A revealed-preference theory of strategic counterfactuals

Marciano Siniscalchi*

October 31, 2011

Preliminary: comments welcome

Abstract

The analysis of extensive-form games involves assumptions concerning players' beliefs at histories off the predicted path of play. However, the revealed-preference interpretation of such assumptions is unclear: how does one elicit probabilities conditional upon events that have zero ex-ante probability? This paper addresses this issue by proposing and axiomatizing a novel choice criterion for an individual who faces a general dynamic decision problem. The individual's preferences are characterized by a Bernoulli utility function and a *conditional probability system* Myerson (1986a). At any decision point, preferences are determined by conditional expected payoffs at the current node, as well as at all subsequent nodes. Thus, prior preferences contain enough information to identify all conditional beliefs. Furthermore, preferences are dynamically consistent, so prior preferences also determine behavior at subsequent decision nodes, including those that have zero ex-ante probability. In particular, the proposed criterion is consistent with, and indeed inspired by the game-theoretic notion of sequential rationality.

*Economics Department, Northwestern University, Evanston, IL 60208; `marciano@northwestern.edu`. I thank Amanda Friedenberg and participants at RUD 2011 for helpful comments on an early draft of this paper.

Keywords: Counterfactuals, dynamic games, conditional probability systems

1 Introduction

The analysis of extensive-form games often hinges on specific assumptions concerning the beliefs about the future actions of opponents that a player holds after one or more deviations from the predicted path of play. For a simple example, consider the four-legged centipede game in Fig. 1. In the usual, textbook analysis of backward induction in this game, Ann expects Bob to choose d at the second and fourth nodes, and Bob expects Ann to choose D at the first and third nodes. Thus, each player chooses rationally at each node, given his or her belief about the opponent's future play. In particular, when Ann contemplates a deviation from her initial choice of A, her beliefs at the *third* node are crucial to the backward-induction argument: if she thought that Bob was going to choose a at the fourth node (with high enough probability), it would be rational for her to choose A at the third node; anticipating this, Bob might want to play a at his *second* node, and the backward-induction prediction would unravel. Thus, for backward induction to "work" as intended, it is essential that Ann believe that Bob will choose d at the second node, but also that, were he to choose a instead, he would nonetheless choose d at the fourth node. In other words, it is essential to specify Ann's beliefs conditional upon an event (Bob's choice of a) to which she initially assigns zero probability.

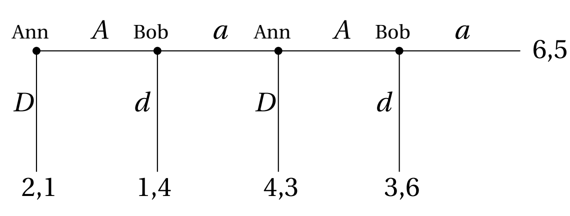

Figure 1: A centipede game

The literature on solution concepts for dynamic games has considered a variety of belief representations that allow the modeler to specify players' beliefs both on the path of play and following "surprise" moves: (consistent) *assessments*, as in Kreps and Wilson (1982), (reasonable) *belief systems*, as in Fudenberg and Tirole (1991), *conditional probability systems*, as in Myerson (1986a), and *consistent conjectures*, as in Pearce (1984). Extended representations of beliefs also play a fundamental role in the more recent literature on the epistemic foundations of solution concepts for dynamic games (see e.g. Ben-Porath, 1997; Stalnaker, 1998; Battigalli and Siniscalchi, 1999, 2002; Battigalli and Friedenberg, 2010, among others). In order to address central notions such as backward and forward induction, these contributions emphasize the importance of assumptions about players' interactive beliefs at different points in the game, including histories and information sets that the players do *not* expect to be reached.

All of the above representations of beliefs constitute a significant departure from the standard, subjective expected-utility paradigm. According to the latter, probabilities are a numerical representation of the individual's betting preferences. The events of interest—for instance, "Bob chooses a at the second node"—are precisely those that are *not* supposed to obtain. Hence, preferences conditional upon such events simply cannot be observed directly; this precludes a straightforward revealed-preference interpretation of the corresponding conditional probabilities. Furthermore, since the Bayesian updating formula does not apply when the conditioning event has zero prior probability, there is also no way to infer the value of such conditional probabilities indirectly, by first eliciting her ex-ante beliefs. To sum up, statements about probabilities conditional upon such counterfactual events cannot have a revealed-preference basis in the expected utility framework.

This paper proposes a way to address this issue. I consider a dynamic decision problem under uncertainty, and an individual characterized by preferences over Anscombe and Aumann (1963)-style acts at each decision point. In a game-theoretic application of the model, acts correspond to reduced-form strategies, or extensive-form plans of actions (that is, equivalence classes of strategies inducing the same outcomes for every choice profile of the opponents), and decision points correspond to histories where the player has a non-trivial choice to make.

I develop and axiomatize a model in which the individual's preferences are represented by a Bernoulli utility function on consequences, as usual, and a *conditional probability system*, or CPS for short. The representation entails a novel decision criterion that reduces to expected-utility maximization if the choice problem is one-shot, but otherwise takes into account the individual's conditional beliefs both at the current decision point, and at all subsequent ones. This makes it possible to uniquely and fully identify the individual's CPS from her ex-ante preferences alone. Furthermore, the individual is dynamically consistent, so ex-ante preferences fully characterize behavior throughout the decision tree.

As noted above, the use of CPSs in game theory was pioneered by Myerson (1986a); subsequent contributions show that deep connections exist between CPSs and more common representations of beliefs in extensive games, such as Kreps-Wilson assessments and Fudenberg-Tirole belief systems (Battigalli, 1996b; Fudenberg and Tirole, 1991). Furthermore, representations of beliefs via CPS have found several applications in the literature on dynamic interactive epistemology. Thus, the results in this paper provide a fully behavioral foundation for a decision model that underlies much of extensive-form game theory.

This paper is organized as follows. Section 2 provides a heuristic discussion of the proposed approach; Section 3 then discusses the connections with the existing literature; in particular, the analysis of CPS due to Myerson (1986b), and the notion of lexicographic expected utility maximization studied by Blume, Brandenburger, and Dekel (1991). Section 4 introduces the decision-theoretic setup. Section 5 provides the main characterization result. All proofs are in the Appendix.

[**Note:** an additional section is planned, focusing on the formal relationship between sequential rationality and sequential EU maximization.]

2 Heuristics

Return to the game of Fig. 1. Bob's strategy set is $S^b = \{dd, da, ad, aa\}$, in obvious notation. However, since the strategies *dd* and *da* are realization-equivalent, they are decision-theoretically indistinguishable from the point of view of Ann; for this reason, following e.g. Rubinstein (1991), we shall take Ann's domain of uncertainty to be Bob's *plans of action*, i.e. $\Omega^b = \{d, ad, aa\}$, in obvious notation. A CPS μ for Ann is then a pair (μ_1, μ_3) of probabilities over Ω^b (subscripts refer to nodes) that satisfies two constraints. First, at node 3, Ann must believe that Bob played a strategy choosing a at the second node:

$$\mu_3(\{ad, aa\}) = 1.$$

Second, μ_3 should be derived from μ_1 via Bayesian updating *whenever possible*:

$$\forall s^b \in \{ad, aa\}, \qquad \mu_1(\{s^b\}) = \mu_1(\{ad, aa\}) \cdot \mu_3(\{s^b\});$$

note that the above formula is written in such a way as to be valid even when $\mu_1(\{ad, aa\}) = 0$.

For simplicity, assume throughout this section that Ann is risk-neutral (or that numbers at terminal nodes represent "utils".)

The objective of this section is to describe how $\mu = (\mu_1, \mu_3)$ can be elicited, or experimentally determined by observing Ann's choice behavior. In particular, the main issue of interest here will be eliciting μ_3 when $\mu_1(\{ad, aa\}) = 0$, as in the backward-induction analysis discussed in the Introduction. I begin by sketching the main idea in Subsec. 2.1; key conceptual issues are discussed in Subsecs. 2.2, 2.3, and 2.4.

2.1 Eliciting beliefs at counterfactual histories

As usual in axiomatic decision theory, eliciting beliefs requires assuming that the individual is able to compare a sufficiently rich set of alternatives, not just those that are relevant to a particular decision problem. Here, Ann will be required to rank uncertain prospects that do not necessarily correspond to strategies in the game of Fig. 1. Specifically, Ann must be able to compare degenerate games in which (i) she can only move Across, so none of Bob's decision points are precluded by Ann's own actions, and (ii) the payoffs at remaining terminal nodes are not necessarily as in Fig. 1. One such comparison is illustrated in Fig. 2.1; note that Bob's payoffs are omitted, and that the degenerate games in question are identified by the corresponding payoff vectors, namely (x,y,z) and (x',y',z'). Thus, degenerate games can be mapped to to Savage-style *acts*; I shall use the two terms interchangeably in this section.

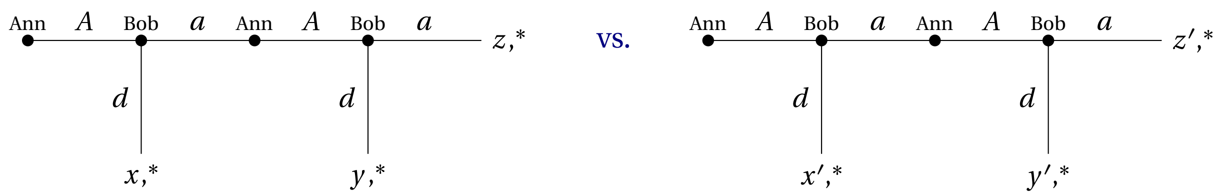

Figure 2: The game (x,y,z) vs. the game (x',y',z')

One crucial assumption is that Ann must be able to compare the degenerate games in Fig. 2.1 *under the assumption that Bob is going to behave exactly as in the original game of Fig. 1.* It is essential to note that it *is* possible to design an experiment that ensures that this is the case: Sec. 2.2 below indicates how to do so.

It will be sufficient to consider Ann's ex-ante preferences over such degenerate games. The decision rule proposed and axiomatized in this paper stipulates that, for every pair of degenerate games (acts) $f=(x,y,z)$ and $g=(x',y',z')$, Ann weakly prefers f to g if:

1. either $\mathrm{E}_{\mu_1}[f] > \mathrm{E}_{\mu_1}[g]$, or

2. $\mathrm{E}_{\mu_1}[f] = \mathrm{E}_{\mu_1}[g]$ and, if $\mu_1(\{ad,aa\})=0$, then $\mathrm{E}_{\mu_3}[f] \geq \mathrm{E}_{\mu_3}[g]$.

As usual, Ann strictly prefers f to g if she weakly prefers f to g, but does not weakly prefer g to f. I defer further discussion to Sec. 2.3; the key feature of this rule is that Ann's preferences at the *first* node are also influenced, in an essential way, by her conditional beliefs at the *third* node.

This feature is essential to elicit Ann's conditional beliefs, as can finally be illustrated. Suppose that Ann's CPS is given by

$$\mu_1(\{da, dd\}) = 1 \qquad \text{and} \qquad \mu_3(\{ad\}) = 1; \tag{1}$$

in other words, Ann's beliefs are consistent with backward induction. To elicit the probability that Ann initially assigns to Bob choosing d at the second node, an experimenter may first ask Ann to rank the acts $(x, 0, 0)$ and $(0, 1, 1)$ for $x \geq 0$. If Ann was an expected-utility maximizer, it would be possible to identify a unique value x^* such that she is just indifferent between $(x^*, 0, 0)$ and $(0, 1, 1)$; such value would identify the required probability. However, given the decision rule described above, and the CPS in Eq. (1), Ann strictly prefers $(x, 0, 0)$ to $(0, 1, 1)$ for any $x > 0$, but strictly prefer $(0, 1, 1)$ to $(0, 0, 0)$. This discontinuous behavior is clearly inconsistent with EU maximization. However, the fact that $(x, 0, 0)$ is strictly preferred, no matter how small $x > 0$ is, is enough to conclude (both intuitively and in light of the postulated decision criterion) that $\mu_1(\{dd, da\}) = 1$.

The experimenter can then ask Ann to rank the degenerate games $(z, x, 0)$ and $(z, 0, 1)$, for arbitrary z and $x \geq 0$. Under the maintained assumptions, Ann strictly prefers $(z, x, 0)$ regardless of the value of z, and for all $x > 0$; if $x = 0$, Ann is just indifferent. While this behavior is continuous, it could not arise if Ann were an EU maximizer who assigns probability zero to Bob's choice of a at the second node. However, (again, both intuitively and formally), this preference pattern reveals that Ann deems d at the fourth node more likely than a—indeed, that $\mu_3(\{ad\}) = 1$.

The above discussion shows how, if Ann's preferences conform to the postulated decision criterion, one can provide a consistent revealed-preference interpretation of her entire CPS. As noted above, the main formal contribution of this paper is to provide behavioral axioms that characterize the proposed criterion.

2.2 Strategic preferences: an experiment

As noted above, eliciting Ann's beliefs requires a preference ranking over a richer domain than just the strategies that Ann can choose in the game of Fig. 1. Furthermore, comparisons such as those illustrated in Fig. 2.1 are obviously only relevant to the elicitation of Ann's beliefs in the original game if Ann actually expects Bob to behave the same way in Fig. 1 and in the degenerate games of Fig. 2.1.

It is equally obvious that, under reasonable assumptions, Bob will *not* in general behave in the same way across these situations. Suppose for instance that Bob's payoffs in the degenerate games of Fig. 2.1 are exactly as in Fig. 1. Since Ann is explicitly precluded from ever going Down, Bob will clearly find it advantageous to go Across at the second node, and then Down at the fourth. If Ann anticipates this, her beliefs are clearly *not* the ones described in Eq. (1). In other words, the intended elicitation exercise will fail—it identifies Ann's CPS in the "wrong" game(s).

A similar issue was first noted by Mariotti (1995), who concluded that, even in one-shot games, eliciting players' beliefs in an incentive-compatible way may pose difficulties. I shall now suggest an experiment that can potentially overcome them.

The proposed experiment consists of a game with imperfect information that augments the game of interest in Fig. 1. Suppose one is interested in Ann's ranking of the acts (x, y, z) and (x', y', z'),[1] in addition, of course, to her behavior in the original game.

A coin is then flipped; Ann is informed of the outcome of the toss, but Bob is not. If the coin lands heads, then (loosely speaking) the original game is played, except that Bob ignores the outcome of the toss. If it lands tails, then Ann can choose one of two actions, say T or B. In both cases, following Ann's choice, Ann and Bob play a "truncated" version of the original game in which Ann cannot choose D (Bob's actions are unchanged). Payoffs are as follows:

[1]The procedure extends naturally to the ranking of any finite number of pairs of acts, and may be further generalized to elicit Ann's choices from larger sets of acts.

regardless of the outcome of the game, Bob receives a constant payoff, say 0; if Ann chooses T, then her payoffs are given by the vector (x, y, z); otherwise, they correspond to the vector (x', y', z'). Again, whenever he moves, Bob ignores the outcome of the coin toss. The game is depicted in Fig. 3.

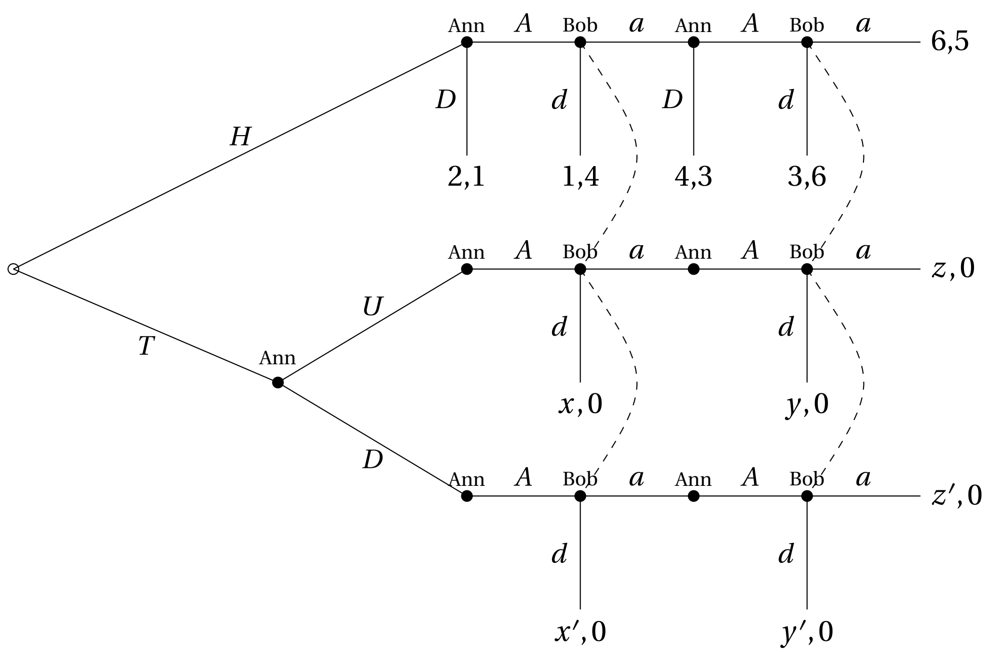

Figure 3: Eliciting Ann's ex-ante preferences: (x, y, z) vs. (x', y', z')

Notice that Bob's strategies in the expanded game are isomorphic to his strategies in the original one: he must choose whether to go $d$own or $a$cross the first time he gets to move, and then again in case Ann follows his choice of a with A. Furthermore, under mild assumptions on Bob's preferences,[2] his ranking of the plans d, ad and aa is the same in the original and in the expanded game. Ann, on the other hand, must specify her choices in the original game, but also effectively choose between the acts (x, y, z) and (x', y', z'). Crucially, the augmented game is set up so that both of *Ann's choices are subject to the same strategic uncertainty.*

[2]Independence with respect to mixtures with constants, as in Gilboa and Schmeidler (1989), suffices.

Assuming that Ann and Bob are indifferent to the timing of the resolution of non-strategic uncertainty, the experiment can also be described as follows. First, Bob is asked to describe his preferred plan of action in the original game; his answer is recorded, but not revealed to Ann. Then, Ann is asked to report her chosen plan of action int he original game, but also her ranking of the acts (x, y, z) and (x', y', z'). Finally, a coin is tossed; if it lands heads up, the original game is played according to the plans specified by Ann and Bob, and the corresponding payoffs are delivered. If instead it lands tails up, then Bob receives a payoff of 0, and Ann receives a payoff determind by which of the two acts (x, y, z) and (x', y', z') she chose, and Bob's recorded plan of action.

I emphasize that, regardless of which version of the experiment one implements, the preference ranking of (x, y, z) vs. (x', y', z') is the one Ann holds before any strategic uncertainty (i.e. any uncertainty about Bob's moves) is resolved. In other words, the experiments elicit Ann's *ex-ante* preferences. As will be clear from the analysis in Sec. 5, this is sufficient for the present purposes. Indeed, as was argued in the Introduction, the fact that it is not necessary to observe Ann's *conditional* preferences is a key desideratum of the present approach.

2.3 Sequential rationality and the sequential EU decision rule:

The proposed decision rule is related to the game-theoretic notion of (weak) *sequential rationality* (Kreps and Wilson, 1982; Reny, 1992; Battigalli, 1996a). Recall a strategy or plan s^a for Ann is sequentially rational given a CPS $\mu = (\mu_1, \mu_2)$ on Bob's strategies or plans if it is conditionally optimal at every node where Ann moves, and that is not ruled out by prior moves specified by s^a.[3] Thus, for any given CPS $\mu = (\mu_1, \mu_2)$, a strategy for Ann that specifies D at the first node

[3]Kreps and Wilson's notion of sequential rationality requires optimality at all nodes, and applies to strategies rather than plans; but choices at nodes that cannot be reached due to prior moves of Ann have no payoff implications, and hence no decision-theoretic significance. For these reasons, beginning with the cited papers by Reny, the literature on non-equilibrium solution concepts, as well as the epistemic foundations literature, have focused on the notion described in the text.

must be unconditionally optimal given μ_1; on the other hand, any strategy that specifies A at the first node must be optimal at the initial node given μ_1, and at the third node given μ_3.

In the example of Fig. 1, Ann can choose one of three plans, denoted D, AD, and AA (obvious notation). These correspond to the Savage-style acts $(2,2,2)$, $(1,4,4)$, and $(1,4,3)$ respectively. If Ann's beliefs are as in Eq. (1), the proposed sequential EU decision rule ranks these acts as follows: $(2,2,2)$ is strictly preferred to $(1,4,4)$, which is itself strictly preferred to $(1,4,3)$.

Thus, in the game of Fig. 1, the unique sequentially rational strategy corresponds to the unique best act according to the proposed decision rule. A bit more formally, if sequential rationality is viewed as defining, for every CPS, a choice correspondence defined over sets of strategies, then the sequential EU decision rule *rationalizes it* in the sense of choice theory.

The connection between sequential rationality and the sequential EU criterion is actually deeper. In the game obtained from Fig. 1 by removing Ann's initial choice of D (but changing nothing else), strategy AD is the unique sequentially rational choice for Ann; as noted above, the corresponding act $(1,4,4)$ is the unique best act in the set $\{(1,4,4),(1,4,3)\}$. Thus, the sequential EU ordering of the acts $(2,2,2),(1,4,4),(1,4,3)$ mirrors the implications of sequential rationality in suitably "nested" games.

[**Note:** formal results generalizing the above discussion can be provided, and will be included in a subsequent draft.]

Finally, notice that, in this simple example, the ranking of acts just described is "almost" lexicographic.[4] While there is a connection with lexicographic utility maximization, there are important differences for general games; see Sec. 3 for additional discussion, and Sec. B.2 for the formal details.

[4]If $\mathrm{E}_{\mu_1}[f] = \mathrm{E}_{\mu_1}[g]$, $\mu_1(\{ad, aa\}) > 0$, and $\mathrm{E}_{\mu_3}[f] < \mathrm{E}_{\mu_3}[g]$, the lexicographic EU ranking strictly prefers g to f, whereas the ranking in the text deems f and g indifferent.

2.4 Null vs. negligible events

According to Ann's CPS in Eq. (1), the event $E = \{ad, aa\}$, i.e. "Bob chooses a at the second node," has zero prior probability.

With expected-utility preferences, a zero-probability event is irrelevant for decisions: formally, the decision-maker is indifferent between any two acts that deliver the same outcomes at states outside the event under consideration. Savage (1954) calls such an event *null*. In the present setting, for the event E to be Savage-null, it must be the case that the generic acts (x, y, z) and (x', y', z') are deemed indifferent whenever $x = x'$. However, it was shown in Subsec. 2.1 that, if Ann's prior preferences conform to the sequential EU criterion, then she strictly prefers the act $(z, 1, 0)$ to the act $(z, 0, 1)$, regardless of the value of z.

Therefore, the event $E = \{ad, aa\}$ is *not* null, despite having zero prior probability. As was shown in Subsec. 2.1, the fact that E does matter for Ann's decisions at the *first* node is precisely what makes it possible to elicit Ann's beliefs conditional upon reaching the *third* node, by observing her prior preferences alone.

That said, the event $\{ad, aa\}$ satisfies a weaker irrelevance condition: given two acts (x, y, z) and (x', y', z'), $x > x'$ always implies that the former is strictly preferred, regardless of the value of the payoffs y, z, y', z'. Loosely speaking, $\{ad, aa\}$ is decision-theoretically irrelevant insofar as *strict preferences* are concerned, but does matter if $x = x'$. I call such an event *negligible*; as will be clear in Sec. 5, negligible events play a key role in the axiomatization of the decision rule proposed here.

Finally, assume that Ann's conditional preferences at the third node (that is, her preferences conditional upon observing that Bob chose a at the second node) are consistent with EU maximization, and represented by the conditional probability μ_3. Then, together with her ex-ante preferences defined above, Ann's conditional preferences satisfy the *Dynamic Consistency* property (i.e. Savage's Sure-Thing Principle). Specifically. her *ex-ante* ranking of the acts (x, y, z) and (x, y, z') is exactly the same as her *conditional* ranking. This is true even though the event

$\{ad, aa\}$ has zero ex-ante probability. One implication of Dynamic Consistency is that the experimenter can infer Ann's conditional preferences by observing her prior ones; this is also possible with standard EU preferences, but only when the conditioning event is not Savage-null.[5]

Dynamic Consistency will be a key axiom in the behavioral characterization of the proposed decision rule.

3 Related Literature

3.1 Conditional EU maximization: Myerson (1986b)

Myerson (1986b) (see also Myerson (1997), Chapter 1), considers a collection of conditional preferences, indexed by non-empty events in a (finite) state space. He provides axioms under which each conditional preference is consistent with EU maximization, and furthermore the corresponding collection of (conditional) probabilities constitute a CPS (i.e. satisfy Bayes' rule whenever possible).

There are two main differences with the present paper. First, because (conditional) preferences in Myerson (1986b) are consistent with EU, they do not actually reflect any consideration of sequential rationality. For instance, given the CPS defined in Eq. (1) for the game in Fig. 1, the strategies *AD* and *AA* yield the same ex-ante expected utility of 1, and hence are deemed indifferent at the initial node. By way of contrast, gien the CPS under consideration, the sequential ordering proposed in this paper deems *AD* strictly superior at the initial node.

Second, in Myerson's approach, it is *not* always possible to elicit conditional beliefs (and preferences) by observing prior preferences alone. The reason is that his axioms allow for the possibility that the conditioning events be Savage-null for the ex-ante preference (or for a preference conditional upon a superset of the event in question). Again, in the game of Fig. 1,

[5]There is a well-known duality between Bayesian updating and dynamic consistency: see e.g. Ghirardato (2002) and Epstein and Le Breton (1993).

with the CPS in Eq. (1), *AD* and *AA* are deemed indifferent at the initial node, even though Ann strictly prefers the former at the third node. Indeed, the event $\{ad, aa\}$ is Savage-null in Myerson's setting. By way of contrast, the proposed approach ensures that all conditioning events—even those that are negligible, and hence have zero ex-ante probability—are not Savage-null. Furthermore, Dynamic Consistency is explicitly assumed. Consequently, a foundation for sequential EU maximization can be provided on the sole basis of prior preferences, which—unlike conditional preferences—are directly observable.

3.2 Lexicographic EU maximization: Blume et al. (1991)

As noted above, in the game of Fig. 1, the decision rule described in Sec. 2.3 is "almost" the same as lexicographic EU maximization. Recall that a lexicographic EU maximizer is characterized by a utility function u and a vector $(\pi^1, \ldots, \pi^n)$ of probabilities over the relevant state space; such an individual then weakly prefers an act f to another act g if the vector $\big(\mathrm{E}_{\pi^1} u(f), \ldots, \mathrm{E}_{\pi^n} u(f)\big)$ is lexicographically greater than the vector $\big(\mathrm{E}_{\pi^1} u(g), \ldots, \mathrm{E}_{\pi^n} u(g)\big)$. Blume et al. (1991) provide an axiomatic foundation for lexicographic EU maximization on a finite state space.

The differences between sequential and lexicographic EU maximization are stark in other games, even relatively simple ones. Consider the following extensive game form: Ann moves first, and chooses between *Out* and *In*. If she chooses *Out*, the game ends; otherwise, Bob chooses among the actions t, m, b. If Bob chooses b, the game ends; otherwise, Ann observes Bob's choice, and a simultaneous-moves game ensues, in which Ann's actions are denoted by U, D and Bob's actions by l, r. Fig. 4 depicts the game and indicates Ann's payoffs at each terminal node (Bob's payoffs are immaterial to the argument, and hence omitted).

Bob's strategy set consists of 12 tuples of the form $a_1 a_t a_m$, where $a_1 \in \{t, m, b\}$ and $a_t, a_m \in \{l, r\}$; however, as in the previous example, it is enough to identify realization-equivalent strategies and focus on the set of plans $S^b = \{a_1 a_2 \ : \ a_1 \in \{t, m, b\}, \ a_2 \in \{l, r\}\}$. Ann's CPS now comprises three probability distributions over S^b: $\mu = (\mu_1, \mu_t, \mu_b)$, where μ_1 is Ann's belief at

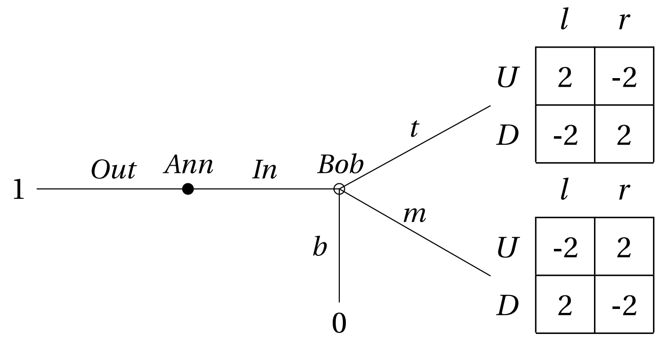

Figure 4: Lexicographic vs. sequential maximization

the initial node, and μ_t and μ_b represent Ann's beliefs after observing Bob's choice of t and b respectively. Assume in particular that

$$\mu_1\left(\left\{a_1a_2 \in S^b : a_1 = b\right\}\right) = 1, \qquad \mu_t(\{tl\}) = \mu_m(\{mr\}) = 1. \tag{2}$$

Given these beliefs, Ann chooses *Out* in any sequentially rational strategy.[6].

In the interest of conciseness, I shall not define here the sequential EU ordering for an arbitrary CPS, but simply indicate what it entails for the specific one defined in Eq. (2): the strategy s is weakly preferred to s' if either it yields a strictly higher expected payoff at the initial node (which is computed with respect to μ_1), or else the two strategies yield the same ex-ante expected payoff, but s yields a weakly higher expected payoff conditional upon reaching *each* of the two simultaneous-moves subgames (which payoffs are computed using μ_t and μ_m respectively). The reader is referred to Sec. 4 and Def. 2 for details.

As a consequence of this definition, a strategy s for Ann that chooses *Out* at the initial node is indeed strictly preferred to any strategy s' that chooses *In* there, because it yields a strictly higher expected payoff ex-ante.[7]. Once again, the proposed decision rule is consistent with

[6]There are four such strategies, which differ in the choices made in the two simultaneous-moves subgames; however, such choices are irrelevant as far as (weak) sequential rationality is concerned, because choosing *Out* prevents the subgames from being reached.

[7]Any two strategies that initially choose *Out* correspond to the constant act that yields 1 regardless of Bob's

sequential rationality.

The key observation is now that *sequential EU maximization may be unable to rank certain inferior strategies*. For instance, consider *InUD* and *InDU*: given μ_1, both yield an expected payoff of 0, but *InUD* yields an expected payoff of 2 given μ_t and -2 given μ_m, whereas the opposite is true for *InDU*. One strategy is better in case Bob unexpectedly chooses t, and the other is better in case Bob unexpectedly chooses m. Intuitively, in order to rank these strategies, one needs to decide which of these unexpected choices is "more likely"; however, a CPS for the game in Fig. 4 cannot provide this information.

Lexicographic expected utility instead yields a *complete* ordering, precisely because it can convey not just Ann's assessment that Bob is virtually certain not to choose t or m, but also her assessment of the "relative likelihood" of t vs. m in case Bob decides not to choose b after all. For instance, take $n=3$, $\pi^1=\mu_1$, $\pi^2=\mu_t$ and $\pi^3=\mu_m$: since $\mu_t(\{a_1a_2 : a_1=t\})=1$, intuitively Bob's choice of t is deemed "infinitely more likely" than his choice of m (strategies that choose t are assigned positive probability by a lower-indexed element of the lexicographic system), and this implies that *InUD* will be strictly preferred to *InDU*.

One way to interpret this fact is to note that the sequential EU criterion is defined here with reference to a specific game under consideration: the conditioning events are precisely those events that a player can actually observe in the game of interest. In Fig. 4, Ann is never confronted with the information that b was not chosen; correspondingly, the sequential EU criterion disregards whatever beliefs Ann may hold conditional upon observing the event "Bob chose either t or m." Indeed, if the game in Fig. 4 is the one of interest, Ann's CPS need not even indicate a belief conditional upon this event.

However, it is worth digging a little deeper. Suppose that the game of 4 is modified by introducing a (dummy) information set for Ann, following Bob's choice of t or m, but before

choices in the game; hence, such strategies all have expected payoff equal to 1 with respect to all three measures μ_1, μ_t, μ_m, and hence are equally good for Ann the simultaneous-moves subgames; at such information set, Ann has a single action available. A CPS for Ann in this modified game would have to include her beliefs conditional upon the event $\{a_1 a_2 : a_1 \in \{t, m\}\}$; even though no actual choice is made at the new information set, this would be sufficient to determine the ranking of *InUD* vs. *InDU* according to the proposed sequential ordering.

Even so, it is important to observe that this ranking is immaterial as far as sequential rationality is concerned. First, the choice of *Out* at the initial node remains strictly preferable if the intermediate dummy information set is added. Second, and perhaps more interestingly, suppose one further modifies the game by removing the choice of *Out* at the initial node. Regardless of the ranking of *InUD* vs. *InDU*, it is obvious that *neither of them is optimal* for Ann: the strategy *InUU* is clearly superior if Ann expects Bob to follow t with l and m with r; it is also sequentially rational in the game with *Out* removed, regardless of whether or not the intermediate information set is present.

To summarize, from a conceptual standpoint, while the proposed ordering can be incomplete (and the axiomatics will need to work around this difficulty), it provides enough information to pin down conditional beliefs (and utilities), *and* also to identify sequentially rational strategies. The additional information provided by lexicographic systems is not required for the purposes of characterizing sequential rationality.

I conclude with a more technical, but nonetheless important observation. The literature on the epistemic foundations of solution concepts indicates that, even if the game one is interested in is finite, it is often necessary, or at least convenient, to embed it in a structure featuring *uncountable* sets of states that describe players' interactive beliefs: see e.g. Brandenburger and Dekel (1993), Battigalli and Siniscalchi (1999). This is the case, for instance, if one is interested in the analysis of forward induction, unadulterated by exogenous assumptions on hierarchical beliefs: for an elaboration of this point, see Battigalli and Siniscalchi (2002). Defining and axiomatizing sequential EU maximization when acts are defined over an uncountable state space does not pose any problem. On the other hand, to the best of my knowledge the existing axiomatizations of lexicographic probability systems requires a finite state space.[8] Indeed, merely defining lexicographic systems on an uncountable state space requires some care: see e.g. Brandenburger, Friedenberg, and Keisler (2008).

4 Setup

4.1 Basics

The following notation is standard. Consider a set Ω (the state space) and a sigma-algebra Σ of subsets of Ω (events). Information is described via an event tree. Formally, fix a finite collection $\mathscr{F} = (\mathscr{F}_0, \ldots, \mathscr{F}_T)$ of progressively finer, measurable partitions of Ω, with $\mathscr{F}_0 = \{\Omega\}$. For every $\omega \in \Omega$ and $t = 0, \ldots, T$, denote by $\mathscr{F}_t(\omega)$ the cell of the partition that contains ω. It is convenient to refer to a pair (t, ω) as a *node*, which evokes the underlying event tree.

As suggested in Sec. 2, if one is interested in modeling a player's beliefs about her opponents' play in a finite extensive game, Ω can be taken to be the set of all profiles of plans of actions, endowed with the discrete sigma-algebra; equivalently, Ω may be defined as the set of strategy profiles S_{-i} played by i's opponents, with Σ being the algebra generated by the equivalence classes of realization-equivalent strategy profiles. Payoff uncertainty for a can be modeled by assuming that $\Omega = \Theta \times S_{-i}$, with Θ a measurable set of parameters. Finally, with an eye toward interactive epistemology, Player i's hierarchical conditional beliefs may be captured by taking $\Omega = S_{-i} \times T_{-i}$, where each t_{-i} represents a profile of hierarchical conditional beliefs for i's opponents.[9]

[8] Fishburn and LaValle (1998) axiomatize a theory of "subjective expected lexicographic utility" for arbitrary state spaces, which is related to, but different from lexicographic EU maximization. Its basic ingredients are a vector of Bernoulli utilities and a "matrix probability" (not an LPS).

[9] To elaborate, one can envision a conditional version of the Epstein and Wang (1996) construction of

Denote by $B_0(\Sigma)$ the set of Σ-measurable real functions with finite range, and by $B(\Sigma)$ its sup-norm closure. The set of bounded, (finitely) additive probability measures on Σ is denoted by $pr(\Sigma)$. For any probability measure $\pi \in pr(\Sigma)$ and function $a \in B(\Sigma)$, let $\mathrm{E}_\pi[a] = \int_\Omega a \, d\pi$, the standard Dunford integral of a with respect to π. When no confusion can arise, I will sometimes omit the square brackets. Finally, $a \circ b : \mathscr{X} \to \mathscr{Z}$ denotes the composition of the functions $b : \mathscr{X} \to \mathscr{Y}$ and $a : \mathscr{Y} \to \mathscr{Z}$.

I adapt Renyi's classic definition (Rényi, 1955) to the setting at hand.

Definition 1 *A **conditional probability system** (CPS) is a collection* $(\mu_{t,\omega})_{t=0,\dots,T;\omega\in\Omega}$ *such that, for every node* (t,ω)*,*

(0) if $\omega' \in \mathscr{F}_t(\omega)$*, then* $\mu_{t,\omega'} = \mu_{t,\omega}$*;*

(1) $\mu_{t,\omega} \in pr(\Sigma$ *and* $\mu_{t,\omega}(\mathscr{F}_t(\omega)) = 1$*;*

(2) for every $E \in \Sigma$ *with* $E \subset \mathscr{F}_{t+1}(\omega)$*,*

$$\mu_{t,\omega}(E) = \mu_{t+1,\omega}(E) \cdot \mu_{t,\omega}(\mathscr{F}_{t+1}(\omega)). \tag{3}$$

The collection of CPS on $(\Omega, \mathscr{F})$ *is denoted by* $pr(\Sigma, \mathscr{F})$*. Given* $a \in B(\Sigma)$*, write* $\mathrm{E}_{t,\omega} a = \mathrm{E}_{\mu_{t,\omega}} a$*.*

Eq. (3) is the usual Bayesian updating formula, except that it is required to hold even if $\mathscr{F}_t(\omega)$ has zero ex-ante probability. Also note that, if $\mu_{t,\omega}(\mathscr{F}_{t+1}(\omega)) = 0$, then also $\mu_{t,\omega}(E) = 0$, so in this case Eq. (3) holds trivially.

4.2 Sequential dominance

Turn now to a key ordering of random variables that is an essential ingredient of this paper.

preference-based type structures. The intuition is that, once players' first-order beliefs have been elicited, with $\Omega = S_{-i}$ for each player i, one can define players' preferences over their opponents' strategies *and* first-order beliefs (or preferences), and then proceed iteratively. At each step, the characterization result of Theorem 1 applies verbatim.

Definition 2 *Fix a node* (t,ω) *and a CPS* $\mu \in pr(\Sigma,\mathscr{F})$. *A node* (τ,ω^*) *is* ***critical given*** μ ***at*** (t,ω) *if* $\tau \geq t$, $\omega^* \in \mathscr{F}_t(\omega)$, *and* $\tau > t$ *implies* $\mu_{\tau-1,\omega^*}(\mathscr{F}_\tau(\omega^*)) = 0$.

Given $a,b \in B(\Sigma)$, a *is* ***sequentially greater than*** b *given* μ *at* (t,ω), *written* $a \geq^{\mu}_{t,\omega} b$, *if, for every node* (τ,ω^*) *that is critical given* μ *at* (t,ω), *and such that* $\mathrm{E}_{\tau,\omega^*}a < \mathrm{E}_{\tau,\omega^*}b$, *there exists another critical node* (σ,ω^*) *for* (t,ω) *given* μ, *with* $\sigma \in \{t,\ldots,\tau-1\}$, *such that* $\mathrm{E}_{\sigma,\omega^*}a > \mathrm{E}_{\sigma,\omega^*}b$.

Furthermore, a *is* ***sequentially strictly greater than*** b *given* μ *at* (t,ω), *written* $a >^{\mu}_{t,\omega} b$, *if* $a \geq^{\mu}_{t,\omega} b$ *and not* $b \geq^{\mu}_{t,\omega} a$.

In words, a node (τ,ω^*) is critical at (t,ω) if it is either (t,ω) itself (in which case the restrictions on conditional probabilities hold trivially), or else it can be reached from (t,ω), but is "surprising," or "unexpected," at time $\tau-1$.

To "unpack" this definition, note first that $a \geq^{\mu}_{t,\omega} b$ requires that $\mathrm{E}_{t,\omega}a \geq \mathrm{E}_{0,\omega}b$: this follows by taking (t,ω) itself as the critical node. If in particular $\mathrm{E}_{t,\omega}a > \mathrm{E}_{t,\omega}b$, then Def. 2 imposes no further requirement. That is, $\mathrm{E}_{t,\omega}a > \mathrm{E}_{t,\omega}b$ is sufficient to conclude that $a \geq^{\mu}_{t,\omega} b$.

If instead $\mathrm{E}_{t,\omega}a = \mathrm{E}_{t,\omega}b$, then comparisons at subsequent critical nodes can break this tie. Loosely speaking, consider the path from the node (t,ω) to some terminal node (T,ω^*) that can be reached from it [so $T \geq t$ and $\omega^* \in \mathscr{F}_t(\omega)$]. Then, for a to be deemed sequentially greater than b, it must be the case that the *first strict ranking* of the conditional expectations of a and b at a node (τ,ω^*) that is critical for (t,ω) is in favor of a: that is, $\mathrm{E}_{\tau,\omega^*}a > \mathrm{E}_{\tau,\omega^*}b$. Furthermore, this must be true along every path from (t,ω) to a consistent terminal node.

It is worth noting that the sequential ordering just defined can also be described as a product of lexicographic orders; the details are in Sec. B.2.

It is intuitively clear that the relation $\geq^{\mu}_{t,\omega}$ is an *incomplete* ordering. For a very simple example, let $\Omega = \{s_1,s_2,s_3,s_4\}$ and represent probability distributions and functions on Ω as tuples $(\alpha,\beta,\gamma,\delta) \in \mathbb{R}^4$. Consider the partitions $\mathscr{F}_0 = \{\Omega\}$ and $\mathscr{F}_1 = \{\{s_1\},\{s_2\},\{s_3\},\{s_4\}\}$, and the (unique) CPS μ such that $\mu_{0,\omega} = (\frac{1}{2},\frac{1}{2},0,0)$. Then $a = (0,0,1,0)$ and $b = (0,0,0,1)$ are not ranked by $\geq^{\mu}_{0,\omega}$. The intuition is clear: interpreting values as payoffs or utilities, a is better conditional on $\{s_3\}$, but b is better conditional on $\{s_4\}$. The notion of sequential comparison in Def. 2 does *not* require the individual to rank the relative likelihood of surprise events, so it does not imply a clear ranking of a vs. b.

Moreover, a function may have strictly smaller conditional expectations at certain nodes relative to another function, and still be sequentially greater. Continuing the example above, consider $a' = (1,0,0,0)$, $b' = (0,\frac{1}{2},0,0)$ and $c' = (0,0,1,0)$: then $a' \geq^{\mu}_{0,\omega} b' \geq^{\mu}_{0,\omega} c'$, even though b' has higher conditional expectation given $\{s_2\}$ and c' beats a' conditional on $\{s_3\}$. Again, the intuition is straightforward: according to logic of Def. 2, comparisons of conditional expectations at later nodes are only used to break ties at earlier ones; in particular, if the ranking is clear on the basis of ex-ante expectations, no further analysis is required.

This last property should not be surprising: even with full-support beliefs, an individual may choose a course of action that turns out to be inferior to another in certain contingencies, provided the probability of such contingencies is small enough. Of course, if "courses of action" correspond to strategies in a game (or decision problem), then neither strategy would be *sequentially rational* in this case. But, one could still say which of the two is less bad!

5 Behavioral analysis

I adopt the decision setting popularized by Anscombe and Aumann (1963): acts are simple measurable maps from Ω to a convex set X of prizes. The set of all acts is denoted by F; for $t = 0,\dots,T-1$, the set of $\mathscr{F}_t$-measurable acts is instead denoted by F_t. For reasons that will become clear in the following, it is also convenient to let $F_T = F$. With the usual abuse of notation, I do not distinguish between the constant act f that delivers the prize x in every state, and x itself.

The main object of interest is a *preference system*, i.e. a collection $(\succsim_{t,\omega})_{t=0,\dots,T;\omega\in\Omega}$ of preferences on $\mathscr{F}$ adapted to $\mathscr{F}$:

$$\forall \omega,\omega' \in \Omega,\ t=0,\dots,T: \qquad \omega' \in \mathscr{F}_t(\omega) \quad \Longrightarrow \quad \succsim_{t,\omega'} = \succsim_{t,\omega}. \tag{4}$$

The axioms follow. I divide them into three groups: axioms that are imposed at each node on *all* acts, axioms imposed at each node *subsets* of acts, and axioms relating preferences at nodes.

Axiom 1 (Preorder) *For all nodes* (t,ω)*,* $\succcurlyeq_{t,\omega}$ *is transitive and reflexive.*

Axiom 2 (Non-Degeneracy) *For all nodes* (t,ω)*, not for all* $f,g \in F$*,* $f \succcurlyeq_{t,\omega} g$*.*

Axiom 3 (Monotonicity) *For all nodes* (t,ω) *and acts* $f,g \in F$*: if* $f(\omega') \succcurlyeq_{t,\omega} g(\omega')$ *for all* $\omega' \in \Omega$*, then* $f \succcurlyeq_{t,\omega} g$*.*

Axiom 4 (Independence) *For all nodes* (t,ω)*, acts* $f,g,h \in F$ *and* $\alpha \in [0,1]$*:* $f \succcurlyeq_{t,\omega} g$ *if and only if* $\alpha f + (1-\alpha)h \succcurlyeq_{t,\omega} \alpha g + (1-\alpha)h$*.*

The next group of axioms consists of versions of completeness and continuity, restricted to suitable subclasses of acts. First, preferences over constant acts satisfy the standard conditions:

Axiom 5 (Prize Completeness) *For all nodes* (t,ω)*, the restriction of* $\succcurlyeq_{t,\omega}$ *to* X *is complete.*

Axiom 6 (Prize Archimedean) *For all nodes* (t,ω) *and prizes* $x,y,z \in X$*: if* $x \succ_{t,\omega} y \succ_{t,\omega} z$*, then there exist* $\alpha,\beta \in [0,1]$ *such that* $\alpha x + (1-\alpha)z \succ_{t,\omega} y$ *and* $y \succ_{t,\omega} \beta x + (1-\beta)z$*.*

Second, I impose completeness and continuity at a node (t,ω) whenever acts are allowed to differ only outside of cells in $\mathscr{F}_{t+1}$ that are suitably "null" for the preference $\succcurlyeq_{t,\omega}$. The intuition is that, when comparing acts that agree on "null" sets, the logic of sequential dominance cannot help break ties, so the ranking of such acts is wholly determined by their conditional evaluation given $\mathscr{F}_t(\omega)$. Therefore, such rankings should satisfy completeness and continuity.

It is necessary to define a suitable notion of null events:

Definition 3 *Fix* $t < T$ *and a node* (t,ω)*. An event* $N \in \Sigma$ *is* $\succcurlyeq_{t,\omega}$***-negligible*** *if:*

1. *there are acts* $f,g \in F$ *such that* $f(\omega') = g(\omega')$ *for all* $\omega' \in N$ *and* $f \succ_{t,\omega} g$*, and*

2. *if $f, g \in F$ satisfy the conditions in (1), and $f', g' \in F$ are such that $f(\omega') = f'(\omega')$ and $g(\omega') = g'(\omega')$ for all $\omega' \notin N$, then $f' \succ_{t,\omega} g'$.*

To interpret, recall that an event $N \in \Sigma$ is *Savage-null* for a preference relation $\succcurlyeq$ on F if, whenever $f, g \in F$ satisfy $f(\omega) = g(\omega)$ for all $\omega \in \Omega \setminus N$, $f \sim g$. Thus, Part 1 of Def. 3 requires that the complement of a negligible event N not be Savage-null for $\succcurlyeq_{t,\omega}$. In other words, the prizes that two acts f, g deliver on the complement of N should matter for the individual.

Part 2 of the definition then adds that prizes delivered on the complement of N should actually matter *overwhelmingly*: if two acts f and g coincide on N and are strictly ranked, then, no matter how f and g are modified at states in N, the same ranking results. In other words, there is no trade-off: if f is better than g at states outside N, it does not matter how much better g is than f at states in N.

The definition implies that $\emptyset$ is always $\succcurlyeq_{t,\omega}$-negligible, whereas Ω never is. Moreover, under Axiom 9 below, any union of elements of $\mathscr{F}_{t+1}$ that do *not* refine $\mathscr{F}_t(\omega)$ is $\succcurlyeq_{t,\omega}$-negligible, whereas the event $\mathscr{F}_t(\omega)$ itself is not.

It is easy to see that a Savage-null event N is negligible.[10] However, the converse is false, as the example in Sec. 2.4 demonstrates. This "wedge" between Savage-null and negligible events is what enables one to elicit preferences at future nodes that are currently assigned zero probability.

Now restrict attention to specific negligible events. Let $\mathscr{N}_{t,\omega}$ be the (possibly empty) collection of *non-empty* $\succcurlyeq_{t,\omega}$-negligible events that are unions of future conditioning events, and follow node (t, ω). Since $\mathscr{F}_T$ is the finest partition under consideration, this is the same as the collection of negligible events that are unions of elements in $\mathscr{F}_T$. Formally, for $t < T$, let

$$\mathscr{N}_{t,\omega} = \left\{ N \in \Sigma \setminus \{\emptyset\} : N = \mathscr{F}_t(\omega) \cap \bigcup_{\omega^0 \in N} \mathscr{F}_T(\omega^0) \text{ and } N \text{ is } \succcurlyeq_{t,\omega} \text{-negligible} \right\}. \tag{5}$$

[10] By Non-degeneracy and Transitivity, the complement of a Savage-null event cannot be Savage-null. Then, if f, g, f', g' are as in Part 2 of Def. 3, if N is negligible, then $f' \sim_{t,\omega} f \succ_{t,\omega} g \sim_{t,\omega} g'$.

Also, let $\mathscr{N}_{T,\omega} = \{\emptyset\}$: this ensures that, by imposing the two axioms below, all conditional preferences at the terminal date T conform to expected utility.

Under Axioms 9 and 10, it turns out that the elements of $\mathscr{N}_{t,\omega}$ are ordered by set inclusion, so their union is also a negligible event. Formally, however, for Axioms 7 and 8 to be well-posed, it is enough that said union be measurable; this follows from the finiteness of $\mathscr{F}_T$. Henceforth, for all nodes (t,ω), let

$$N_{t,\omega} = \bigcup_{N_{t,\omega} \in \mathscr{N}_{t,\omega}} N, \tag{6}$$

with the understanding that $N_{t,\omega} = \emptyset$ if $\mathscr{N}_{t,\omega} = \emptyset$.

The next two axioms state that preferences at a node (t,ω) are complete and continuous when ranking acts that are constant on every negligible event. Intuitively, this means that the only deviations from standard, EU behavior that arise have to do with the evaluation of prizes delivered in negligible events.

Axiom 7 (Non-Negligible Completeness) *For all nodes (t,ω) and acts $f,g \in F$ such that $f(\omega^0) = g(\omega^0)$ for all $\omega^0 \in N_{t,\omega}$: either $f \succcurlyeq_{t,\omega} g$ or $g \succcurlyeq_{t,\omega} f$ (or both).*

Axiom 8 (Non-Negligible Archimedean) *For all nodes (t,ω) and acts $f,g,h \in F$ with $f(\omega^0) = g(\omega^0) = h(\omega^0)$ for all $\omega^0 \in N_{t,\omega}$: if $f \succ_{t,\omega} g \succ_{t,\omega} h$, then there exist $\alpha,\beta \in [0,1]$ such that $\alpha f + (1-\alpha)h \succ_{t,\omega} g$ and $g \succ_{t,\omega} \beta f + (1-\beta)h$.*

Finally, turn to the axioms pertaining to dynamic choice. The terminology is borrowed from Epstein and Schneider (2003). The first axiom ensures that, when conditioning on an event, only prizes delivered at states in that event matter. The second, key axiom is the standard consistency requirement, applied to decision trees based upon the filtration $\mathscr{F}$. However, unlike standard treatments, I do *not* restrict it to conditioning events that are not Savage-null, even when strict preferences are involved.[11]

[11] Similarly, Blume et al. (1991) impose State-Independence at all states, rather than at non-null states only.

Axiom 9 (Conditional Preference—CP) *For all nodes (t,ω) and $f, f' \in F$: if $f(\omega') = f'(\omega')$ for all $\omega' \in \mathscr{F}_t(\omega)$, then $f \sim_{t,\omega} f'$.*

Axiom 10 (Dynamic Consistency—DC) *For every node (t,ω) and $f, f' \in F$ such that $f(\omega') = f'(\omega')$ for $\omega' \notin \mathscr{F}_{t+1}(\omega)$: $f \succcurlyeq_{t+1,\omega} f'$ if and only if $f \succcurlyeq_{t,\omega} f'$.*

Note that, as a consequence of Axiom 10 and Non-degeneracy, *no* conditioning event $\mathscr{F}_{t+1}(\omega)$ is Savage-null for the relation $\succcurlyeq_{t,\omega}$ at the preceding time-t node. As noted above, this enables the elicitation of conditional preferences at all relevant nodes—even those that, in the CPS representation, will receive zero probability.

The main result follows.

Theorem 1 *If Axioms 1–10 hold, then there exists a non-constant, affine function $u : X \to \mathbb{R}$ and a CPS $\mu \in pr(\Sigma, \mathscr{F})$ such that, for every pair of acts $f, g \in F$, and every node (t,ω),*

$$u \circ f \geq^{\mu}_{t,\omega} u \circ g \Longrightarrow f \succcurlyeq_{t,\omega} g \quad \text{and} \quad u \circ f >^{\mu}_{t,\omega} u \circ g \Longrightarrow f \succ_{t,\omega} g. \tag{7}$$

Furthermore, u is unique up to positive affine transformations, and μ is unique.

Conversely, fix a non-constant, affine function $u : X \to \mathbb{R}$ and a CPS $\mu \in pr(\Sigma, \mathscr{F})$, and for every node (t,ω), let $\succcurlyeq_{t,\omega}$ be the (setwise) minimal relation that satisfies Eq. (7). Then the preference system $(\succcurlyeq_{t,\omega})$ satisfies Axioms 1–10.

In either case, for any node (t,ω), $N_{t,\omega} = \{\omega^0 \in \mathscr{F}_t : \mu_{t,\omega}(\mathscr{F}_T(\omega^0)) = 0\}$, and $N_{t,\omega} \in \mathscr{N}_{t,\omega}$.

Notice the form of the first statement (sufficiency of the axioms): if a preference system satisfies the axioms, then a unique CPS μ and a cardinally unique utility function u can be identified so that, for any two acts f, g, if $u \circ f$ is sequentially (strictly) greater than $u \circ g$ at (t,ω) given μ, then f is (strictly) preferred to g at that node. This allows for the possibility that the individual's preference ranking may be richer than is implied by Eq. (7).

The necessity statement may instead be rephrased as follows. Fix a CPS μ and a non-constant Bernoulli utility function u, and define each preference relation $\succcurlyeq_{t,\omega}$ by

$$u \circ f \geq^{\mu}_{t,\omega} u \circ g \Longleftrightarrow f \succcurlyeq_{t,\omega} g \quad \text{and} \quad u \circ f >^{\mu}_{t,\omega} u \circ g \Longleftrightarrow f \succ_{t,\omega} g. \tag{8}$$

Clearly, $\succcurlyeq_{t,\omega}$ is then the minimal (by set inclusion) binary relation that satisfies Eq. (7), and Theorem 1 states that this preference satisfies the axioms described above. In general, one cannot ensure that preferences may be further extended in a manner consistent with the axioms. However, as suggested in Sec. 2, the sequential EU ordering characterized in Theorem 1 is sufficient for the purposes of identifying sequential best replies.

Finally, Theorem 1 also states that the union of all negligible events in $\mathscr{N}_{t,\omega}$ is itself negligible, and characterizes it in terms of the CPS μ.

A Preliminaries

I first establish a simple but useful property of critical nodes. **Note**: here and in the following, I will say "critical for (t,ω)" and omit "given μ" when the CPS μ is clear from the context.

Remark 1 *Fix a CPS $\mu \in pr(\Omega, \mathscr{F})$ and a node (t,ω). If a node (τ,ω^0) is critical for (t,ω) given μ, then $\mu_{\sigma,\omega^0}(\mathscr{F}_\tau(\omega^0)) = 0$ for $\sigma = t,\dots,\tau-1$.*

Conversely, if $\tau > t$, $\omega^0 \in \mathscr{F}_t(\omega)$, and $\mu_{t,\omega}(\mathscr{F}_\tau(\omega^0)) = 0$, then there is $\sigma \in \{t,\dots,\tau-1\}$ such that $(\sigma+1,\omega^0)$ is critical for (t,ω) given μ.

Thus, for $\tau > t$ and $\omega^0 \in \mathscr{F}_t(\omega)$, $\mu_{t,\omega}(\mathscr{F}_\tau(\omega^0)) = 0$ if and only if one of the nodes on the path from (t,ω) to (τ,ω^0) is critical for (t,ω): equivalently, if and only if the individual will be surprised at least once along this path.

Proof: If $\tau = t$, then there is nothing to prove; thus, assume that $\tau > t$. The statement is true by definition for $\sigma = \tau - 1$; thus, assume it is true for $\sigma + 1 \in \{t+1,\dots,\tau-1\}$. Then, by Bayesian updating, i.e. Part (2) of Def. 1, since $\mathscr{F}_\tau(\omega^0) \subset \mathscr{F}_{\sigma+1}(\omega^0)$,

$$\mu_{\sigma,\omega^0}(\mathscr{F}_\tau(\omega^0)) = \mu_{\sigma,\omega^0}(\mathscr{F}_{\sigma+1}(\omega^0)) \cdot \mu_{\sigma+1,\omega^0}(\mathscr{F}_\tau(\omega^0)) = 0.$$

For the converse, apply Part (2) of Def 1 iteratively to get,

$$0=\mu_{t,\omega}(\mathscr{F}_\tau(\omega^0))=\mu_{t,\omega^0}(\mathscr{F}_\tau(\omega^0))=\mu_{t,\omega^0}(\mathscr{F}_{t+1}(\omega^0))\cdot\mu_{t+1,\omega^0}(\mathscr{F}_\tau(\omega^0))=\ldots=\prod_{\sigma=t}^{\tau-1}\mu_{\sigma,\omega^0}(\mathscr{F}_{\sigma+1}(\omega^0));$$

it then follows that at least one of the factors $\mu_{\sigma,\omega^0}(\mathscr{F}_{\sigma+1}(\omega^0))$ in the rightmost term must be zero, so the corresponding node $(\sigma+1,\omega^0)$ is critical for (t,ω). ■

Next, I characterize the "sequentially greater than" order in Def. 2 inductively.

Proposition 1 *Fix any CPS $\mu \in pr(\Sigma,\mathscr{F})$, node (t,ω), and functions $a,b \in B(\Sigma)$. For $t = T$, $a \succeq^\mu_{T,\omega} b$ if and only if $\mathrm{E}_{T,\omega}a \geq \mathrm{E}_{T,\omega}b$. For $t<T$, $a \succeq^\mu_{t,\omega} b$ if and only if either*

(1) $\mathrm{E}_{t,\omega}a > \mathrm{E}_{t,\omega}b$*, or*

(2) $\mathrm{E}_{t,\omega}a = \mathrm{E}_{t,\omega}b$ *and, for all nodes (τ,ω^0) that are critical for (t,ω), with $\tau > t$, $a \succeq^\mu_{t+1,\omega^0} b$; moreover, if $\mathrm{E}_{\tau,\omega^0}a < \mathrm{E}_{\tau,\omega^0}b$ and $\mathrm{E}_{\sigma,\omega^0}a \leq \mathrm{E}_{\sigma,\omega^0}b$ for all $\sigma \in \{t+2,\ldots,\tau-1\}$ such that (σ,ω^0) is critical for $(t+1,\omega^0)$, then $(t+1,\omega^0)$ is critical for (t,ω).*

Proof: The equivalence is clear for $t=T$. For $t<T$, assume $a \succeq^\mu_{t,\omega} b$. As noted in the main text, this implies $\mathrm{E}_{t,\omega}a \geq \mathrm{E}_{t,\omega}b$. Case (1) in the Proposition corresponds to a strict inequality. If we have an equality instead, we must check the remaining conditions in Case (2) of the Proposition; thus, consider a critical node (τ,ω^0) for (t,ω) with $\tau>t$. Furthermore, consider a critical node (τ^*,ω^*) for $(t+1,\omega^0)$. Such a node is also critical for (t,ω); therefore, if $\mathrm{E}_{\tau^*,\omega^*}a < \mathrm{E}_{\tau^*,\omega^*}b$, there is $\sigma \in \{t,\ldots,\tau^*-1\}$ such that (σ,ω^*) is critical for (t,ω) and $\mathrm{E}_{\sigma,\omega^*}a > \mathrm{E}_{\sigma,\omega^*}b$; but since $\mathrm{E}_{t,\omega}a = \mathrm{E}_{t,\omega}b$, it must be the case that $\sigma \geq t+1$. Since (τ^*,ω^*) was arbitrary, conclude that $a \succeq^\mu_{t+1,\omega^0} b$, i.e. Case (2) holds. In particular, note that, if $\sigma=t+1$ is the largest σ that satisfies this condition, then $a \succeq^\mu_{t,\omega} b$ requires that $(t+1,\omega^0)$ be critical for (t,ω).

For the converse, fix a node (t,ω) with $t<T$, and consider a critical node (τ,ω^*) for which $\mathrm{E}_{\tau,\omega^*}a < \mathrm{E}_{\tau,\omega^*}b$. If Case (1) in the Proposition applies, then clearly $\tau>t$, and furthermore, as noted in the main text, $\mathrm{E}_{\sigma,\omega^*}a > \mathrm{E}_{\sigma,\omega^*}b$ for $\sigma = t$. If Case (2) holds, then again $\tau > t$; furthermore, since (τ,ω^*) is also critical for $(t+1,\omega^*)$, the assumption that $a \succeq^{\mu}_{t+1,\omega^*} b$ implies that there is $\sigma \in \{t+1,\dots,\tau-1\} \subset \{t,\dots,\tau-1\}$ such that (σ,ω^*) is critical for $(t+1,\omega^*)$, and $\mathrm{E}_{\sigma,\omega^*}a > \mathrm{E}_{\sigma,\omega^*}b$. If $\sigma > t+1$, then automatically (σ,ω^*) is critical for (t,ω) as well; if instead the largest σ that satisfies these conditions is $\sigma = t+1$, then Case (2) explicitly indicates that $(t+1,\omega^*)$ must be critical for (t,ω). Repeating the argument for all critical nodes for (t,ω) shows that $a \succeq^{\mu}_{t,\omega}$, as claimed. ■

The following characterization of "sequentially strictly greater than" is convenient.

Remark 2 *Fix a node* (t,ω)*, a CPS* $\mu \in pr(\Sigma,\mathscr{F})$ *and functions* $a,b \in B(\Sigma)$*. Then* $a \succ^{\mu}_{t,\omega} b$ *if and only if* $a \succeq^{\mu}_{t,\omega} b$ *and there is a node* (τ,ω^*) *that is critical for* (t,ω) *and such that* $\mathrm{E}_{\tau,\omega^*}a > \mathrm{E}_{\tau,\omega^*}b$*.*

Proof: If $a \succ^{\mu}_{t,\omega} b$, then in particular it is not the case that $b \succeq^{\mu}_{t,\omega} a$, so by definition there must be a critical node (τ,ω^*) with the properties in the claim.

Conversely, suppose that $a \succeq^{\mu}_{t,\omega} b$ and there is a critical node (τ_1,ω^*) for (t,ω) with $\mathrm{E}_{\tau_1,\omega^*}a > \mathrm{E}_{\tau_1,\omega^*}b$. Suppose by contradiction that also $b \succeq^{\mu}_{t,\omega} a$; then there is $\tau_2 \in \{t,\dots,\tau_1-1\}$ such that (τ_2,ω^*) is also critical for (t,ω) and $\mathrm{E}_{\tau_1,\omega^*}a < \mathrm{E}_{\tau_1,\omega^*}b$. But then, $a \succeq^{\mu}_{t,\omega} b$ requires that there be $\tau_3 \in \{t,\dots,\tau_2-1\}$ such that $\mathrm{E}_{\tau_3,\omega^*}a > \mathrm{E}_{\tau_3,\omega^*}b$. Inductively, there is an infinite sequence $(\tau_\ell)_{\ell\geq 1}$ such that each (τ_ℓ,ω^*) is critical for (t,ω) and, for $\ell > 1$, $\tau_\ell < \tau_{\ell-1}$ and $\mathrm{E}_{\tau_\ell,\omega^*}a > \mathrm{E}_{\tau_\ell,\omega^*}b$ (resp. $<$) if $\mathrm{E}_{\tau_{\ell-1},\omega^*}a < \mathrm{E}_{\tau_{\ell-1},\omega^*}b$ (resp. $>$). Since there are finitely many time periods between τ_1 and t, this is a contradiction. Thus, not $b \succeq^{\mu}_{t,\omega} a$, and so $a \succ^{\mu}_{t,\omega} b$ as claimed. ■

B Proof of Theorem 1

Terminology: two acts f,g **agree** on $E \in \Sigma$ if $f(\omega) = g(\omega)$ for all $\omega \in E$.

Notation: $f_{t,\omega} g$ denotes the act that agrees with f on $\mathscr{F}_t(\omega)$, and with g elsewhere. More generally, fEg is the act that agrees with f on $E \in \Sigma$ and with g elsewhere. 1

B.1 Sufficiency of the axioms

Begin by noting three consequences of Axiom 10 (Dynamic Consistency), along with the other axioms.

Remark 3 *Fix two nodes $(t,\omega),(\tau,\omega^*)$ with $\omega^* \in \mathscr{F}_t(\omega)$ and $\tau > t$; also fix $f, g \in F$.*

(1) if $f \succ_{\tau,\omega^*} g$ and $f(\omega') = g(\omega')$ for $\omega' \notin \mathscr{F}_\tau(\omega^*)$, then $f \succ_{t,\omega} g$.

(2) if $f \succcurlyeq_{t+1,\omega'} g$ for all $\omega' \in \Omega$, then $f \succcurlyeq_{t,\omega} g$; and if $f \succ_{t+1,\omega^*} g$ for some $\omega^* \in \mathscr{F}_t(\omega)$, then $f \succ_{t,\omega} g$.

(3) if $f \succcurlyeq_{\tau,\omega'} g$ for all $\omega' \in \Omega$, then $f \succcurlyeq_{t,\omega} g$; and if $f \succ_{\tau,\omega^*} g$ for some $\omega^* \in \mathscr{F}_t(\omega)$, then $f \succ_{t,\omega} g$.

Notice that (2) is equivalent to Axiom 10 in case preferences are complete.

Proof: (1) For $t = \tau - 1$, the assertion follows directly from Axiom 10 (recall that $f \succ_{t,\omega} g$ means "$f \succcurlyeq_{t,\omega} g$ and not $g \succcurlyeq_{t,\omega} f$"). Thus, assume the statement holds for some $t \le \tau - 1$. Fix $\omega^* \in \mathscr{F}_{t-1}(\omega)$. Then $f \succ_{\tau,\omega^*} g$ implies $f \succ_{t,\omega^*} g$. Applying Axiom 10 at time $t-1$ then yields $f \succ_{t-1,\omega} g$, as required.

(2) By Axiom 9, we may as well assume that f, g agree outside $\mathscr{F}_t(\omega)$. Fix $\omega^1, \ldots, \omega^n \in \mathscr{F}_t(\omega)$ such that $\omega^i \notin \mathscr{F}_{t+1}(\omega^j)$ for $i \neq j$, and $\bigcup_i \mathscr{F}_{t+1}(\omega^i) = \mathscr{F}_t(\omega)$. Let $f^0 = f$ and, for $i = 1, \ldots, n$, $f^i = g_{t+1,\omega^i} f^{i-1}$. Note that $f^n = g$. Furthermore, for $i = 1, \ldots, n$, $f^{i-1} \sim_{t+1,\omega^i} f \succcurlyeq_{t+1,\omega^i} g \sim_{t+1,\omega^i} f^i$, where the indifferences follow from Axiom 9 and the weak preference holds by assumption. Since f^i and f^{i-1} only differ on $\mathscr{F}_{t+1}(\omega^i)$, Axiom 10 implies that $f^{i-1} \succcurlyeq_{t,\omega} f^i$. Hence, by Transitivity, $f \succcurlyeq_{t,w} g$. Furthermore, if one of the time-$(t+1)$ preferences is strict, so is the time-t preference.

(3) Combine (1) and (2). If $\tau = t+1$, then (3) is just (2). By induction, suppose the statement is true for $t \leq \tau - 1$. Choose $\omega^1, \ldots, \omega^n \in \mathscr{F}_{t-1}(\omega)$ as in (2). The induction hypothesis implies that $f \succcurlyeq_{t,\omega^i} g$ for all $i = 1, \ldots, n$, and again, we may as well assume that f, g agree outside $\mathscr{F}_{t-1}(\omega)$. Then (2) applies and yields $f \succcurlyeq_{t-1,\omega} g$, as required. Furthermore, if $f \succ_{\tau,\omega^*} g$ for some $\omega^* \in \mathscr{F}_{t-1}(\omega)$, then, letting i be such that $\omega^* \in \mathscr{F}_t(\omega^i)$, $f \succ_{t,\omega^i} g$, so again (2) yields $f \succ_{t-1,\omega} g$.
■

Bernoulli utility. At any (t,ω), the standard assumptions hold on X, so we obtain a non-constant Bernoulli utility $u_{t,\omega}$ that represents preferences over X at that node. Now suppose that $x \succcurlyeq_{0,\omega} y$ at some $(0,\omega)$, but $x \prec_{1,\omega} y$: then, by Reflexivity and Axiom 10 (Dynamic Consistency), $x_{1,\omega} y \prec_{0,\omega} y$, which contradicts Axiom 3 (Monotonicity) as applied to $\succcurlyeq_{0,\omega}$. Thus, $x \succcurlyeq_{0,\omega} y$ implies $x \succcurlyeq_{1,\omega} y$. By Corollary B.3 in Ghirardato, Maccheroni, and Marinacci (2004), $u_{0,\omega}$ coincides with $u_{1,\omega}$ up to a positive affine transformation; since $u_{0,\omega}$ is independent of ω, denote it by u: one can then take $u_{1,\omega} = u$ for all $\omega \in \Omega$. Repeating the argument, take $u_{t,\omega} = u$ for all nodes (t,ω).

Henceforth, unless otherwise noted, I focus on preferences at a node (t,ω) with $t < T$. The delicate case is of course that of a node (t,ω) with $\mathscr{N}_{t,\omega} \neq \emptyset$. Consider this case first.

Claim: the elements of $\mathscr{N}_{t,\omega}$ are nested. Fix two distinct $N, M \in \mathscr{N}_{t,\omega}$ (if $\mathscr{N}_{t,\omega} = \emptyset$, or there is a single $N_{t,\omega} \in \mathscr{N}_{t,\omega}$, there is nothing to show), and introduce the following notation: (f, g, h', h'') is the act that agrees with f on $N_{t,\omega} \setminus M$, with g on $M \setminus N$, with h' on $M \cap N$, and with h'' elsewhere. Notice that all events involved are unions of elements of $\mathscr{F}_T$. Also, some of these events may be empty.

Fix $x, y \in X$ with $u(x) > u(y)$ and an arbitrary $k \in F$. Suppose that both $N \setminus M$ and $N_{t,\omega} \setminus M$ are non-empty. Then inductively invoking Axiom 10 (cf. Part (3) of Remark 3) yields $(x,k,k,k) \succ_{t,\omega} (y,k,k,k)$; since M is negligible, it follows that also $(x,y,k,k) \succ_{t,\omega} (y,x,k,k)$, because the union of the second and third events is M, and we can modify acts arbitrarily on M without affecting preferences. But by a similar argument, one obtains $(k,x,k,k) \succ_{t,\omega} (k,y,k,k)$, and now the assumption that N is negligible yields $(y,x,k,k) \succ_{t,\omega} (x,y,k,k)$: contradiction. Notice that the events $M \cap N$ and $\Omega \setminus (M \cup N)$ may be empty without invalidating this argument. Therefore, M and N are nested, as claimed.

Note that, as a consequence, $N_{t,\omega} \in \mathscr{N}_{t,\omega}$.

<u>EU representation off $N_{t,\omega}$</u>: $u_{t,\omega;k}, \pi_{t,\omega;k}$. For every $k \in F$, consider the class $F_{t,\omega}(k)$ of acts $f \in F$ such that $f(\omega) = k(\omega)$ for all $\omega \in N_{t,\omega}$. By assumption, $N_{t,\omega} \neq \emptyset$; furthermore, consider two acts $f, g \in F$ that agree on $\mathscr{F}_t(\omega)$: then Axiom 9 (Conditional Preferences) implies that $f \sim_{t,\omega} g$, so $\mathscr{F}_t(\omega)$ is not $\succcurlyeq_{t,\omega}$-negligible. Since $\mathscr{N}_{t,\omega}$ consists of negligible events that are at the same time subsets of $\mathscr{F}_t(\omega)$ and unions of cells of $\mathscr{F}_T$, there is some $\omega^1 \in \mathscr{F}_t(\omega)$ such that $\mathscr{F}_T(\omega) \cap N_{t,\omega} = \emptyset$.

By Non-degeneracy and Monotonicity, there are $x, y \in X$ with $x \succ_{T,\omega^1} y$; then for any $k \in F$, invoking Axiom 10 (Dynamic Consistency) inductively (cf. Part (1) of Remark 3), $x_{T,\omega^1} k \succ_{t,\omega} y_{T,\omega^1} k$. Furthermore, $x_{T,\omega^1} k, y_{T,\omega^1} k \in F_{t,\omega}(k)$.

It follows that, for fixed $k \in F$, the preference $\hat{\succcurlyeq}$ on F defined by $f \hat{\succcurlyeq} g$ iff $kN_{t,\omega} f \succcurlyeq_{t,\omega} kN_{t,\omega} g$ satisfies the Anscombe-Aumann axioms: in addition to Transitivity, Monotonicity and Independence, which follow from Axioms 1, 3 and 4 respectively, Completeness and the Archimedean property follow from Axioms 7 and 8, whereas Non-Degeneracy follows by taking $f = x_{T,\omega^1} k$ and $g = y_{T,\omega^1} k$, where $x, y \in X$ and $\omega^1 \in \mathscr{F}_t(\omega)$ are as described above. Hence, we obtain a non-constant, cardinally unique utility $u_{t,\omega;k}$ and a unique measure $\pi_{t,\omega;k} \in pr(\Sigma)$ such that, for all $f, g \in F$, $f \hat{\succcurlyeq} g$ iff $kN_{t,\omega} f \succcurlyeq_{t,\omega} kN_{t,\omega} g$ iff $\mathrm{E}_{\pi_{t,\omega;k}} u_{t,\omega;k} \circ f \geq \mathrm{E}_{\pi_{t,\omega;k}} u_{t,\omega;k} \circ g$. In particular, there must exist (by Non-degeneracy) $\hat{x}, \hat{y} \in X$ with $\hat{x} \hat{\succ} \hat{y}$;[12] by construction of $\hat{\succcurlyeq}$, $\hat{x} N_{t,\omega} \hat{y} \hat{\succcurlyeq} \hat{y}$ iff $kN_{t,\omega} \hat{y} = kN_{t,\omega}(\hat{x} N_{t,\omega} \hat{y}) \succcurlyeq_{t,\omega} kN_{t,\omega} \hat{y}$, so $\hat{x} N_{t,\omega} \hat{y} \hat{\sim} \hat{y}$, and therefore $\pi_{t,\omega;k}(N_{t,\omega}) = 0$. Note also that, for $f, g \in F_{t,\omega}(k)$, $f \hat{\succcurlyeq} g$ iff $f \succcurlyeq_{t,\omega} g$; thus the map $f \mapsto \mathrm{E}_{\pi_{t,\omega;k}} u_{t,\omega;k} \circ f$ also provides a representation of $\succcurlyeq_{t,\omega}$ on $F_{t,\omega}(k)$.

[12] We are not yet able to say that the restriction of $\hat{\succcurlyeq}$ to X coincides with that of $\succcurlyeq_{t,\omega}$, but we shall do so shortly.

Furthermore, recall that, by construction, $N_{t,\omega} \subset \mathscr{F}_t(\omega)$. With $\hat{x}, \hat{y}$ as in the preceding paragraph, let f be the act that agrees with k on $N_{t,\omega}$, and equals $\hat{y}$ on $\mathscr{F}_t(\omega) \setminus N_{t,\omega}$ and with $\hat{x}$ elsewhere; then Axiom 9 implies that $f \sim_{t,\omega} kN_{t,\omega}\hat{y}$, so that $\hat{y}\mathscr{F}_t(\omega)\hat{x} \hat{\sim} \hat{y}$, so $\pi_{t,\omega;k}(\mathscr{F}_t(\omega)) = 1$.

Claim: $u_{t,\omega;k} = u$. To see this, fix x, y with $u(x) \geq u(y)$, so that in particular $x \succcurlyeq_{t,\omega} y$ for all ω'. Then Axiom 3 (Monotonicity) implies that $kN_{t,\omega}x \succcurlyeq_{t,\omega} kN_{t,\omega}y$, so $x \hat{\succcurlyeq} y$ and therefore $u_{t,\omega;k}(x) \geq u_{t,\omega;k}(y)$. Arguing as in the "Bernoulli utility" part of this proof, the claim follows.

Claim: $\pi_{t,\omega;k}$ is independent of k. That is, for $k, k' \in F$, $\pi_{t,\omega;k} = \pi_{t,\omega;k'}$. This follows by noting that Savage's Postulate P2 (Savage, 1972) holds for $\succcurlyeq_{t,\omega}$. Suppose that $kN_{t,\omega}f \succcurlyeq_{t,\omega} kN_{t,\omega}g$ for some $k \in F$. Apply Axiom 4 mixing with $k'N_{t,\omega}f$ and weights $\frac{1}{2} : \frac{1}{2}$ to obtain $k''N_{t,\omega}f \succcurlyeq_{t,\omega} k''N_{t,\omega}(\frac{1}{2}f + \frac{1}{2}g)$, where $k'' = \frac{1}{2}k + \frac{1}{2}k'$. But $k''N_{t,\omega}f = \frac{1}{2}k'N_{t,\omega}f + \frac{1}{2}kN_{t,\omega}f$ and $k''N_{t,\omega}(\frac{1}{2}f + \frac{1}{2}g) = \frac{1}{2}k'N_{t,\omega}g + \frac{1}{2}kN_{t,\omega}f$, so Axiom 4 again implies that $k'N_{t,\omega}f \succcurlyeq_{t,\omega} k'N_{t,\omega}g$. Conclude that, $kN_{t,\omega}f \succcurlyeq_{t,\omega} kN_{t,\omega}g$ if and only if $k'N_{t,\omega}f \succcurlyeq_{t,\omega} k'N_{t,\omega}g$. In particular, if $u(x) = \mathrm{E}_{\pi_{t,\omega;k}} u \circ f$ for some $x \in X$, so that $kN_{t,\omega}f \sim_{t,\omega} kN_{t,\omega}x$ (because $\pi_{t,\omega;k}(N_{t,\omega}) = 0$), it is also the case that $k'N_{t,\omega}f \sim_{t,\omega} k'N_{t,\omega}x$ and therefore $u(x) = \mathrm{E}_{\pi_{t,\omega;k'}} u \circ f$ (again because $\pi_{t,\omega;k'}(N_{t,\omega}) = 0$ as well). This implies the claim.

Henceforth, write $\mu_{t,\omega} = \pi_{t,\omega;k}$ for an arbitrary $k \in F$. Note that, for every $k \in F$, acts in $F_{t,\omega}(k)$ are ranked according to EU, with utility u and probability measure $\mu_{t,\omega}$; furthermore, $\mu_{t,\omega}(N_{t,\omega}) = 0$ and $\mu_{t,\omega}(\mathscr{F}_t(\omega)) = 1$.

Characterization of $N_{t,\omega}$ Suppose there is $\omega^* \in \mathscr{F}_t(\omega) \setminus N_{t,\omega}$ with $\mu_{t,\omega}(\mathscr{F}_T(\omega^*)) = 0$. Again, fix $x, y \in X$ with $u(x) > u(y)$. Then, since $x_{T,\omega^*}y, y \in F_{t,\omega}(y)$, again these acts are ranked according to $u, \mu_{t,\omega}$, and so $x_{T,\omega^*}y \sim_{t,\omega} y$, which again contradicts Axiom 10 (cf. Remark 3). Therefore, $\mu_{t,\omega}(\mathscr{F}_T(\omega^*)) > 0$ for all $\omega^* \in \mathscr{F}_t(\omega) \setminus N_{t,\omega}$, or

$$N_{t,\omega} = \{\omega^0 \in \mathscr{F}_t(\omega) : \mu_{t,\omega}(\mathscr{F}_T(\omega^0)) = 0\}. \tag{9}$$

Construction of the CPS. Finally, since each preference $\succcurlyeq_{T,\omega}$ at a time-T node is complete and Archimedean, and so is every preference $\succcurlyeq_{t,\omega}$ for which $\mathscr{N}_{t,\omega} = \emptyset$, in such cases standard arguments yield a corresponding measure $\mu_{t,\omega}$ that represents it together with u; by Axiom 9, again $\mu_{t,\omega}(\mathscr{F}_t(\omega))=1$.

We thus obtain a collection $(\mu_{t,\omega})$ of measures that are adapted to $\mathscr{F}$ and satisfy Conditions 0 and 1 in Def. 1. Turn now to Condition 2, Bayes' Rule; thus, fix $t<T$, $\omega\in\Omega$.

The case $t=T-1$ requires separate treatment. Fix $E\subset\mathscr{F}_T(\omega)$. If $\mu_{T-1,\omega}(\mathscr{F}_T(\omega))=0$, then also $\mu_{T-1,\omega}(E)=0$ and Bayes' Rule holds. Otherwise, fix $x,y\in X$ with $u(x)>u(y)$; indeed, wlog assume $u(x)=1$ and $u(y)=0$. Since $\succcurlyeq_{T,\omega}$ is a EU preference represented by $u,\mu_{T,\omega}$, there is $z\in X$ s.t. $xEy\sim_{T,\omega} z\sim_{T,\omega} z_{T,\omega}y$, where the second indifference follows from Axiom 9. Thus, $u(z)=\mu_{T,\omega}(E)$. Furthermore, by Axiom 10, also $xEy\sim_{T-1,\omega} z_{T,\omega}y$. Since $\mathscr{F}_T(\omega)$ has positive probability, it does not belong to $N_{T-1,\omega}$. Hence the ranking of these acts at $(T-1,\omega)$ is EU and represented by $u,\mu_{T-1,\omega}$, so $\mu_{T-1,\omega}(E)=\mu_{T-1,\omega}(\mathscr{F}_T(\omega))\mu_{T,\omega}(E)$, as required.

Now consider $t<T-1$. I first claim that

$$\mu_{t,\omega}(\mathscr{F}_{t+1}(\omega))>0 \quad\Rightarrow\quad N_{t+1,\omega}=N_{t,\omega}\cap\mathscr{F}_{t+1}(\omega). \tag{10}$$

To see this, note first that, since $\mu_{t,\omega}(\mathscr{F}_{t+1}(\omega))>0$, there is $\omega^*\in\Omega$ with $\mathscr{F}_T(\omega^*)\subset\mathscr{F}_{t+1}(\omega^*)\setminus N_{t,\omega}$. As usual, for $x,y\in X$ with $u(x)>u(y)$, $x_{T,\omega^*}y\succ_{t,\omega} y$; these acts agree on $N_{t,\omega}\cap\mathscr{F}_{t+1}(\omega)$, and by Axioms 9 and 10, $x_{T,\omega^*}y\succ_{t+1,\omega} y$ as well. Furthermore, suppose $f,g\in F$ agree on $N_{t,\omega}\cap\mathscr{F}_{t+1}(\omega)$ and satisfy $f\succ_{t+1,\omega} g$. By Axioms 9 and 10, for $k\in F$, also $f_{t+1,\omega}k\succ_{t,\omega} g_{t+1,\omega}k$. But the acts $f_{t+1,\omega}k, g_{t+1,\omega}k$ agree on $N_{t,\omega}$, and so, in particular, if f',g' agree with f,g on $\mathscr{F}_{t+1}(\omega)\setminus N_{t,\omega}$, also $f'_{t+1,\omega}k\succ_{t,\omega} g'_{t+1,\omega}k$. Axioms 9 and 10 then imply that $f'\succ_{t+1,\omega} g'$, and by Axiom 9 we may as well assume that f',g' agree with f,g outside $\mathscr{F}_{t+1}(\omega)$. Therefore, $N_{t,\omega}\cap\mathscr{F}_{t+1}(\omega)$ is negligible for $\succcurlyeq_{t+1,\omega}$.

Furthermore, suppose that $N_{t+1,\omega}\supsetneq N_{t,\omega}\cap\mathscr{F}_{t+1}(\omega)$, so by Eq. (9) there is $\omega^0\in\mathscr{F}_{t+1}(\omega)\setminus N_{t,\omega}$ with $\mu_{t+1,\omega}(\mathscr{F}_T(\omega^0))=0$. Then, with x,y as above, $x_{T,\omega^0}y\sim_{t+1,\omega} y$, because again both acts belong to $F_{t+1,\omega}(y)$ and their ranking at $(t+1,\omega)$ is represented by $u,\mu_{t+1,\omega}$. By Axiom 10, $x_{T,\omega^0}y\sim_{t,\omega} y$ as well. But $\mu_{t,\omega}(\mathscr{F}_T(\omega^0))>0$ because $\omega^0\notin N_{t,\omega}$, and $x_{T,\omega^0}y, y\in F_{t,\omega}(y)$ as well; thus, since the ranking of these acts at (t,ω) is represented by $u,\mu_{t,\omega}$, we obtain a contradiction.

Therefore, Eq. (10) holds.

Now take $E \subset \mathscr{F}_{t+1}(\omega)$. If $\mu_{t,\omega}(\mathscr{F}_{t+1}(\omega)) = 0$, then $\mu_{t,\omega}(E) = 0$ as well, so Bayes' Rule holds. Otherwise, suppose first that $E \subset \mathscr{F}_{t+1}(\omega) \setminus N_{t+1,\omega}$; then, for x, y as above, and $z \in X$ such that $u(z) = \mu_{t+1,\omega}(E)u(x) + [1 - \mu_{t+1,\omega}(E)]u(y)$, we have $xEy, z[\mathscr{F}_{t+1}(\omega) \setminus N_{t+1,\omega}]y \in F_{t+1,\omega}(N_{t+1}, y)$, so their ranking is represented by $u, \mu_{t+1,\omega}$ and therefore $xEy \sim_{t+1,\omega} yN_{t+1,\omega}z$, because $\mu_{t+1,\omega}(N_{t+1,\omega}) = 0$. By Axiom 10, $xEy \sim_{t,\omega} z[\mathscr{F}_{t+1}(\omega) \setminus N_{t+1,\omega}]y$, and since, as was just shown, $N_{t,\omega} = N_{t+1,\omega} \cup [N_{t,\omega} \cap \mathscr{F}_{t+1}(\omega)^c]$, both acts are in $F_{t,\omega}(y)$. Therefore, their ranking is represented by $u, \mu_{t,\omega}$, and so $\mu_{t,\omega}(E)u(x) + [1 - \mu_{t,\omega}(E)]u(y) = \mu_{t,\omega}(\mathscr{F}_{t+1}(\omega) \setminus N_{t+1,\omega})u(z) + [1 - \mu_{t,\omega}(\mathscr{F}_{t+1}(\omega) \setminus N_{t+1,\omega})u(y)$. We can w.l.o.g. assume that $u(x) = 1$ and $u(y) = 0$, in which case $u(z) = \mu_{t+1,\omega}(E)$ and $\mu_{t,\omega}(E) = \mu_{t,\omega}(\mathscr{F}_{t+1}(\omega) \setminus N_{t+1,\omega})\mu_{t+1,\omega}(E)$; but since $N_{t+1,\omega} \subset N_{t,\omega}$, it follows that $\mu_{t,\omega}(N_{t+1,\omega}) \le \mu_{t,\omega}(N_{t,\omega}) = 0$, so we obtain $\mu_{t,\omega}(E) = \mu_{t,\omega}(\mathscr{F}_{t+1}(\omega))\mu_{t+1,\omega}(E)$, as required.

Finally, suppose that $E = F \cup N$, where $E \subset \mathscr{F}_{t+1}(\omega) \setminus N_{t+1,\omega}$ and $N \subset N_{t+1,\omega}$. Then $\mu_{t+1,\omega}(N) = 0$, so $\mu_{t+1,\omega}(E) = \mu_{t+1,\omega}(F)$. Furthermore, since $N_{t+1,\omega} \subset N_{t,\omega}$, also $N \subset N_{t,\omega}$ and so $\mu_{t,\omega}(N) = 0$, hence $\mu_{t,\omega}(E) = \mu_{t,\omega}(F)$. But then, by the argument in the paragraph above, $\mu_{t,\omega}(E) = \mu_{t,\omega}(F) = \mu_{t,\omega}(\mathscr{F}_{t+1}(\omega))\mu_{t+1,\omega}(F) = \mu_{t,\omega}(\mathscr{F}_{t+1}(\omega))\mu_{t+1,\omega}(E)$, as required. Thus, Bayes' Rule holds.

<u>Representation: weak preferences</u> For $t = T$, the representation in Eq. (7) holds because each $\succcurlyeq_{T,\omega}$ is an EU preference represented by $u, \mu_{T,\omega}$, and $\ge^{\mu}_{T,\omega}$ is simply the usual ordering $\ge$. Argue by induction and assume that the representation holds for $t + 1 \le T$, and fix a node (t, ω) and two acts $f, g \in F$. Consider a weak ordering first: that is, suppose that $u \circ f \ge^{\mu}_{t,\omega} u \circ g$. If $\mathscr{N}_{t,\omega} = \emptyset$, preferences at (t, ω) are EU and represented by $\mu_{t,\omega}, u$; since $u \circ f \ge^{\mu}_{t,\omega} u \circ g$ implies $\mathrm{E}_{t,\omega} u \circ f \ge \mathrm{E}_{t,\omega} u \circ g$, it follows that $f \succcurlyeq_{t,\omega} g$. Thus, assume now that $N_{t,\omega} \ne \emptyset$.

If $\mathrm{E}_{t,\omega} u \circ f > \mathrm{E}_{t,\omega} u \circ g$, then also $\mathrm{E}_{t,\omega} u \circ (kN_{t,\omega} f) > \mathrm{E}_{t,\omega} u \circ (kN_{t,\omega} g)$ for any $k \in F$, because $\mu_{t,\omega}(N_{t,\omega}) = 0$. Since u and $\mu_{t,\omega}$ represent preferences over $F_{t,\omega}(k)$, $kN_{t,\omega} f \succ_{t,\omega} kN_{t,\omega} g$. Finally, since $N_{t,\omega}$ is negligible, $f \succ_{t,\omega} g$, as required.

If instead $\mathrm{E}_{t,\omega} u \circ f = \mathrm{E}_{t,\omega} u \circ g$, then let $x \in X$ be such that $u(x) = \mathrm{E}_{t,\omega} u \circ f$. Then $f = fN_{t,\omega} f \sim_{t,\omega} fN_{t,\omega} x$, because $fN_{t,\omega} f, fN_{t,\omega} x \in F_{t,\omega}(f)$, so the ranking of f and $fN_{t,\omega} x$ is represented by $u,\mu_{t,\omega}$, and $\mathrm{E}_{t,\omega} u \circ f = \mathrm{E}_{t,\omega} f N_{t,\omega} x$. Similarly, $g \sim_{t,\omega} g N_{t,\omega} x$. Hence, by Transitivity, to show that $f \succcurlyeq_{t,\omega} g$ it is enough to prove that $f N_{t,\omega} x \succcurlyeq_{t,\omega} g N_{t,\omega} x$. Thus, fix $\omega^* \in \mathscr{F}_t(\omega)$. It will be argued that $f N_{t,\omega} x \succcurlyeq_{t+1,\omega^*} g N_{t,\omega} x$ for each such ω^* (for all other ω^*, the same ranking holds by Reflexivity); together with Remark 3, this implies the result.

If $\mathscr{F}_{t+1}(\omega^*) \cap N_{t,\omega} = \emptyset$, then Reflexivity implies that $f N_{t,\omega} x \sim_{t+1,\omega^*} g N_{t,\omega} x$. Thus, suppose that $\mathscr{F}_{t+1}(\omega^*) \cap N_{t,\omega} \neq \emptyset$. If in particular $\mathscr{F}_{t+1}(\omega^*) \subset N_{t,\omega}$, then $\mu_{t,\omega}(\mathscr{F}_{t+1}(\omega^*)) = 0$, so $(t+1,\omega^*)$ is critical for (t,ω). Since $u \circ f \geq^{\mu}_{t,\omega} u \circ g$ and $\mathrm{E}_{t,\omega} u \circ f = \mathrm{E}_{t,\omega} u \circ g$, by Proposition 1 $u \circ f \geq^{\mu}_{t+1,\omega^*} u \circ g$. Since $\mathscr{F}_{t+1}(\omega^*) \subset N_{t,\omega}$, and hence $\mathscr{F}_\tau(\omega') \subset N_{t,\omega}$ for all $\tau \geq t$ and $\omega' \in \mathscr{F}_{t+1}(\omega^*)$, the acts f and $f N_{t,\omega} x$, as well as g and $g N_{t,\omega} x$ agree at all nodes following $(t+1,\omega^*)$, so $u \circ f \geq^{\mu}_{t+1,\omega^*} u \circ g$ implies $u \circ (f N_{t,\omega} x) \geq^{\mu}_{t+1,\omega^*} u \circ (g N_{t,\omega} x)$. By the induction hypothesis, $f N_{t,\omega} x \succcurlyeq_{t+1,\omega^*} g N_{t,\omega} x$, as required.[13]

We are left with the case $\mathscr{F}_{t+1}(\omega^*) \cap N_{t,\omega} \neq \{\emptyset, \mathscr{F}_{t+1}(\omega^*)\}$. Note that then $\mu_{t,\omega}(\mathscr{F}_{t+1}(\omega^*)) > 0$. Therefore, by Eq. (9), $N_{t+1,\omega^*} = N_{t,\omega^*} \cap \mathscr{F}_{t+1}(\omega^*) = N_{t,\omega} \cap \mathscr{F}_{t+1}(\omega^*)$.

To show that $f N_{t,\omega} x \succcurlyeq_{t+1,\omega^*} g N_{t,\omega} x$, I again argue that $u \circ (f N_{t,\omega} x) \geq^{\mu}_{t+1,\omega^*} u \circ (g N_{t,\omega} x)$ and invoke the inductive hypothesis. Thus, suppose that there is a node (τ,ω^0) that is critical for $(t+1,\omega^*)$ and such that $\mathrm{E}_{\tau,\omega^0} u \circ (f N_{t,\omega} x) < \mathrm{E}_{\tau,\omega^0} u \circ (g N_{t,\omega} x)$. Since $\mu_{t+1,\omega^*}(N_{t,\omega}) = \mu_{t+1,\omega^*}(N_{t,\omega} \cap \mathscr{F}_{t+1}(\omega^*)) = \mu_{t+1,\omega^*}(N_{t+1,\omega^*}) = 0$ and $\omega^0 \in \mathscr{F}_{t+1}(\omega^*)$, $\mathrm{E}_{t+1,\omega^0} u \circ (f N_{t,\omega} x) = \mathrm{E}_{t+1,\omega^0} u \circ (g N_{t,\omega} x)$; thus, $\tau > t+1$, so node (τ,ω^0) is also critical for (t,ω); thus, by Remark 1, $\mu_{t,\omega}(\mathscr{F}_\tau(\omega^0)) = 0$. By Eq. (9), $\mathscr{F}_\tau(\omega^0) \subset N_{t,\omega}$: therefore, $\mathrm{E}_{\tau,\omega^0} u \circ f = \mathrm{E}_{\tau,\omega^0} u \circ (f N_{t,\omega} x) < \mathrm{E}_{\tau,\omega^0} u \circ (g N_{t,\omega} x) = \mathrm{E}_{\tau,\omega^0} u \circ g$.

But then, the assumption that $u \circ f \geq_{t,\omega} u \circ g$ implies that there is $\sigma \in \{t,\ldots,\tau-1\}$ such that $\mathrm{E}_{\sigma,\omega^0} u \circ f > \mathrm{E}_{\sigma,\omega^0} u \circ g$ and (σ,ω^0) is critical for (t,ω); since $\mathrm{E}_{t,\omega} u \circ f = \mathrm{E}_{t,\omega} u \circ g$, it must actually be the case that $\sigma \geq t+1$. By the definition of critical node, (σ,ω^0) is then critical for $(t+1,\omega^*)$ as well, either because $\sigma = t+1$, or because $\sigma > t+1$ and $\mu_{\sigma-1,\omega^0}(\mathscr{F}_\sigma(\omega^0)) = 0$.

[13] From $f \geq_{t+1,\omega^*} g$ we could also invoke Axiom 9 and reach the same conclusion. However, the statement that $u \circ (f N_{t,\omega} x) \geq^{\mu}_{t+1,\omega^*} u \circ (g N_{t,\omega} x)$ will be used in the argument for strict preferences below.

Moreover, since (σ,ω^0) is critical for (t,ω) and $\sigma > t$, by Remark 1 $\mu_{t,\omega}(\mathscr{F}_\sigma(\omega^0))=0$, so that $\mathscr{F}_\sigma(\omega^0)\subset N_{t,\omega}$. Therefore, $\mathrm{E}_{\sigma,\omega^0} u\circ(fN_{t,\omega}x)=\mathrm{E}_{\sigma,\omega^0} u\circ f > \mathrm{E}_{\sigma,\omega^0} u\circ g=\mathrm{E}_{\sigma,\omega^0} u\circ(gN_{t,\omega}x)$. Since (τ,ω^0) was arbitrary, $u\circ(fN_{t,\omega}x)\geq^\mu_{t+1,\omega^0} u\circ(gN_{t,\omega}x)$, so $fN_{t,\omega}x \succcurlyeq_{t+1,\omega^0} gN_{t,\omega^0}x$ again follows from the inductive hypothesis.

Thus, for all $\omega^*\in\Omega$, $fN_{t,\omega}x \succcurlyeq_{t+1,\omega^*} gN_{t,\omega}x$. Invoking Remark 3 and Transitivity then yields $f\sim_{t,\omega} fN_{t,\omega}x \succcurlyeq_{t,\omega} gN_{t,\omega}x \sim_{t,w} g$, as required. Note for future reference that the argument also showed that $u\circ(fN_{t,\omega}x)\geq^\mu_{t+1,\omega^*} u\circ(gN_{t,\omega}x)$ for all $\omega^*\in\mathscr{F}_t(\omega)$.

Representation: strict preferences. As above, the case $t=T$ is immediate, so assume that $t<T$, $u\circ f >^\mu_{t,\omega} u\circ g$, and by induction suppose that the representation result for strict preferences has been established for time $t+1$.

If $N_{t,\omega}=\emptyset$, then by Eq. (9), $\mu_{t,\omega}(\mathscr{F}_T(\omega^*))>0$ for all $\omega^*\in\mathscr{F}_t(\omega)$, so also $\mu_{t,\omega}(\mathscr{F}_\tau(\omega^*))>0$ for all $\tau=t+1,\dots,T-1$. Hence, by Remark 1, there are no critical nodes for (t,ω) except itself. In this case, $u\circ f >^\mu_{t,\omega} u\circ g$ implies $\mathrm{E}_{t,\omega} u\circ f > \mathrm{E}_{t,\omega} u\circ g$, and since $N_{t,\omega}=\emptyset$ implies that preferences at (t,ω) are EU and represented by $u,\mu_{t,\omega}$, it follows that $f\succ_{t,\omega} g$.

Thus, assume that $N_{t,\omega}\neq\emptyset$. It was shown above that, in this case, too, $\mathrm{E}_{t,\omega} u\circ f > \mathrm{E}_{t,\omega} u\circ g$ actually implies $f\succ_{t,\omega} g$. Thus, it remains to consider the case $\mathrm{E}_{t,\omega} u\circ f = \mathrm{E}_{t,\omega} u\circ g$.

Arguing as in the proof for weak preferences, let $x\in X$ satisfy $u(x)=\mathrm{E}_{t,\omega} u\circ f$. Then, as noted above, $f\sim_{t,\omega} fN_{t,\omega}x$ and $g\sim_{t,\omega} gN_{t,\omega}x$. Furthermore, it was shown that, for all $\omega^*\in\mathscr{F}_t(\omega)$, $fN_{t,\omega}x\succcurlyeq_{t+1,\omega^*} gN_{t,\omega}x$. I now claim that this preference is strict for at least one such ω^*.

Since $u\circ f >^\mu_{t,\omega} u\circ g$, by Remark 2 there is a critical node (τ,ω^{**}) for (t,ω) with $\mathrm{E}_{\tau,\omega^{**}} u\circ f > \mathrm{E}_{\tau,\omega^{**}} u\circ g$. By assumption, $\mathrm{E}_{t,\omega} u\circ f = \mathrm{E}_{t,\omega} u\circ g$, so $\tau\geq t+1$. Since (τ,ω^{**}) is critical for (t,ω), Remark 1 shows that $\mu_{t,\omega}(\mathscr{F}_\tau(\omega^{**}))=0$; but then, by Eq. (9), $\mathscr{F}_\tau(\omega^{**})\subset N_{t,\omega}$. Hence, $\mathrm{E}_{\tau,\omega^{**}} u\circ(fN_{t,\omega}x)=\mathrm{E}_{\tau,\omega^{**}} u\circ f > \mathrm{E}_{\tau,\omega^{**}} u\circ g=\mathrm{E}_{\tau,\omega^{**}} u\circ(gN_{t,\omega}x)$.

Now, it was already argued in the proof for weak preferences that $u\circ(fN_{t,\omega})\geq^\mu_{t+1,\omega^*} u\circ(gN_{t+1,\omega^*}x)$ for all $\omega^*\in\mathscr{F}_t(\omega)$, so this holds for ω^{**} as well. Furthermore, we have identified a node (τ,ω^{**}) that is critical for (t,ω) and such that $\tau\geq t+1$, where $\mathrm{E}_{\tau,\omega^{**}} u\circ(fN_{t,\omega}x\mathrm{E}_{\tau,\omega^{**}} u\circ$

$(gN_{t,\omega}x)$. The node (τ,ω^{**}) is also critical for $(t+1,\omega^{**})$, so Remark 2 implies that $u\circ(fN_{t,\omega}x)>_{t+1,\omega^{**}} u\circ(gN_{t,\omega}x)$. Then, by the inductive hypothesis, $fN_{t,\omega}x \succ_{t+1,\omega^{**}} gN_{t,\omega}x$, as claimed.

To sum up, we have $fN_{t,\omega}x \succcurlyeq_{t+1,\omega^*} gN_{t,\omega}x$ for all $\omega^*\in\Omega$, and $fN_{t,\omega}x \succ_{t+1,\omega^{**}} gN_{t,\omega}x$ for at least one $\omega^{**}\in\mathscr{F}_t(\omega)$. By Remark 3, $fN_{t,\omega}x \succ_{t,\omega} gN_{t,\omega}x$, and so $f\succ_{t,\omega} g$, as required.

B.2 Necessity of the axioms

Assume that preferences $(\succcurlyeq_{t,\omega})$ are defined via Eq. (7).

Axioms 1–6. Recall that the lexicographic ordering on vectors in $\mathbb{R}^n$ is defined as follows: for $r=(r_1,\dots,r_n), s=(s_1,\dots,s_n)$, r is lexicographically greater than s iff, for all $i\in\{1,\dots,n\}$, $r_i<s_i$ implies that there is $j<i$ with $r_i>s_i$.

This order is complete and transitive. It is also monotonic: $r_i\ge s_i$ for all $i=1,\dots,n$ implies $(r_1,\dots,r_n)\ge_L(s_1,\dots,s_n)$. Finally, it is positively homogeneous and affine, in the sense that, for $r,s,t\in\mathbb{R}^n$, $r\ge_L s$ if and only if $\alpha r+\beta t\ge_L \alpha s+\beta t$, for all $\alpha,\beta\in\mathbb{R}^n$ with $\alpha>0$.

Now fix a CPS μ. The ordering $\ge^{\mu}_{t,\omega}$ on $B(\Sigma)$ can be described as a product of lexicographic orderings: $a\ge^{\mu}_{t,\omega} b$ iff, for every $\omega^*\in\mathscr{F}_t(\omega)$, $n\ge1$, and $\tau_1,\dots,\tau_n\in\{t,\dots,T\}$ such that

(i) $t=\tau_1<\tau_2\ldots<\tau_n$,

(ii) each (τ_i,ω^*) is critical for (t,ω), and

(iii) no (σ,ω^*) with $\sigma\notin\{\tau_1,\dots,\tau_n\}$ is critical for (t,ω),

the vector $\big(\mathrm{E}_{\tau_i,\omega^*}a\big)_{i=1,\dots,n}$ is lexicographically greater than the vector $\big(\mathrm{E}_{\tau_i,\omega^*}b\big)_{i=1,\dots,n}$.

Note that the collection of indices $(\tau_1,\dots,\tau_n)$ that satisfy (i)–(iii) is entirely determined by the CPS μ, not the specific functions a and b being compared. Therefore, $\ge^{\mu}_{t,\omega}$ is the product of the lexicographic orders of conditional expectations corresponding to each such collection.

It is then immediate to verify that $\succcurlyeq_{t,\omega}$ satisfies Axioms 1, 3 and 4. Since u is non-constant, $\succcurlyeq_{t,\omega}$ also satisfies 2; and since u is affine on X, Axioms 5 and 6 hold as well.

The set $N_{t,\omega}$. I now claim that, as per Theorem 1, for every node (t,ω), either $\mathscr{N}_{t,\omega}=\emptyset$, so $N_{t,\omega}=\emptyset$ as well, or $N_{t,\omega}=\{\omega^0\in\mathscr{F}_t(\omega):\mu_{t,\omega}(\mathscr{F}_{t+1}(\omega^0))=0\}\equiv N^{\mu}_{t,\omega}$; furthermore, $N_{t,\omega}\in\mathscr{N}_{t,\omega}$.

To see this, I first show that, if $N \in \mathscr{N}$, then $\mu_{t,\omega}(N) = 0$. Thus, fix $N \in \mathscr{N}$ and choose $x, y \in X$ with $u(x) > u(y)$; also fix $k \in F$. Def. 3 implies that $N \subsetneq \mathscr{F}_t(\omega)$, so there is $\omega^* \in \mathscr{F}_t(\omega)$ such that $\mathscr{F}_T(\omega) \subset \mathscr{F}_t(\omega) \setminus N$. Then $u \circ (x_{T,\omega^*} y) >^{\mu}_{t,\omega} u \circ y$: applying Def. 2, there is no node (τ, ω^0), critical or not, where $\mathrm{E}_{\tau,\omega^0} u \circ (x_{T,\omega^*} y) < u(y)$, so $u \circ (x_{T,\omega^*} y) \geq^{\mu}_{t,\omega} u(y)$; furthermore, if $\mathrm{E}_{\tau,\omega^0} u \circ (x_{T,\omega^*} y) > u(y)$ at some critical node (τ, ω^0), there can be no $\sigma \in \{t, \ldots, \tau - 1\}$ with $\mathrm{E}_{\sigma,\omega^0} u \circ (x_{T,\omega^*} y) < u(y)$, so that it is not the case that $u(y) \geq^{\mu}_{t,\omega} u \circ (x_{T,\omega^*} y)$. Thus, $x_{T,\omega^*} y \succ_{t,\omega} y$. But since N is negligible, for any $x', y' \in X$, it must also be the case that $f' \equiv x' N (x_{T,\omega^*} y) \succ_{t,\omega} y' N y \equiv g'$. But note that

$$\mathrm{E}_{t,\omega} u \circ f' = u(x')\mu_{t,\omega}(N) + u(x)\mu_{t,\omega}(\mathscr{F}_T(\omega^*)) + u(y)[1 - \mu_{t,\omega}(N) - \mu_{t,\omega}(\mathscr{F}_T(\omega^*))],$$

whereas

$$\mathrm{E}_{t,\omega} u \circ g' = u(y')\mu_{t,\omega}(N) + u(y)[1 - \mu_{t,\omega}(N)].$$

It is then clear that, if $\mu_{t,\omega}(N) > 0$, by choosing $u(x) - u(y)$ positive but small, and $u(y') - u(x)$ large, one can obtain $\mathrm{E}_{t,\omega} u \circ f' < \mathrm{E}_{t,\omega} u \circ g'$. This readily implies $u \circ f' <^{\mu}_{t,\omega} u \circ g'$, and so $f' \prec_{t,\omega} g'$: contradiction. Thus, $\mu_{t,\omega}(N) = 0$ for every $N \in \mathscr{N}_{t,\omega}$.

It follows that $N \subset N^{\mu}_{t,\omega}$ for all $N \in \mathscr{N}_{t,\omega}$. Hence, in particular, if $N^{\mu}_{t,\omega} = \emptyset$, then $\mathscr{N}_{t,\omega} = \emptyset$. Thus, assume that $N^{\mu}_{t,\omega} \neq \emptyset$: it will now be shown that $N^{\mu}_{t,\omega} \in \mathscr{N}_{t,\omega}$, so that $N_{t,\omega} = N^{\mu}_{t,\omega}$, as claimed.

Note first that $N^{\mu}_{t,\omega}$ is a union of cells in $\mathscr{F}_T$, and by assumption it is non-empty. Furthermore, it clearly cannot coincide with $\mathscr{F}_t(\omega)$; hence, there is $\omega^* \in \mathscr{F}_t(\omega) \setminus N^{\mu}_{t,\omega}$ such that $\mu_{t,\omega}(\mathscr{F}_T(\omega^*)) > 0$. Then, with x, y as above, $\mathrm{E}_{t,\omega} u \circ (x_{T,\omega^*} y) > u(y)$, so $u \circ (x_{T,\omega^*} y) >^{\mu}_{t,\omega} u(y)$ and hence $x_{T,\omega^*} y \succ_{t,\omega} y$. Next, suppose that $f, g \in F$ agree on $N^{\mu}_{t,\omega}$ and are such that $f \succ_{t,\omega} g$, which by assumption means that $u \circ f >^{\mu}_{t,\omega} u \circ g$. By Remark 1, if (τ, ω^0) is critical for (t, ω) and $\tau > t$, then $\mu_{t,\omega}(\mathscr{F}_\tau(\omega^0)) = 0$, so that $\mathscr{F}_\tau(\omega^0) \subset N^{\mu}_{t,\omega}$. Since f, g agree on $N^{\mu}_{t,\omega}$, $\mathrm{E}_{\tau,\omega^0} u \circ f = \mathrm{E}_{\tau,\omega^0} u \circ g$; hence, $u \circ f >^{\mu}_{t,\omega} u \circ g$ implies that $\mathrm{E}_{t,\omega} u \circ f > \mathrm{E}_{t,\omega} u \circ g$. But then, if f', g' agree with f, g on $\Omega \setminus N^{\mu}_{t,\omega}$, since $\mu_{t,\omega}(N^{\mu}_{t,\omega}) = 0$, also $\mathrm{E}_{t,\omega} u \circ f' > \mathrm{E}_{t,\omega} u \circ g'$, so $u \circ f' >^{\mu}_{t,\omega} u \circ g'$ and hence finally $f' \succ_{t,\omega} g'$. This shows that $N^{\mu}_{t,\omega}$ is negligible, so $N^{\mu}_{t,\omega} \in \mathscr{N}_{t,\omega}$ and therefore $N_{t,\omega} = N^{\mu}_{t,\omega}$.

Axioms 7 and 8. If there is no negligible event for $\succeq_{t,\omega}$, the only critical node for (t, ω) is (t, ω) itself (up to the specification of $\omega' \in \mathscr{F}_t(\omega)$ as usual), so $\succeq_{t,\omega}$ is represented by $u, \mu_{t,\omega}$, and hence it satisfies Axioms 7 and 8. If instead $N_{t,\omega} \neq \emptyset$, then, as just noted, since any critical node (τ, ω^0) for (t, ω) with $\tau > t$ is a subset of $N_{t,\omega}$, any two acts f, g agree on $N_{t,\omega}$ satisfy $\mathrm{E}_{\tau,\omega^0} u \circ f = \mathrm{E}_{\tau,\omega^0} u \circ g$. Therefore, $u \circ f \geq^{\mu}_{t,\omega} u \circ g$ holds if and only if $\mathrm{E}_{t,\omega} u \circ f \geq \mathrm{E}_{t,\omega} u \circ g$: if $\mathrm{E}_{t,\omega} u \circ f > (<) \mathrm{E}_{t,\omega} u \circ g$, then $u \circ f >^{\mu}_{t,\omega} (<^{\mu}_{t,\omega}) u \circ g$, whereas if $\mathrm{E}_{t,\omega} u \circ f = \mathrm{E}_{t,\omega} u \circ g$, then there is no critical node at which a strict preference holds so that both $u \circ f \geq^{\mu}_{t,\omega} u \circ g$ and $u \circ f \leq^{\mu}_{t,\omega} u \circ g$.

Hence, $\succcurlyeq_{t,\omega}$ is represented by $u, \mu_{t,\omega}$ when comparing acts that agree on $N_{t,\omega}$, so that Axioms 7 and 8 hold.

<u>Axioms 9 and 10</u>. To conclude, Axiom 9 holds because, in the definition of $\geq^{\mu}_{t,\omega}$, only expectations conditional on $\mathscr{F}_t(\omega)$ and its subsets are considered.

As for Axiom 10, fix a node (t, ω) and acts $f, g \in F$ that agree on $\Omega \setminus F_{t+1}(\omega)$. Assume that $u \circ f \geq^{\mu}_{t,\omega} u \circ g$. Consider a critical node (τ, ω^0) for $(t+1, \omega)$ with $\mathrm{E}_{\tau,\omega^0} u \circ f < \mathrm{E}_{\tau,\omega^0} u \circ g$.

If $\mu_{t,\omega}(\mathscr{F}_{t+1}(\omega)) > 0$, then it must be the case that $\tau > t+1$, because f, g agree outside of $\mathscr{F}_{t+1}(\omega)$ and so the preceding inequality would imply that also $\mathrm{E}_{t,\omega} u \circ f < \mathrm{E}_{t,\omega} u \circ g$ and so $u \circ f <^{\mu}_{t,\omega} u \circ g$, a contradiction. But then (τ, ω^0) is also critical for (t, ω), so there must be (σ, ω^0) critical for (t, ω) with $\sigma \in \{t, \ldots, \tau - 1\}$ and $\mathrm{E}_{\sigma,\omega^0} u \circ f > \mathrm{E}_{\sigma,\omega^0} u \circ g$. If $\sigma \geq t+1$, then (σ, ω^0) is critical for $(t+1, \omega)$ as well; if instead $\sigma = t$, then since $\mu_{t,\omega}(\mathscr{F}_{t+1}(\omega)) > 0$, Bayesian updating and the assumption that f, g agree outside $\mathscr{F}_{t+1}(\omega)$ imply that also $\mathrm{E}_{t+1,\omega^0} u \circ f > \mathrm{E}_{t+1,\omega^0} u \circ g$, and of course $(t+1, \omega^0)$ is critical for $(t+1, \omega)$.

If instead $\mu_{t,\omega}(\mathscr{F}_{t+1}(\omega)) = 0$, then $(t+1, \omega^0)$ is critical for (t, ω). Then the assumed inequality for $\tau = t+1$ contradicts the assumption that $u \circ f \geq^{\mu}_{t,\omega} u \circ g$, because the fact that $\mathscr{F}_{t+1}(\omega) = \mathscr{F}_{t+1}(\omega^0)$ has zero probability also implies that $\mathrm{E}_{t,\omega} u \circ f = \mathrm{E}_{t,\omega} u \circ g$, as f and g agree outside $\mathscr{F}_{t+1}(\omega)$. Hence, $\tau > t+1$, so again (τ, ω^0) is critical for (t, ω) and there is $\sigma \in \{t, \ldots, \tau - 1\}$ with (σ, ω^0) critical for (t, ω) and such that $\mathrm{E}_{\sigma,\omega^0} u \circ f > \mathrm{E}_{\sigma,\omega^0} u \circ g$. Since $\mu_{t,\omega}(\mathscr{F}_{t+1}(\omega^0)) = 0$ and f and g agree outside $\mathscr{F}_{t+1}(\omega)$, it must be the case that $\sigma \geq t+1$, so (σ, ω^0) is also critical for $(t+1, \omega)$ (possibly because $\sigma = t+1$). This shows that $u \circ f \geq^{\mu}_{t+1,\omega} u \circ g$, i.e. $f \succcurlyeq_{t+1,\omega} g$.

Conversely, suppose that $f \succcurlyeq_{t+1,\omega} g$, so $u \circ f \geq^{\mu}_{t+1,\omega} g$. Let (τ, ω^0) be critical for (t, ω) and such that $\mathrm{E}_{\tau,\omega^0} u \circ f < \mathrm{E}_{\tau,\omega^0} u \circ g$; since f, g agree outside $\mathscr{F}_{t+1}(\omega)$, $\omega^0 \in \mathscr{F}_{t+1}(\omega)$.

Suppose that $\tau = t$: then it must be the case that $\mu_{t,\omega}(\mathscr{F}_{t+1}(\omega)) > 0$, because f, g agree outside $\mathscr{F}_{t+1}(\omega)$, and so a strict inequality could not obtain otherwise. But then, by Bayesian updating, $\mathrm{E}_{t+1,\omega} u \circ f < \mathrm{E}_{t+1,\omega} u \circ g$, which contradicts the assumption that $u \circ f \geq^{\mu}_{t+1,\omega} u \circ g$. Thus, $\tau > t$.

Therefore, (τ, ω^0) is also critical for $(t+1, \omega)$. Since $u \circ f \geq^{\mu}_{t+1,\omega} u \circ g$ by assumption, there must be $\sigma \in \{t+1, \ldots, \tau-1\}$ with (σ, ω^0) critical for $(t+1, \omega^0)$ and such that $\mathrm{E}_{\sigma,\omega^0} u \circ f > \mathrm{E}_{\sigma,\omega^0} u \circ g$. If $\sigma > t+1$, then (σ, ω^0) is also critical for (t, ω). If instead $\sigma = t+1$, we must consider two cases. If $\mu_{t,\omega}(\mathscr{F}_{t+1}(\omega)) = 0$, then again $(\sigma, \omega^0) = (t+1, \omega)$ is critical for (t, ω). If instead $\mu_{t,\omega}(\mathscr{F}_{t+1}(\omega)) > 0$, then Bayesian updating implies that also $\mathrm{E}_{t,\omega} u \circ f > \mathrm{E}_{t,\omega} u \circ g$, because f, g agree outside $\mathscr{F}_{t+1}(\omega)$. Thus, in any case, one can find a node (σ', ω^0) that is critical for (t, ω) and such that $\mathrm{E}_{\sigma',\omega^0} u \circ f > \mathrm{E}_{\sigma',\omega^0} u \circ g$. Therefore, $u \circ f \geq^{\mu}_{t,\omega} u \circ g$, so $f \succcurlyeq_{t,\omega} g$.

References

Frank J. Anscombe and Robert J. Aumann. A definition of subjective probability. *Annals of Mathematical Statistics*, 34:199–205, 1963.

P. Battigalli. Strategic rationality orderings and the best rationalization principle. *Games and Economic Behavior*, 13(2):178–200, 1996a.

P. Battigalli. Strategic independence and perfect Bayesian equilibria. *Journal of Economic Theory*, 70(1):201–234, 1996b. ISSN 0022-0531.

P. Battigalli and A. Friedenberg. Forward induction reasoning revisited. *Theoretical Economics*, 2010. ISSN 1555-7561.

P. Battigalli and M. Siniscalchi. Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games. *Journal of Economic Theory*, 88(1):188–230, 1999.

P. Battigalli and M. Siniscalchi. Strong belief and forward induction reasoning. *Journal of Economic Theory*, 106(2):356–391, 2002.

E. Ben-Porath. Rationality, Nash equilibrium and backwards induction in perfect-information games. *The Review of Economic Studies*, pages 23–46, 1997.

L. Blume, A. Brandenburger, and E. Dekel. Lexicographic probabilities and choice under uncertainty. *Econometrica: Journal of the Econometric Society*, 59(1):61–79, 1991.

A. Brandenburger and E. Dekel. Hierarchies of beliefs and common knowledge. *Journal of Economic Theory*, 59:189–189, 1993.

A. Brandenburger, A. Friedenberg, and H.J. Keisler. Admissibility in games. *Econometrica*, 76 (2):307, 2008.

Larry G. Epstein and Michel Le Breton. Dynamically consistent beliefs must be bayesian. *Journal of Economic Theory*, 61:1–22, 1993.

Larry G. Epstein and Martin Schneider. Recursive multiple-priors. *Journal of Economic Theory*, 113:1–31, 2003.

L.G. Epstein and T. Wang. " Beliefs about beliefs" without probabilities. *Econometrica*, 64(6): 1343–1373, 1996. ISSN 0012-9682.

P.C. Fishburn and I.H. LaValle. Subjective expected lexicographic utility: axioms and assessment. *Annals of Operations Research*, 80:183–206, 1998.

D. Fudenberg and J. Tirole. Perfect Bayesian equilibrium and sequential equilibrium* 1. *Journal of Economic Theory*, 53(2):236–260, 1991. ISSN 0022-0531.

Paolo Ghirardato. Revisiting Savage in a conditional world. *Economic Theory*, 20:83–92, March 2002.

Paolo Ghirardato, Fabio Maccheroni, and Massimo Marinacci. Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, 118:133–173, October 2004.

Itzhak Gilboa and David Schmeidler. Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics*, 18:141–153, 1989.

D.M. Kreps and R. Wilson. Sequential equilibria. *Econometrica: Journal of the Econometric Society*, 50(4):863–894, 1982.

M. Mariotti. Is Bayesian rationality compatible with strategic rationality? *The Economic Journal*, 105(432):1099–1109, 1995. ISSN 0013-0133.

R.B. Myerson. Multistage games with communication. *Econometrica*, 54(2):323–358, 1986a. ISSN 0012-9682.

R.B. Myerson. *Game theory: analysis of conflict*. Harvard Univ Pr, 1997.

Roger B. Myerson. Axiomatic foundations of bayesian decision theory. Discussion Paper 671, The Center for Mathematical Studies in Economics and Management Science, Northwestern University, January 1986b.

D.G. Pearce. Rationalizable strategic behavior and the problem of perfection. *Econometrica: Journal of the Econometric Society*, 52(4):1029–1050, 1984. ISSN 0012-9682.

P.J. Reny. Backward induction, normal form perfection and explicable equilibria. *Econometrica*, 60(3):627–649, 1992. ISSN 0012-9682.

A. Rényi. On a new axiomatic theory of probability. *Acta Mathematica Hungarica*, 6(3):285–335, 1955. ISSN 0236-5294.

Ariel Rubinstein. Comments on the interpretation of game theory. *Econometrica*, 59:909–924, 1991.

Leonard J. Savage. *The Foundations of Statistics*. Wiley, New York, 1954.

L.J. Savage. *The foundations of statistics*. Dover Pubns, 1972.

R. Stalnaker. Belief revision in games: forward and backward induction1. *Mathematical Social Sciences*, 36(1):31–56, 1998. ISSN 0165-4896.