\ \delta^TD + \varepsilon^TE = 0$.

(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.4 (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 $\delta \in \mathbb{Z}^{|\succ|}_+{\sim}\{0\}$ and $\varepsilon \in \mathbb{Z}^{|\approx|}$ such that $\delta^TD + \varepsilon^TE = 0$. 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 $\varepsilon^TE = \varepsilon_+^TE$. Thus we have $\delta \in \mathbb{Z}^{|\succ|}_+{\sim}\{0\}$ and $\varepsilon_+ \in \mathbb{Z}^{|\approx|}_+$ such that $\delta^TD + \varepsilon_+^TE = 0$. 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 $\delta^TD + \varepsilon_+^TE = 0$ yields that this sequence satisfies $(\forall a)\ |\{m|a{\in}s^m\}| = |\{m|a{\in}t^m\}|$, and

$\delta \in \mathbb{Z}_+^{|\prec|} \sim \{0\}$ 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.3, there is a vector $\pi \in \mathbb{Z}^{|A|}$ such that $D\pi \gg 0$ and $E\pi = 0$. By the definitions of D and E, this is equivalent to π being a mass function representing a completion of $\succcurlyeq$.

Necessity of the Cancellation Law.[10] Suppose π 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,

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

Yet by representation,

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

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. $\square$

Lemma A.5. *Suppose that $\succcurlyeq$ is the plausibility relation of (σ, β), and that $\langle (\sigma_n, \beta_n) \rangle_n$ is a sequence of full-support Bayesian assessments that converges to (σ, β). Then*

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

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

Proof. This paragraph shows

$$(\forall\, t^1 \overset{\sigma}{\succ} t^2) \ \lim_n \frac{\Pi_{a' \in t^2} (\rho \cup \sigma_n)(a')}{\Pi_{a' \in t^1} (\rho \cup \sigma_n)(a')} = 0 \ . \tag{13}$$

Accordingly, suppose $t^1 \overset{\sigma}{\succ} t^2$. By the definition of $\overset{\sigma}{\succ}$, there exists a such that $\sigma(a)$=0, $a$$\in$$F(t^1)$, and $t^1 \cup \{a\} = t^2$. Thus, since $a$$\notin$$t^1$ by the definition of F,

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

[10]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.

This paragraph shows

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

Suppose $t^1 \overset{\sigma}{\approx} t^2$. By the definition of $\overset{\sigma}{\approx}$, there exists a with $\sigma(a){>}0$ such that either, $a{\in}F(t^1)$ and $t^1{\cup}\{a\}{=}t^2$, or, $a{\in}F(t^2)$ and $t^2{\cup}\{a\}{=}t^1$. In the first case, $a{\notin}t^1$ by the definition of F and thus

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

and in the second case, $a{\notin}t^2$ by the definition of F and thus

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

In a similar fashion, this paragraph shows

$$(15)\qquad (\forall\, t^1 \overset{c}{\approx} t^2)\ \lim_n \frac{\Pi_{a'\in t^2}(\rho\cup\sigma_n)(a')}{\Pi_{a'\in t^1}(\rho\cup\sigma_n)(a')} \in (0,\infty)\ .$$

Suppose $t^1 \overset{c}{\approx} t^2$. By the definition of $\overset{c}{\approx}$, there exists $a{\in}A^c$ such that either, $a{\in}F(t^1)$ and $t^1{\cup}\{a\}{=}t^2$, or, $a{\in}F(t^2)$ and $t^2{\cup}\{a\}{=}t^1$. In the first case, $a{\notin}t^1$ and thus

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

and in the second case, $a{\notin}t^2$ and thus

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

Finally, note that if t^1 and t^2 are distinct nodes in some $h{\in}H^s$, and if $\beta(t^1) > 0$, then

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

where the first equality holds because at least one of the nodes must be nonempty, the second follows from (2), and the third follows from consistency and the assumption that $\beta(t^1) > 0$. Thus by the definitions of $\overset{\beta}{\succ}$ and $\overset{\beta}{\approx}$ we have

$$(16) \qquad (\forall\, t^1 \overset{\beta}{\succ} t^2)\ \lim_n \frac{\Pi_{a\in t^2}(\rho\cup\sigma_n)(a)}{\Pi_{a\in t^1}(\rho\cup\sigma_n)(a)} = \frac{\beta(t^2)}{\beta(t^1)} = 0 \ \text{ and}$$

$$(17) \qquad (\forall\, t^1 \overset{\beta}{\approx} t^2)\ \lim_n \frac{\Pi_{a\in t^2}(\rho\cup\sigma_n)(a)}{\Pi_{a\in t^1}(\rho\cup\sigma_n)(a)} = \frac{\beta(t^2)}{\beta(t^1)} \in (0,\infty)\ .$$

The lemma's conclusion follows from (13)–(17) and the definitions of $\succ$ and $\approx$. □

A.3. "Counting steps below path"

Proof A.6 (for Lemma 4.3)**.** (a) Take any $a{\in}A^c$. By Lemma A.1(a), there is some $t{\in}T^c$ such that $a{\in}F(t)$. Then $t \approx t{\cup}\{a\}$ by the definition of $\overset{c}{\approx}$, which implies $\Sigma_{a'\in t}\pi(a') = \Sigma_{a'\in t\cup\{a\}}\pi(a')$ by representation (4), which implies $\pi(a) = 0$ since $a{\notin}t$ by $a{\in}F(t)$ and the definition of F.

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

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

A.4. Links to PJP and KW

Lemma A.7. *$\pi{:}A{\to}\mathbb{R}$ represents a completion of (σ,β)'s plausibility relation iff π satisfies*

$$(18a) \qquad \pi \textit{ is nonpositive-valued}\ ,$$

$$(18b) \qquad (\forall a{\in}A^c)\ \pi(a) = 0\ ,$$

$$(18c) \qquad (\forall a{\in}A^s)\ \sigma(a) > 0 \textit{ iff } \pi(a) = 0\ , \textit{ and}$$

$$(18d) \qquad (\forall h{\in}H^s)(\forall t{\in}h)\ \beta(t) > 0 \textit{ iff } t \in \textit{argmax}_{t'\in h}\, \Sigma_{a\in t'}\pi(a)\ .$$

Proof. Take any π and (σ,β). We are to show that representation (4) is equivalent to (18).

Necessity of (18). (18a) follows from Lemma 4.3's last sentence. (18b) follows from Lemma 4.3(a). The forward direction of (18c) follows from Lemma 4.3(b). The contrapositive of the reverse direction of (18c) holds because (a) not $\sigma(a) > 0$ implies $\sigma(a) = 0$ by the non-negativity of σ, which (b) implies $\pi(a) < 0$ by Lemma 4.3(c), which (c) implies not $\pi(a) = 0$.

(18d) is established by this paragraph. Take any $h{\in}H^s$ and any $t{\in}h$. On the one hand, suppose $\beta(t){=}0$. Since $\beta|_h$ is a probability distribution, there is some $t^*{\in}h$ such that $\beta(t^*){>}0$. This implies $t^* \succ t$ by the definition of $\overset{\beta}{\succ}$, which implies $\Sigma_{a\in t^*}\pi(a) > \Sigma_{a\in t}\pi(a)$ by representation (4), which implies $t \notin \operatorname{argmax}_{t'\in h} \Sigma_{a\in t'}\pi(a)$. On the other hand, suppose $\beta(t){>}0$. Then consider any other $t'{\in}h$. By the definitions of $\overset{\beta}{\succ}$ and $\overset{\beta}{\approx}$, either $t \overset{\beta}{\succ} t'$ or $t \overset{\beta}{\approx} t'$, and thus in either event we have $t \succsim t'$. Hence by (4), $\Sigma_{a\in t}\pi(a) \geq \Sigma_{a\in t'}\pi(a)$. Since this has been demonstrated for any $t'{\in}h$, we have that $t \in \operatorname{argmax}_{t'\in h} \Sigma_{a\in t'}\pi(a)$.

Sufficiency of (18). (a) Take any t^1 and t^2 such that $t^1 \succ t^2$. By the definition of $\succ$, we have $t^1 \overset{\sigma}{\succ} t^2$ or $t^1 \overset{\beta}{\succ} t^2$.

In the first case, $t^1{\in}T^s$ and there is an $a{\in}F(t^1)$ such that $t^2{=}t^1{\cup}\{a\}$ and $\sigma(a){=}0$. Note that $\{a\} = t^2{\sim}t^1$ because $t^2{=}t^1{\cup}\{a\}$ and because $a{\notin}t^1$ by $a{\in}F(t^1)$ and the definition of F. Therefore $\Sigma_{a'\in t^1}\pi(a') > \Sigma_{a'\in t^2}\pi(a')$ because $\pi(a){<}0$ by $\sigma(a){=}0$ and (18c).

In the second case, there is an $h{\in}H^s$ such that $\{t^1,t^2\}{\subseteq}h$, $\beta(t^1){>}0$, and $\beta(t^2){=}0$. Thus $\Sigma_{a\in t^1}\pi(a) = \max_{t'\in h}\Sigma_{a\in t'}\pi(a) > \Sigma_{a\in t^2}\pi(a)$ by two applications of (18d).

(b) Take any t^1 and t^2 such that $t^1 \approx t^2$. By the definition of $\approx$, we have $t^1 \overset{\sigma}{\approx} t^2$, or $t^1 \overset{\beta}{\approx} t^2$, or $t^1 \overset{c}{\approx} t^2$.

In the first case, either [a] $t^1{\in}T^s$ and there is an $a{\in}F(t^1)$ such that $t^2{=}t^1{\cup}\{a\}$ and $\sigma(a){>}0$, or symmetrically [b] $t^2{\in}T^s$ and there is an $a{\in}F(t^2)$ such that $t^1{=}t^2{\cup}\{a\}$ and $\sigma(a){>}0$. In subcase [a], $\Sigma_{a\in t^2}\pi(a) = \Sigma_{a\in t^1\cup\{a\}}\pi(a) = \Sigma_{a\in t^1}\pi(a)$, where the first equality holds by $t^2{=}t^1{\cup}\{a\}$ and the second holds because $\pi(a){=}0$ by $\sigma(a){>}0$ and (18c). Subcase [b] can be treated by a symmetric argument.

In the second case, there is an $h{\in}H^s$ such that $\{t^1,t^2\}{\subseteq}h$, $\beta(t^1){>}0$, and $\beta(t^2){>}0$. Thus $\Sigma_{a\in t^1}\pi(a) = \max_{t'\in h}\Sigma_{a\in t'}\pi(a) = \Sigma_{a\in t^2}\pi(a)$ by two applications of (18d).

In the third case, either [a] $t^1{\in}T^c$ and there is an $a{\in}F(t^1)$ such that $t^2{=}t^1{\cup}\{a\}$ or symmetrically [b] $t^2{\in}T^c$ and there is an $a{\in}F(t^2)$ such that $t^1{=}t^2{\cup}\{a\}$. To treat subcase [a], note $a{\in}A^c$ by $a{\in}F(t^1)$, $t^1{\in}T^c$, and Lemma A.1(a). Thus $\pi(a) = 0$ by (18b). Hence $\Sigma_{a\in t^2}\pi(a) =$

$\Sigma_{a\in t^1\cup\{a\}}\pi(a) = \Sigma_{a\in t^1}\pi(a)$ by $t^2{=}t^1\cup\{a\}$ and the previous sentence. Subcase [b] can be treated with a symmetric argument. □

Proof A.8 (for Lemma 6.1)**.** Take any (σ, β) and any π. Because of Lemma A.7, Lemma 6.1 can be established by showing that (18) is equivalent to

$$\pi^{-1}(0) = A^c \cup A^+(\sigma) \ , \tag{19a}$$

$$(\forall a{\in}A^0(\sigma))\ e^{\pi(a)} \in (0,1) \ , \text{ and} \tag{19b}$$

$$(\forall h{\in}H^s)(\forall t{\in}h)\ \beta(t) > 0 \text{ iff } t \in \text{argmax}_{t'\in h}\, \Pi_{a\in t'\cap A^0(\sigma)} e^{\pi(a)} \ . \tag{19c}$$

As an initial step, this paragraph shows that (19a) implies

$$\begin{aligned}(\forall h{\in}H^s)\ \ &\text{argmax}_{t'\in h}\, \Pi_{a\in t'\cap A^0(\sigma)} e^{\pi(a)} \\ =\ \ &\text{argmax}_{t'\in h}\, \Sigma_{a\in t'}\pi(a) \ .\end{aligned} \tag{20}$$

To see this, note that the second of the following equalities is implied by (19a) and by the fact that $\{A^c, A^0(\sigma), A^+(\sigma)\}$ is a collection of disjoint sets whose union is A:

$$\begin{aligned}(\forall t')\ \ &\ln\big(\Pi_{a\in t'\cap A^0(\sigma)} e^{\pi(a)}\big) \\ =\ \ &\Sigma_{a\in t'\cap A^0(\sigma)}\pi(a) \\ =\ \ &\Sigma_{a\in t'}\pi(a) \ .\end{aligned}$$

This equality then implies that $\Pi_{a\in t'\cap A^0(\sigma)} e^{\pi(a)}$ and $\Sigma_{a\in t'}\pi(a)$ are ordinally equivalent over any $h{\in}H^s$.

(18)⇒(19). Since (1) the domain of π is A and A is the union of the disjoint sets of $\{A^c, A^s\}$, and (2) $A^+(\sigma) \subseteq A^s$ by the definition of $A^+(\sigma)$, we have that (19a) is equivalent to

$$\begin{aligned}&(\forall a{\in}A^c)\ \pi(a) = 0 \text{ iff } a{\in}A^c \text{ and} \\ &(\forall a{\in}A^s)\ \pi(a) = 0 \text{ iff } a{\in}A^+(\sigma) \ .\end{aligned}$$

Thus (19a) holds: the first equivalence holds (vacuously) by (18b), and the second is identical to (18c) by the definition of $A^+(\sigma)$. Since A^s is partitioned by $\{A^+(\sigma), A^0(\sigma)\}$ because of the nonnegativity of σ, the second of the above equivalences implies that $(\forall a{\in}A^0(\sigma))\ \pi(a) \neq 0$. Thus (19b) follows from (18a). Finally, (20) follows from the previous paragraph and (19a), which has already been derived. Thus by (20), (19c) follows from (18d).

(19)⇒(18). (18a) follows from (19a,b) because the domain of π is A, which is the union of $\{A^c, A^+(\sigma), A^0(\sigma)\}$. (18b) follows from (19a).

The forward direction of (18c) holds because (1) $\sigma(a)>0$ is equivalent to $a{\in}A^+(\sigma)$ by the definition of $A^+(\sigma)$, which (2) implies $\pi(a)=0$ by (19a). The contrapositive of the reverse direction of (18c) holds because (1) not $\sigma(a)>0$ implies $\sigma(a)=0$ by the nonnegativity of $\sigma(a)$, which (2) is equivalent to $a{\in}A^0(\sigma)$, which (3) implies $\pi(a)<0$ by (19b). Finally, note that (20) holds by (19a) and the paragraph before the last one. Thus by (20), (18d) follows from (19c). □

Proof A.9 (for Lemma 6.2)**.** This paragraph shows that for any K and (σ,β), K labels the support of (σ,β) iff

(21a) $$(\forall a{\in}A^s)\ \ \sigma(a)>0 \ \text{ iff } \ K(a)=0\ ,\ \text{and}$$

(21b) $$(\forall h{\in}H^s)(\forall t{\in}h)\ \ \beta(t)>0 \ \text{ iff } \ t\in\operatorname{argmin}_{t'\in h}\Sigma_{a\in t'}K(a)\ .$$

To see this, take any K and (σ,β), and then let b be the support of (σ,β). By the definition (9) of labelling, K labels b iff

$$(\forall h{\in}H^s)(\exists a{\in}F(h))\ \ K(a)=0\ ,$$
$$(\forall a{\in}A^s)\ \ a\in b \ \text{ iff } \ K(a)=0\ ,\ \text{and}$$
$$(\forall h{\in}H^s)(\forall t{\in}h)\ \ t\in b \ \text{ iff } \ t\in\operatorname{argmin}_{t'\in h}\Sigma_{a\in t'}K(a)\ .$$

Thus, by the definition of support, K labels the support b iff

(22a) $$(\forall h{\in}H^s)(\exists a{\in}F(h))\ \ K(a)=0\ ,$$

(22b) $$(\forall a{\in}A^s)\ \ \sigma(a)>0 \ \text{ iff } \ K(a)=0\ ,\ \text{and}$$

(22c) $$(\forall h{\in}H^s)(\forall t{\in}h)\ \ \beta(t)>0 \ \text{ iff } \ t\in\operatorname{argmin}_{t'\in h}\Sigma_{a\in t'}K(a)\ .$$

Note that b is absent from (22). Also note that (22b) implies (22a) because each $\sigma|_{F(h)}$ is a probability distribution by assumption.[11] Thus K labels the support of (σ,β) iff (21) holds.

This paragraph shows that for any π and (σ,β), $-\pi|_{A^s}$ labels the support of (σ,β) iff

(23a) $$\pi|_{A^s} \text{ assumes nonpositive integer values}\ ,$$

(23b) $$(\forall a{\in}A^s)\ \sigma(a)>0 \text{ iff } -\pi(a)=0\ ,\ \text{and}$$

(23c) $$(\forall h{\in}H^s)(\forall t{\in}h)\ \beta(t)>0 \text{ iff } t\in\operatorname{argmin}_{t'\in h}\Sigma_{a\in t'}{-\pi(a)}\ .$$

First assume (23). Since (23a) yields $-\pi|_{A^s}{:}A^s{\rightarrow}\mathbb{Z}_+$, we may substitute $-\pi|_{A^s}$ for K in the previous paragraph. Thus (23b,c) yields that

[11] Here (22a)=(9a) is superfluous because b is assumed to be the support of some assessment. However, it is not superfluous in the first paragraph of Lemma A.10's proof, for there it has not been assumed that b is the support of some assessment.

$-\pi|_{A^s}$ labels the support of (σ,β). Conversely, suppose that $-\pi|_{A^s}$ labels the support of (σ,β). By the definition of a labelling, we have $-\pi|_{A^s}{:}A^s{\rightarrow}\mathbb{Z}_+$, which implies (23a). Further, the previous paragraph implies (23b,c).

We are now in a position to prove the lemma. Take any π and (σ,β). By Lemma A.7 and the previous paragraph, the task is to show that (18) and the integer-valuedness of π is equivalent to (23) and $\pi|_{A^c}=0$. Assume (18) and integer-valuedness. Then (23a) follows from (18a) and integer-valuedness. Further, (23b), (23c) and $\pi|_{A^c}=0$ follows from (18c), (18d) and (18b). Conversely, assume (23) and $\pi|_{A^c}=0$. Then (18a) and integer-valuedness follow from (23a) and $\pi|_{A^c}=0$. Further, (18b), (18c) and (18d) follow from $\pi|_{A^c}=0$, (23b) and (23c). □

A.5. THE SUFFICIENCY OF LABELLING

Lemma A.10. *If a basis can be labelled, then it supports a consistent assessment.*

Proof. Suppose that b is labelled by K. To avoid a lengthy argument, we will employ Streufert (2012b, Theorem 3). Accordingly, define $\pi = -K$ and define κ at each $h{\in}H^s$ and each $a{\in}h$ by

$$\kappa(a) = \left(\begin{array}{cc} \frac{1}{|b\cap F(h)|} & \text{if } a{\in}b \\ 1 & \text{if } a{\notin}b \end{array} \right) .$$

Note κ is well-defined because $|b{\cap}F(h)| \geq 1$ by (9a) in the definition of labelling. (Unfortunately the κ notation in Streufert (2012b) almost conflicts with the K notation in Kreps and Wilson (1982). They are very different objects.)

Now define (σ,β) by $(\forall a{\in}A^s)$

$$\sigma(a) = \left(\begin{array}{cc} \kappa(a) & \text{if } \pi(a)=0 \\ 0 & \text{if } \pi(a)<0 \end{array} \right) ,$$

and $(\forall h{\in}H^s)(\forall t{\in}h)$

$$\beta(t) = \left(\begin{array}{cc} \dfrac{\Pi_{a\in x\cap A^c}\,\rho(a)\times\Pi_{a\in x\cap A^s}\,\kappa(a)}{\Sigma_{t'\in T_h^*(\pi)}\,\Pi_{a\in t'\cap A^c}\,\rho(a)\times\Pi_{a\in t'\cap A^s}\,\kappa(a)} & \text{if } t\in T_h^*(\pi) \\ 0 & \text{if } t\notin T_h^*(\pi) \end{array} \right) .$$

where $T_h^*(\pi) = \text{argmax}_{t' \in h}\, \Sigma_{a \in t' \cap A^s} \pi(a)$. This (σ, β) is an assessment because (a) $(\forall h)\ \Sigma_{t \in h} \beta(t) = 1$ by inspection and (b) $(\forall h {\in} H^s)$

$$\begin{aligned}\Sigma_{a \in F(h)}\, \sigma(a) &= \Sigma_{a \,|\, a \in F(h) \text{ and } \pi(a)=0}\, \kappa(a) \\ &= \Sigma_{a \,|\, a \in F(h) \text{ and } K(a)=0}\, \kappa(a) \\ &= \Sigma_{a \,|\, a \in F(h) \text{ and } a \in b}\, \kappa(a) \\ &= \Sigma_{a \in b \cap F(h)} \tfrac{1}{|b \cap F(h)|} = 1\end{aligned}$$

by the definition of σ, the definition of π, (9b) in the definition of labelling, and the definition of κ. Further, (σ, β) is consistent by Streufert (2012b, Theorem 3) and the definition of (σ, β).

It remains to be shown that b supports (σ, β). For any $h {\in} H^s$ and any $a {\in} F(h)$, we have

$$\sigma(a) > 0 \;\Leftrightarrow\; \pi(a) = 0 \;\Leftrightarrow\; K(a) = 0 \;\Leftrightarrow\; a \in b$$

by the definition of σ, the definition of π, and (9b) in the definition of labelling. For any $h {\in} H^s$ and any $t {\in} h$, we have

$$\begin{aligned}\beta(t) > 0 \;&\Leftrightarrow\; t \in T_h^*(\pi) \\ &\Leftrightarrow\; t \in \text{argmax}_{t' \in h}\ \Sigma_{a \in t' \cap A^s} \pi(a) \\ &\Leftrightarrow\; t \in \text{argmax}_{t' \in h}\ \Sigma_{a \in t' \cap A^s} {-K(a)} \\ &\Leftrightarrow\; t \in \text{argmin}_{t' \in h}\ \Sigma_{a \in t' \cap A^s} K(a) \\ &\Leftrightarrow\; t \in b\end{aligned}$$

by the definition of β, the definition of $T_h^*(\pi)$, the definition of π, and (9c) in the definition of a labelling. □

APPENDIX B. TANGENTIAL ARGUMENTS

We have separated the following subsections from Appendix A because each is a logical tangent with its own relatively extensive notation.

B.1. THE IRRELEVANCE OF TERMINAL NODES

Figures 2 and 3 suppress their terminal nodes. This is justified by the following lemma which shows that terminal nodes are irrelevant to the existence of a mass-function representation.

The lemma requires that we generalize the definition of mass-function representation. Accordingly, consider any set $\hat{A}$ and any $\hat{\succcurlyeq}$ comparing subsets t of $\hat{A}$. By definition, $\hat{\pi}{:}\hat{A}{\to}\mathbb{R}$ *represents a completion*[12] of $\hat{\succcurlyeq}$ if for all subsets t^1 and t^2 of $\hat{A}$,

$$\begin{aligned} t^1 \mathrel{\hat{\succ}} t^2 \;&\Rightarrow\; \Sigma_{a\in t^1}\hat{\pi}(a) > \Sigma_{a\in t^2}\hat{\pi}(a) \text{ and} \\ t^1 \mathrel{\hat{\approx}} t^2 \;&\Rightarrow\; \Sigma_{a\in t^1}\hat{\pi}(a) = \Sigma_{a\in t^2}\hat{\pi}(a)\ , \end{aligned} \tag{24}$$

where $\hat{\succ}$ and $\hat{\approx}$ are the asymmetric and symmetric parts of $\hat{\succcurlyeq}$. This definition of a mass-function representation reduces to that of the text when $\hat{A}$ is A and $\hat{\succcurlyeq}$ is the plausibility relation $\succcurlyeq$ of an assessment.

The definition is also used in the following lemma when $\hat{A}$ is $\bigcup(T{\sim}Z)$ and $\hat{\succcurlyeq}$ is $\succcurlyeq|_{T\sim Z}$, where $\succcurlyeq|_{T\sim Z}$ is defined to be $\succcurlyeq \cup\, (T{\sim}Z)^2$. This $\succcurlyeq|_{T\sim Z}$ is the restriction of $\succcurlyeq$ to the set $T{\sim}Z$ of nonterminal nodes, and $\bigcup(T{\sim}Z)$ is the set of actions that lead to them.

Lemma B.1. *Suppose $\succcurlyeq$ is the plausibility relation of some assessment. Then $\succcurlyeq$ has a completion represented by some $\pi{:}A{\to}\mathbb{R}$ iff $\succcurlyeq|_{T\sim Z}$ has a completion represented by some $\pi^o{:}\bigcup(T{\sim}Z){\to}\mathbb{R}$.*

Proof. Suppose $\succcurlyeq$ is the plausibility relation of (σ,β).

First suppose $\pi{:}A{\to}\mathbb{R}$ represents a completion of $\succcurlyeq$. Equivalently, π and $\succcurlyeq$ satisfy (24) for all (t^1,t^2) in $\succcurlyeq$. Thus, since $\succcurlyeq|_{T\sim Z}$ equals $\succcurlyeq \cap\, (T{\sim}Z)^2$ by assumption, π and $\succcurlyeq|_{T\sim Z}$ satisfy (24) for all (t^1,t^2) in $\succcurlyeq|_{T\sim Z}$. Further, since that fact only depends upon the values of π over $\bigcup(T{\sim}Z)$, we have that $\pi|_{T\sim Z}$ and $\succcurlyeq|_{T\sim Z}$ satisfy (24) for all (t^1,t^2) in $\succcurlyeq|_{T\sim Z}$. Equivalently, $\pi|_{T\sim Z}$ represents a completion of $\succcurlyeq|_{T\sim Z}$.

Conversely, suppose $\pi^o{:}\bigcup(T{\sim}Z){\to}\mathbb{R}$ represents a completion of $\succcurlyeq|_{T\sim Z}$. This is equivalent to π^o and $\succcurlyeq|_{T\sim Z}$ satisfying (24) for all (t^1,t^2) in $\succcurlyeq|_{T\sim Z}$. Also, because $\succcurlyeq|_{T\sim Z}$ equals $\succcurlyeq\cap(T{\sim}Z)^2$ by definition, we note that

$$\begin{aligned} t^1 \succ t^2 \quad &\text{iff} \quad t^1 \succ|_{T\sim Z}\, t^2 \quad \text{and} \\ t^1 \approx t^2 \quad &\text{iff} \quad t^1 \approx|_{T\sim Z}\, t^2 \end{aligned}$$

for all (t^1,t^2) in $\succcurlyeq|_{T\sim Z}$, where $\succ|_{T\sim Z}$ and $\approx|_{T\sim Z}$ are the asymmetric and symmetric parts of $\succcurlyeq|_{T\sim Z}$. The last two sentences imply that π^o and $\succcurlyeq$ satisfy (24) for all (t^1,t^2) in $\succcurlyeq|_{T\sim Z}$.

[12] This abstract concept is cleanest when the domain of the completion is left unspecified. The domain of the completion can be any collection of subsets of $\hat{A}$ that contains the subsets that $\hat{\succcurlyeq}$ compares. It is natural to let the domain be T when $\hat{A}$ is A and $\hat{\succcurlyeq}$ is $\succcurlyeq$, and to let the domain be $T{\sim}Z$ when $\hat{A}$ is $\bigcup(T{\sim}Z)$ and $\hat{\succcurlyeq}$ is $\succcurlyeq|_{T\sim Z}$.

Since A is the union of the disjoint sets A^c and A^s, we can let $\pi{:}A{\rightarrow}\mathbb{R}$ be the extension of π^o defined by

$$(\forall a)\ \pi(a) = \begin{pmatrix} \pi^o(a) & \text{if } a \in \bigcup(T{\sim}Z) \\ 0 & \text{if } a \in A^c{\sim}\bigcup(T{\sim}Z) \\ 0 & \text{if } a \in A^s{\sim}\bigcup(T{\sim}Z) \text{ and } \sigma(a) > 0 \\ -1 & \text{if } a \in A^s{\sim}\bigcup(T{\sim}Z) \text{ and } \sigma(a) = 0 \end{pmatrix} .$$

Our task is to show that π represents a completion of $\succcurlyeq$. Equivalently, our task is to show that π and $\succcurlyeq$ satisfy (24) for all (t^1,t^2) in $\succcurlyeq$. Since π is an extension of π^o, the previous paragraph shows that π and $\succcurlyeq$ satisfy (24) for all (t^1,t^2) in $\succcurlyeq|_{T\sim Z}$. Hence it remains to show that π and $\succcurlyeq$ satisfy (24) for all (t^1,t^2) in $\succcurlyeq \sim \succcurlyeq|_{T\sim Z}$. Accordingly, take any (t^1,t^2) in $\succcurlyeq \sim \succcurlyeq|_{T\sim Z}$. (This (t^1,t^2) will remain fixed through the remainder of the proof.)

Recall from Lemma 3.1 that $\succcurlyeq$ is partitioned by $\{\succ,\approx\}$, that $\succ$ is partitioned by $\{\overset{\sigma}{\succ},\overset{\beta}{\succ}\}$, and that $\approx$ is partitioned by $\{\overset{\sigma}{\approx},\overset{c}{\approx},\overset{\beta}{\approx}\}$. Since both $\overset{\beta}{\succ}$ and $\overset{\beta}{\approx}$ consist of pairs of nonterminal nodes, both $\overset{\beta}{\succ}$ and $\overset{\beta}{\approx}$ are subsets of $\succcurlyeq|_{T\sim Z}$. Thus, since (t^1,t^2) lies outside of $\succcurlyeq|_{T\sim Z}$ by assumption, (t^1,t^2) cannot be an element of either $\overset{\beta}{\succ}$ or $\overset{\beta}{\approx}$. Hence (t^1,t^2) must be an element of $\overset{\sigma}{\succ}$ or $\overset{\sigma}{\approx}$ or $\overset{c}{\approx}$. Thus it remains to show that

(25a) $$t^1 \overset{\sigma}{\succ} t^2 \;\Rightarrow\; \Sigma_{a\in t^1}\pi(a) > \Sigma_{a\in t^2}\pi(a)\ ,$$

(25b) $$t^1 \overset{\sigma}{\approx} t^2 \;\Rightarrow\; \Sigma_{a\in t^1}\pi(a) = \Sigma_{a\in t^2}\pi(a)\ , \text{ and}$$

(25c) $$t^1 \overset{c}{\approx} t^2 \;\Rightarrow\; \Sigma_{a\in t^1}\pi(a) = \Sigma_{a\in t^2}\pi(a)\ .$$

These implications are derived in the following three paragraphs.

Suppose $t^1 \overset{\sigma}{\succ} t^2$. By the definition of $\overset{\sigma}{\succ}$, $t^1{\in}T^s$ and there is some $\tilde{a}{\in}F(t^1)$ such that $t^1{\cup}\{\tilde{a}\} = t^2$ and $\sigma(\tilde{a}) = 0$. On the one hand, suppose $\tilde{a}{\notin}\bigcup(T{\sim}Z)$. Since $\pi(\tilde{a}) = -1$ by the definition of π, and since $\tilde{a}{\notin}t^1$ by $\tilde{a}{\in}F(t^1)$, we have that

(26) $$\Sigma_{a\in t^1}\pi(a) > \Sigma_{a\in t^1}\pi(a) + \pi(\tilde{a}) = \Sigma_{a\in t^2}\pi(a)\ .$$

On the other hand, suppose $\tilde{a}{\in}\bigcup(T{\sim}Z)$. Then $\tilde{a}$ is an element of some nonterminal node. Thus since every node has a last action, there must be nonterminal nodes $\tilde{t}^1$ and $\tilde{t}^2$ such that $\tilde{a}{\notin}\tilde{t}^1$ and $\tilde{t}^1{\cup}\{\tilde{a}\} = \tilde{t}^2$. Since $\sigma(\tilde{a}) = 0$, we have $\tilde{t}^1 \overset{\sigma}{\succ} \tilde{t}^2$. This implies $\Sigma_{a\in\tilde{t}^1}\pi^o(a) > \Sigma_{a\in\tilde{t}^2}\pi^o(a)$ because π^o represents a completion of $\succcurlyeq|_{T\sim Z}$ by assumption. This then implies $\pi^o(\tilde{a}) < 0$, which implies $\pi(\tilde{a}) < 0$. Therefore since $a{\notin}t^1$ by $a{\in}F(t^1)$, we again have (26). This establishes (25a).

Suppose $t^1 \overset{\sigma}{\approx} t^2$. By the definition of $\overset{\sigma}{\approx}$, either $t^1 \in T^s$ and there is some $\tilde{a} \in F(t^1)$ such that $t^1 \cup \{\tilde{a}\} = t^2$ and $\sigma(\tilde{a}) > 0$, or $t^2 \in T^s$ and there is some $\tilde{a} \in F(t^2)$ such that $t^2 \cup \{\tilde{a}\} = t^1$ and $\sigma(\tilde{a}) > 0$. Without loss of generality, consider the first case. On the one hand, suppose $\tilde{a} \notin \bigcup(T{\sim}Z)$. Since $\pi(\tilde{a}) = 0$ by the definition of π, we have that

$$\Sigma_{a \in t^1} \pi(a) = \Sigma_{a \in t^2} \pi(a) \ . \tag{27}$$

On the other hand, suppose $\tilde{a} \in \bigcup(T{\sim}Z)$. Then $\tilde{a}$ is an element of some nonterminal node. Thus since every node has a last action, there must be nonterminal nodes $\tilde{t}^1$ and $\tilde{t}^2$ such that $\tilde{a} \notin \tilde{t}^1$ and $\tilde{t}^1 \cup \{\tilde{a}\} = \tilde{t}^2$. Since $\sigma(\tilde{a}) > 0$, we have $\tilde{t}^1 \overset{\sigma}{\approx} \tilde{t}^2$. This implies $\Sigma_{a \in \tilde{t}^1} \pi^o(a) = \Sigma_{a \in \tilde{t}^2} \pi^o(a)$ because π^o represents a completion of $\succcurlyeq|_{T \sim Z}$ by assumption. This then implies $\pi^o(\tilde{a}) = 0$, which implies $\pi(\tilde{a}) = 0$. Therefore we again have (27). This establishes (25b).

Finally suppose $t^1 \overset{c}{\approx} t^2$. By the definition of $\overset{c}{\approx}$, either $t^1 \in T^c$ and there is some $\tilde{a} \in F(t^1)$ such that $t^1 \cup \{\tilde{a}\} = t^2$, or $t^2 \in T^c$ and there is some $\tilde{a} \in F(t^2)$ such that $t^2 \cup \{\tilde{a}\} = t^1$. Without loss of generality, consider the first case. Because $t^1 \in T^c$ and $\tilde{a} \in F(t^1)$, we have that $\tilde{a} \in A^c$ by Lemma A.1(a). On the one hand, suppose $\tilde{a} \notin \cup(T{\sim}Z)$. Since $\pi(\tilde{a}) = 0$ by the definition of π, we have that

$$\Sigma_{a \in t^1} \pi(a) = \Sigma_{a \in t^2} \pi(a) \ . \tag{28}$$

On the other hand, suppose $\tilde{a} \in \bigcup(T{\sim}Z)$. Then $\tilde{a}$ is an element of some nonterminal node. Thus since every node has a last action, there must be nonterminal nodes $\tilde{t}^1$ and $\tilde{t}^2$ such that $\tilde{a} \notin \tilde{t}^1$ and $\tilde{t}^1 \cup \{\tilde{a}\} = \tilde{t}^2$. Since $\tilde{a} \in A^c$, we have $\tilde{t}^1 \overset{c}{\approx} \tilde{t}^2$. This implies $\Sigma_{a \in \tilde{t}^1} \pi^o(a) = \Sigma_{a \in \tilde{t}^2} \pi^o(a)$ because π^o represents a completion of $\succcurlyeq|_{T \sim Z}$ by assumption. This then implies $\pi^o(\tilde{a}) = 0$, which implies $\pi(\tilde{a}) = 0$. Therefore we again have (28). This establishes (25c). □

B.2. ADDITIVE SEPARABILITY

Lemma B.2. $(\forall t)(\forall h)\ |t \cap F(h)| \in \{0, 1\}$. *In other words, a node t can contain no more than one element from each $F(h)$.*

Proof.[13] Consider any $(A, T, H, I, i^c, \rho, u)$. By Streufert (2012a, Theorem 1), this set-tree game is isomorphic to a sequence-tree game $(A, \bar{T}, \bar{H}, \bar{I}, \bar{\imath}^c, \rho, \bar{u})$ with agent recall.

---

[13]Although the lemma could be proved directly from the definition of a set-tree game, it is quicker to use Streufert (2012a, Theorem 1)'s isomorphism.

Now suppose there are t and h such that $|t \cap F(h)| > 1$. Then there are distinct a and a' such that $\{a, a'\} \subseteq t \cap F(h)$. First let $\bar{h} = (R_1|_{\mathcal{P}(\bar{T})})^{-1}(h)$, and note that $\{a, a'\} \subseteq \bar{F}(\bar{h})$ because $\bar{F}(\bar{h}) = F(h)$ by Streufert (2012a, Lemma A.5(c)). Second let $\bar{t} = (R|_{\bar{T}})^{-1}(t)$, and note that there are distinct m and n such that $\bar{t}_m = a$ and $\bar{t}_n = a'$. By the last two sentences there are distinct m and n such that $\{\bar{t}_m, \bar{t}_n\} \in \bar{F}(\bar{h})$. By Streufert (2012a, Lemma A.6(d)) this violates agent recall. □

Proof B.3 (for Lemma 5.1)**.** V is well-defined because Lemma B.2 shows that $t \cap F(h)$ has no more than one element for any t and any h.

It remains to show that, for any t, t is equal to $\{\check{t}_h | \check{t}_h \in A\}$ evaluated at $\check{t} = V(t)$. Accordingly take any t and note that

$$
\begin{aligned}
& \{\ \check{t}_h \mid \check{t}_h \in A\ \} \text{ evaluated at } \check{t} = V(t) \\
=\ & \{\ [V(t)]_h \mid [V(t)]_h \in A\ \} \\
=\ & \{\ a \mid (\exists h)\ a = [V(t)]_h\ \} \\
=\ & \{\ a \mid (\exists h)\ \{a\} = t \cap F(h)\ \} \\
=\ & \textstyle\bigcup_h (t \cap F(h)) \\
=\ & t\ ,
\end{aligned}
$$

where the first equality follows from manipulation, the second from $A = \bigcup_h F(h)$ by Streufert (2012a, Lemma A.2), the third from the definition of $V(t)$, the fourth from Lemma B.2, and the last from $A = \bigcup_h F(h)$ by Streufert (2012a, Lemma A.2). □

Proof B.4 (for Lemma 5.2)**.** First we argue that for all $\check{t} \in V(T)$

$$
\begin{aligned}
\text{(29)} \qquad \Sigma_h \varphi_h(\check{t}_h) &= \Sigma_{h|\check{t}_h \in F(h)}\, \varphi_h(\check{t}_h)\ +\ \Sigma_{h|\check{t}_h = o}\, \varphi_h(\check{t}_h) \\
&= \Sigma_{h|\check{t}_h \in F(h)}\, \pi(\check{t}_h)\ +\ \Sigma_{h|\check{t}_h = o}\, 0 \\
&= \Sigma_{h|\check{t}_h \in A}\, \pi(\check{t}_h) \\
&= \Sigma_{a \in \{\check{t}_h | \check{t}_h \in A\}} \pi(a) \\
&= \Sigma_{a \in V^{-1}(\check{t})} \pi(a)\ .
\end{aligned}
$$

The first equality holds because each φ_h has domain $\check{F}(h) = F(h) \cup \{o\}$, the second because each $\varphi_h = \pi|_{F(h)} \cup \{(o, 0)\}$ by assumption, and the third because each $\check{t}_h$ is in $F(h)$ iff it is in A. Then the fourth equality holds by manipulation and the fifth by Lemma 5.1.

Now suppose π represents a completion of $\succcurlyeq$. Then by the definition of $\check{\succcurlyeq}$ and Lemma 5.1, by representation (4), and by (29), we have that for all $\check{t}^1$ and $\check{t}^2$

$$\begin{aligned}
\check{t}^1 \check{\succ} \check{t}^2 &\Rightarrow V^{-1}(\check{t}^1) \succ V^{-1}(\check{t}^2) \\
&\Rightarrow \Sigma_{a \in V^{-1}(\check{t}^1)} \pi(a) > \Sigma_{a \in V^{-1}(\check{t}^2)} \pi(a) \\
&\Rightarrow \Sigma_h \varphi_h(\check{t}^1_h) > \Sigma_h \varphi_h(\check{t}^2_h) \ .
\end{aligned}$$

Identical reasoning shows $\check{t}^1 \approx \check{t}^2$ implies $\Sigma_h \varphi_h(\check{t}^1_h) = \Sigma_h \varphi_h(\check{t}^2_h)$. Thus $\langle \varphi_h \rangle_h$ represents a completion of $\check{\succsim}$.

Conversely suppose $\langle \varphi_h \rangle_h$ represents a completion of $\check{\succsim}$. Then by the definition of $\check{\succsim}$, by representation (8), by (29), and by inspection, we have that for all t^1 and t^2

$$\begin{aligned}
t^1 \succ t^2 &\Rightarrow V(t^1) \check{\succ} V(t^2) \\
&\Rightarrow \Sigma_h \, \varphi_h([V(t^1)]_h) > \Sigma_h \, \varphi_h([V(t^2)]_h) \\
&\Rightarrow \Sigma_{a \in V^{-1} \circ V(t^1)} \, \pi(a) > \Sigma_{a \in V^{-1} \circ V(t^2)} \, \pi(a) \\
&\Rightarrow \Sigma_{a \in t^1} \, \pi(a) > \Sigma_{a \in t^2} \, \pi(a) \ .
\end{aligned}$$

Identical reasoning shows $t^1 \approx t^2$ implies $\Sigma_{a \in t^1} \pi(a) = \Sigma_{a \in t^2} \pi(a)$. Thus π represents a completion of $\succsim$. □

References

Anderlini, L., D. Gerardi, and R. Lagunoff (2008): "A Super Folk Theorem for Dynastic Repeated Games," *Economic Theory*, 37, 357–394.

Battigalli, P. (1996): "Strategic Independence and Perfect Bayesian Equilibrium," *Journal of Economic Theory*, 70, 201–234.

Blackorby, C., D. Primont, and R. R. Russell (1978): *Duality, Separability, and Functional Structure: Theory and Economic Applications*. North-Holland.

Blume, L., A. Brandenburger, and E. Dekel (1991a): "Lexicographic Probabilities and Choice Under Uncertainty," *Econometrica*, 59, 61–79.

——— (1991b): "Lexicographic Probabilities and Equilibrium Refinements," *Econometrica*, 59, 81–98.

Blume, L. E., and W. R. Zame (1994): "The Algebraic Geometry of Perfect and Sequential Equilibrium," *Econometrica*, 62, 783–794.

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.

Fudenberg, D., and D. K. Levine (2006): "Superstition and Rational Learning," *American Economic Review*, 96, 630–651.

Gorman, W. M. (1968): "The Structure of Utility Functions," *Review of Economic Studies*, 35, 376–390.

GOVINDAN, S., AND T. KLUMPP (2002): "Perfect Equilibrium and Lexicographic Beliefs," *International Journal of Game Theory*, 31, 229–243.

GOVINDAN, S., AND R. WILSON (2001): "Direct Proofs of Generic Finiteness of Nash Equilibrium Outcomes," *Econometrica*, 69, 765–769.

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.

KOBAYASHI, H. (2007): "Folk Theorems for Infinitely Repeated Games Played by Organizations with Short-Lived Members," *International Economic Review*, 48, 407–549.

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.

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.

——— (2012a): "Specifying Nodes as Sets of Actions," University of Western Ontario, 34 pages.

——— (2012b): "Algebraic Characterizations of Consistency," University of Western Ontario, in progress.

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

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