# Dynamic Choice under Ambiguity

Marciano Siniscalchi*

June 21, 2004
PRELIMINARY AND SLIGHTLY INCOMPLETE
(and the Introduction needs more work!)

## 1 Introduction

Models of ambiguity-sensitive preferences have received considerable attention in the decision-theoretic literature; a number of promising economic applications of these models have also appeared in recent years. This paper contributes to the comparatively smaller literature on dynamic choice in the presence of ambiguity.

It is well-known, at least since Epstein and Le Breton's contribution [3], that imposing standard *dynamic consistency* restrictions on conditional preferences limits the extent to which ambiguity-sensitive behavior can be described; see also Ghirardato [5]. Thus, loosely speaking, models that combine dynamic choice and ambiguity must necessarily compromise on the degree of dynamic consistency assumed, and/or accept limitatons on the ability to represent ambiguity-sensitive preferences.

Existing contributions, most notably Epstein and Schneider [4] as well as a number of related papers coauthored by Epstein, essentially resolve this trade-off in favor of relatively strong notions of dynamic consistency. This paper, on the other hand, imposes no restrictions on attitudes towards ambiguity, and replaces standard dynamic consistency assumptions with a less restrictive *sophistication* requirement: while the decision-maker (DM) is allowed to hold different preferences over Savage-style acts at different points in time, *she is required to correctly anticipate her future*

*Economics Department, Northwestern University, Evanston, IL 60208; Email `marciano@northwestern.edu`. Web: `faculty.econ.northwestern.edu/faculty/siniscalchi`. I have greatly benefited from conversations with Peter Klibanoff. I also thank Larry Epstein, Paolo Ghirardato, and Fabio Maccheroni for extensive discussions. All errors are my own.

*choices, and incorporate them in her current decisions.* This is formalized in a setting wherein the objects of choice are decision trees rather than Savage acts.

Within this framework, this paper:

- Characterizes *backward induction* for decision trees via Sophistication and Weak Commitment axioms (Theorem 4). This is the main contribution of this paper, and applies to a variety of decision models that account for non-neutral ambiguity attitudes.

- Provides axioms on conditional preferences that make it possible to *elicit* the latter from prior preferences, and conversely shows how to *define* conditional preferences from prior preferences in such a way as to satisfy those axioms. This mirrors analogous, standard results in the expected-utility setting.

- Provides a characterization of "full Bayesian updating" for maxmin expected utility preferences, and adapts the general characterization of Backward Induction to this decision model. This yields a complete behavioral theory of dynamic choice for maxmin EU preferences that does not rely on full dynamic consistency, and hence does not restrict the DM's static preferences in any way.

## Motivating Examples

The key issues that arise in the analysis of dynamic choice under ambiguity, and the approach taken in this paper to address them, are best understood with the aid of examples. Consider the celebrated three-color-urn example by Daniel Ellsberg [1]. An urn contains 90 balls, of which 30 are red and 60 are either yellow or blue. The choices made available to the DM, as well as the modal preferences, are depicted in Fig. 1.

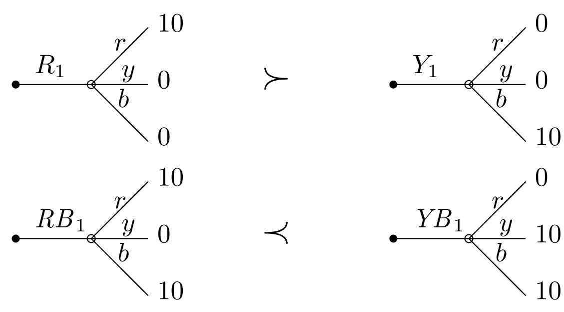


Figure 1: The static Ellsberg paradox.

The four choices $R_1$, $Y_1$, $RB_1$ and $YB_1$ are represented as (very simple) "decision trees" to introduce the graphical conventions adopted here to describe more interesting dynamic choice situations. Filled dots are *decision points*, whereas empty circles represent *chance nodes*—or, more precisely, points in the decision problem where some information is revealed to the DM. Labels at chance nodes have the obvious interpretation: $r, y, b$ are the states in the usual description of the Ellsberg three-color urn experiment. In general, actions at decision points are similarly labelled. Note that each tree has a unique action at the initial node, which is always a decision node; this convention is consistent with the formalism described in the next section, and turns out to be convenient for the purposes of composing decision trees out of other, smaller trees. Essentially, the action at the initial decision node may be thought as representing the choice of a particular decision tree; the labelling adopted in the figures presented here is consistent with this interpretation.

The preferences $R_1 \succ Y_1$ and $RB_1 \prec YB_1$ are consistent with the usual interpretation of the Ellsberg paradox: the former reflects the DM's concern that the urn might contain no yellow balls at all; the latter is informed by the consideration that, in fact, the urn may contain no blue balls. As is well-known, these rankings are consistent with the assumption that the DM's preferences admit a maxmin-expected utility (MEU) representation (Gilboa and Schmeidler [6]), with utility function $u$ such that $u(10) > u(0)$ and a set of priors $C$ that, for example, contains all probability distributions $q$ over $\{r, y, b\}$ with $q(r) = \frac{1}{3}$.

Now consider a prize $x \in \{0, 10\}$, and a corresponding, alternative decision problem, denoted $f_x$. The DM is first informed whether the ball drawn from the urn is blue, in which case she receives $x$. If the ball drawn is not blue, the she can choose whether to bet on red on yellow. The corresponding decision tree is depicted in Fig. 2. At her second decision point, the DM has two actions available, labelled $R$ and $Y$, with the natural interpretation. Observe that each of the subtrees beginning with that decision point and continuing with either $R$ or $Y$ can be viewed as a *conditional* decision tree in its own right; these subtrees (drawn in red and yellow colors respectively in Fig. 2) will be denoted by $R$ and $Y$ respectively.

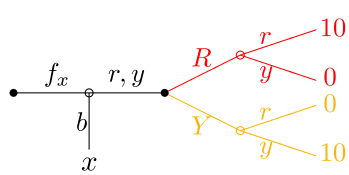


Figure 2: Deferred Choice

It is clear that, since there are two decision points in the tree of Fig. 2, in order to describe the DM's behavior it is necessary to consider two preference relations—her unconditional, or *a priori* preferences, and her preferences conditional upon learning that the ball drawn is not blue. It is natural to assume that, conditional upon the event $E = \{r, y\}$, the DM has preferences over all "conditional acts", i.e. maps from $E$ to a suitable outcome space; alternatively, conditional preferences may be defined over all maps from $\Omega = \{r, y, b\}$ to the outcome space, i.e. all Savage acts.[1] Denote such conditional preferences by $\succcurlyeq_E$.

But what should the domain of the DM's *un*conditional preferences be? It is typically assumed that such preferences are (also) defined over all Savage acts. However, there are two difficulties with this approach. To highlight them, consider the following question:

> Suppose the DM has a choice between $f_{10}$ and $RB_1$ (i.e. if she can choose to "play" either decision tree). Which one will she pick?

Clearly, if the DM's unconditional preferences are defined on a domain that does *not* include $f_{10}$, this question cannot be answered directly.

Dynamic consistency (cf. Axiom 4.1 in Section 3 for a formal statement) provides a way to answer it indirectly in a coherent fashion, but leads to difficulties in the presence of ambiguity. One implication of dynamic consistency is that

$$R \succcurlyeq_E Y \quad \text{if and only if} \quad R_1 \succcurlyeq Y_1 \quad \text{if and only if} \quad RB_1 \succcurlyeq YB_1,$$

where $\succcurlyeq$ denotes unconditional preferences. That is: since $R_1(b) = Y_1(b)$ and $RB_1(b) = YB_1(b)$, preferences between $R_1$ and $Y_1$ and, respectively, $RB_1$ and $YB_1$ should be determined by the DM's preferences conditional upon $E = \{r, y\}$. Disregarding the actual preferences in the Ellsberg example for the moment, note that, if these conditions hold, then the above question *can* easily be answered. The DM can regard $f_{10}$ as providing a deferred choice between a bet on "red or blue" and a bet on "yellow or blue"; if, for instance, she prefers the latter, Dynamic Consistency implies that she will also prefer $Y$ to $B$ conditional upon $E$; hence, if she decides to "play" the decision tree $f_{10}$, she can also *commit* to choosing $Y$ at her second decision point. Thus, from the point of view of a dynamically consistent DM who prefers $YB_1$ to $RB_1$, $f_{10}$ is strictly preferred to $RB_1$ ex ante. The problem is of course that, as noted above, Dynamic Consistency *restricts ex-ante preferences* in a manner that is inconsistent with those typically reported in the Ellsberg paradox: as reported above, it requires that $R_1 \succcurlyeq Y_1$ hold if and only if $RB_1 \succcurlyeq YB_1$.

On the other hand, if Dynamic Consistency is dropped and only preferences over Savage acts are observed, there is no way to answer the above question. Suppose that one is willing to *assume* that the DM's conditional preferences are derived from her prior preferences via some well-specified updating rule: for instance, prior-by-prior, or "full", Bayesian updating. For the set of priors $C$

[1] As is well known, the second approach can be made equivalent to the former by imposing a suitable *consequentialism* assumption. The former approach is notationally more convenient, once full-blown decision trees are taken into account.

defined above, the relevant set of posteriors is $C_E = \{q : \frac{1}{3} \leq q(r) = 1 - q(y) \leq 1\}$. It is immediate to verify that $R \succ_E Y$; notice that this constitutes a violation of Dynamic Consistency. It is now no longer clear whether the DM should prefer $f_{10}$ to $RB_1$: ex-ante, she would like to commit to choosing $Y$ at her second decision node, but ex-post she will strictly prefer $R$. *If she anticipates this*, then ex-ante the DM should be indifferent between the decision trees $f_{10}$ and $RB_1$: her behavior at the second decision node in $f_{10}$ is such that, in every state of the world, she will receive the same final prizes in both trees. But *whether or not she is able to anticipate her future choices cannot be ascertained by observing only her conditional and unconditional preferences over acts.*[2]

In the framework adopted in this paper, the DM has preferences over decision trees. The above question can thus be answered directly; more importantly, it is possible to formalize the assumption that the DM correctly anticipates her future behavior. In the setting under consideration, the requirement is that

$$\text{If } R \succ_E Y \quad \text{then} \quad f_{10} \sim RB_1 \text{ and } f_0 \sim R_1.$$

That is: the DM evaluates the trees $f_{10}$ and $f_0$ *as if* they simply did not contain continuation trees that she will surely not chose. This assumption will be deemed *Sophistication*.

It is useful to think of this as an assumption about the DM's ex-ante beliefs about the behavior of her future selves: specifically, Sophistication captures one implication of the assumption that such beliefs are *correct*. In this respect, Sophistication may be viewed as having a "game-theoretic" character, although formally it is purely decision-theoretic (and fully behavioral).

Since it requires that correct beliefs be reflected in the DM's ex-ante preferences, it seems plausible to expect that Sophistication should play a role in guaranteeing that *conditional preferences can be elicited from ex-ante preferences*. Proposition 2 shows that this is indeed the case; it also clarifies the behavioral character of the approach adopted here.

Furthermore, interpreting Sophistication as an assumption about beliefs clarifies the limits to the kind of restrictions that it can generate. Consider the decision tree in Fig. 3, denoted $f$.

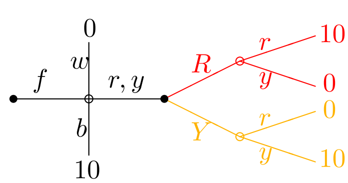


Figure 3: Sophistication and Weak Commitment

[2]Incidentally, if one is *not* willing to assume that conditional preferences are observable (or just postulate their functional form), it is no longer clear how the latter can be elicited from observed prior preferences. This is evident in the Ellsberg example: since $R_1 \succ Y_1$ but $RB_1 \prec YB_1$, it is not clear how one can conclude that, in fact, $R \succ_E Y$.

The state space comprises four states, denoted $r, y, b, w$; suppose that the DM has MEU preferences, with priors $C = \{q : q(r) = q(b),\, q(y) = q(w)\}$. Also let $E = \{r, y\}$ and suppose that her posterior preferences $\succcurlyeq_E$ are determined by full Bayesian updating; the set of posteriors is then $C_E = \{q : q(r) = 1 - q(y)\}$.

Notice that $R \sim_E Y$; moreover, observe that the DM strictly prefers the Savage act $YB_1$ that yields 10 if $y$ or $b$ obtain, and 0 otherwise, to the Savage act $RB_1$ that yields 10 if $r$ or $b$ obtain, and 0 otherwise. Hence,[3] if the trees $f_R$, $f_Y$ are obtained from $f$ by eliminating the yellow and, respectively, the red subtrees, the DM should prefer $f_Y$ to $f_R$ ex-ante. Thus, it cannot be the case that $f \sim f_Y$ and $f \sim f_R$ both hold.

Intuitively, this is consistent with the correct beliefs interpretation of Sophistication. The DM may equally well choose $R$ or $Y$ upon reaching her second node; therefore, ex-ante, if she *believes* that $Y$ will be chosen, she will actually strictly prefer $f$ to $f_R$. Thus, the correct beliefs interpretation does *not* support a restriction such as "if $R \sim_E Y$, then $f \sim f_R \sim f_Y$", and indeed, the formal Sophistication axiom does *not* require this.

On the other hand, this example illustrates that Sophistication alone may not be sufficient to pin down the DM's behavior in all decision trees: it must be complemented by some *tie-breaking assumption*. One possible approach, which will be fully developed in Section 2, is to require that the DM may be able to commit to specific future choices, if the DM's future self does not have opposite strict preferences. This assumption, referrred to as *Weak Commitment*, formalizes the notion of *consistent planning* first proposed by Strotz [16].

Together, Sophistication and Weak Commitment provide a full characterization of *backward induction* for decision trees. This is the main contribution of this paper; the details and formal statement of this result can be found in Section 3.2.2.

Assuming that preferences are defined over decision trees is much in the same spirit as Kreps's seminal contribution on menu choice (Kreps [10]) and the temporal resolution of uncertainty (Kreps [11]). A menu choice framework has been recently adopted by Epstein [2] to study "non-Bayesian" updating. Klibanoff and Ozdenoren [9] characterize a subjective version of Kreps's [11] model of recursive expected utility; they do not consider ambiguity-sensitive preferences, but their decision setting is similar to the one adopted here.

Finally, Klibanoff [8] addresses the conflict between dynamic consistency and ambiguity by allowing for non-neutral preferences for the timing of resolution of uncertainty (as well as a form of state- and history-dependence of preferences over outcomes).

[3] If the DM evaluates trees where no "actual" choices are made by reducing them to acts: cf. Assumption 4.3 in Section 3.

## 2 Decision Setting

### 2.1 Histories and Trees

The notation employed in this paper adapts the notion of "perfect-information game tree" in Osborne-Rubinstein [13]. The basic building block in the description of a decision tree is the *history*: an ordered list of the DM's actions and "Nature's actions" which describes a possible (partial or complete) unfolding of occurrences in the decision tree under consideration. Specifically, the DM's actions are labels representing *choices* available to the DM; Nature's actions represent *information* that the DM may receive. A decision tree can then be represented by a suitable collection of histories, together with an assignment of prizes to terminal (i.e. complete) histories.

Formally, fix a finite set $\Omega$ of states, a countable collection $A$ of action labels that contains a distinguished element "$*$", and a collection $X$ of prizes; assume the latter is a connected separable topological space. A **history of length** $T \geq 0$ **starting at** $E \subset \Omega$ is a sequence of ordered pairs, denoted

$$h = [(a_1, E_1), \ldots, (a_T, E_T)],$$

such that, for all $t = 1, \ldots, T$, $a_t \in A$, $E_t \subset E$, and in particular $a_1 = *$. Throughout this note, it is notationally convenient to also consider the "empty history" $h = \emptyset$. If $h$ has length $T$ and $1 \leq t \leq T$, then $h_t$ denotes the history consisting of the first $t$ elements of $h$; furthermore, $h_0 = \emptyset$ for every history $h$. Similarly, $h_{t_1:t_2}$ denotes the history $[(a_{t_1}, E_{t_1}), (a_{t_1+1}, E_{t_1+1}), \ldots, (a_{t_2}, E_{t_2})]$.

The length of history $h$ is denoted by $\lambda(h)$; also, $\lambda(\emptyset) = 0$. The action and event that appear in the last ordered pair of a history $h$ (i.e. $a_T$ and $E_T$ above) are denoted by $a(h)$ and $E(h)$ respectively; it is also convenient to let $E(\emptyset) = E$, for reasons that will be clear below. Compositions of histories are denoted in an obvious way as follows: $[h, h']$, $[h, (a, E)]$; in particular, $[\emptyset, h] = [h, \emptyset] = h$. Finally, a history $h$ is a **subhistory** of another history $h'$, written $h \leq h'$, if $h = h'_t$ for some $t \in \{0, \ldots, \lambda(h')\}$; it is a **strict subhistory** of $h'$ if $t < \lambda(h')$.

**Definition 1** A **decision tree starting at** $E$ is a tuple $f = (E, H, x)$, where $E \subset \Omega$, $H$ is a finite collection of histories starting at $E$ such that

1. if $h \in H$ and $\lambda(h) > 0$, then $h_{\lambda(h)-1} \in H$;
2. for every $h, h' \in H$ such that $\lambda(h) = \lambda(h') \equiv \bar{\lambda} > 1$: if $h_{\bar{\lambda}} = h'_{\bar{\lambda}}$, then $a(h) \neq a(h')$;
3. for every $h \in H$ and every $a \in A$ such that $[h, (a, F)] \in H$ for some $F \subset E$, the collection $\{F : [h, (a, F)] \in H\}$ is a non-trivial partition of $E(h)$;
4. for every $\omega \in \Omega$, there exists $h \in H$ such that $E(h) = \{\omega\}$;

and $x : \{h : |E(h)| = 1\} \to X$. The set of all decision trees starting at $E$ is denoted by $F_E$.

A few clarifying comments. Action labels are just that—labels. In every decision tree, the "time-1" action $a_1 = *$ is dummy: think of it as corresponding to the choice of the decision tree under consideration; Nature's move, on the other hand, is real. Also, for every decision tree, by Condition 1, $\emptyset \in H$. Condition 2 requires that all actions avaliable at a history be labelled differently. Condition 3 obviously implies that events in a history form a decreasing subsequence that, by 4, ends with a singleton.

It is convenient to refer to histories $h$ such that $E(h)$ is a singleton as **terminal histories**; all other histories are **non-terminal**. The set of terminal histories is sometimes denoted $Z$.

Special cases of decision trees: first, if $A = \{*\}$ and $H = \{[(*, \{\omega\})] : \omega \in E\}$, then $f = (E, H, x)$ corresponds to the **Savage act** conditional on $E$ that yields the outcome $x(\omega)$ in state $\omega$; in particular, if $x(\cdot)$ is constant and equal to some $\bar{x} \in X$, $f$ is a constant act, and will be denoted by $\bar{x}_E$; indeed, in an extension of the usual abuse of notation, whenever the set $E$ is clear from the context, this subscript will be dropped.

Another special case obtains when, for any non-terminal history $h$ of $f$, there is a unique $a \in A$ such that $[h, (a, E)]$ for some $E \subset \Omega$. That is, there is a unique action available at every non-terminal history; intuitively, there is simply no choice to be made at any point in the $f$. Such trees will be called **plans**. Notice that a Savage act, and in particular a constant act, is necessarily a plan, but the converse is not true.

For a more interesting example, fix a sequence $\mathcal{F}_1, \ldots, \mathcal{F}_T$ of progressively finer partitions of $\Omega$, such that $\mathcal{F}_T = \{\{\omega\} : \omega \in \Omega\}$; also, for every $F \subset \Omega$ such that $F \in \mathcal{F}_t$ for some $t$, let $A(F)$ be a collection of action labels. Then one can construct a decision history by specifying histories inductively, as follows: for 1, $[(a_1, F)] \in H$ iff $F \in \mathcal{F}_1$; then, for $t = 1, \ldots, T$, and for every $h \in H$ such that $\ell(h) = t - 1$, say that $[h, (a, F)] \in H$ iff $a \in A(E(h))$ and $F \in \mathcal{F}_t$. One feature of such a tree is that Nature's moves are independent of the DM's choices; the notation adopted here allows for greater flexibility (e.g. buying different signals about the prevailing state of the world).

The preceding example suggests the following derived notation and terminology. For a given decision tree $f = (E, H, x) \in F_E$, and for any $h \in H \backslash Z$, let $A_f(h) = \{a \in A : \exists F \subset E \text{ s.t. } [h, (a, F)] \in H\}$, the set of **actions available to the DM at** $h$ (or, more precisely, immediately after $h$); also, for every $h \in H \setminus Z$ and every $a \in A_f(h)$, let $\mathcal{F}_f(h, a) = \{F : [h, (a, F)] \in H\}$ be the **"actions" available to Nature following history** $h$ **and the DM's choice of** $a$. Thus, Condition (b) states that, for every $h = H \cup \{\emptyset\}$ and $a \in A(h)$, $\mathcal{F}_f(h, a)$ is a partition of $E(h)$.

### 2.2 Tree Surgery

First, consider $f = (E, H, x) \in F_E$, $\{f_1, \ldots, f_n\} \subset F_{E'}$ such that $f_i = (E', H_i, x_i)$ for all $i = 1, \ldots, n$, and $h \in H$ such that $E(h) = E'$. Then the **replacement tree** $\{f_1, \ldots, f_n\}_h f$ is the element of $F_E$ defined by the tuple $(E, \bar{H}, \bar{x})$, where:

- $a_1^*, \ldots, a_n^* \in A$ are $n$ distinct action labels;
- $\bar{H} = \{h' \in H : h \not< h'\} \cup \bigcup_i \{[h, (a_i^*, E(h_1')), h'_{2:\lambda(h')}] : h' \in H_i\}$;
- $\bar{x}(z) = x(z)$ if $z \in Z$ (the set of terminal histories for $f$) and $h \not< z$; $\bar{x}([h, (a_i^*, E(z_1')), z'_{2:\lambda(z')}]) = x_i(z')$ if $z' \in Z_i$ (terminal histories for $f_i$).

The key point is that histories consist of histories from $f$ that do not strictly follow $h$, plus histories that begin with $h$ and continue with some history from one of the trees $f_i$. Recall that histories $h'$ from all trees $f_1, \ldots, f_n$ all begin with the action $a_1 = *$ by definition; hence, before the history $h'$ from $f_i$ is joined with $h$, its time-1 action $a(h_1') = *$ is replaced with a suitable label $a_i^*$, ensuring that no duplicate labels are used for the trees $f_1, \ldots, f_n$. Notice also that, for $h = \emptyset$, $\{f_1, \ldots, f_n\}_h f = f$.

Second, consider $f = (E, A, H, x) \in F_E$ and $h \in H$ non-terminal. Then any $a \in A(h)$ identifies a **continuation tree**, denoted $f(h, a)$ and defined by the tuple $(E', H', x') \in F_{E'}$, where:

- $E' = E(h)$;
- $H' = \{[(*, E(h_1')), h'_{2:\lambda(h')}] : [h, h'] \in H \text{ and } a(h_1') = a\}$;
- $x'([(*, E(z_1')), z'_{2:\lambda(z')}]) = x'([h, z'])$ for all $z'$ such that $[h, z'] \in Z$.

That is: every continuation $h'$ of $h$ wherein $a$ is chosen at $h$ becomes a history in the continuation tree. However, the initial action choice in $h'$ is changed from $a$ to $*$. Again, according to this definition, for $h = \emptyset$, $f(h, a) = f$, where of course $a = *$.

## 3 Axioms and Results

Throughout this section, the main object of interes is a collection $\{\succcurlyeq_E\}_{\neq E \subset \Omega}$ of binary relations; in particular, for all non-empty $E \subset \Omega$, $\succcurlyeq_E$ is a preference relation on $F_E$. For notational simplicity, $\succcurlyeq_\Omega$ will be denoted by $\succcurlyeq$.

Subsection 3.1 shows that, under suitable axioms, conditional preferences can be either elicited or defined from unconditional ones; this result clarifies the behavioral significance of the analysis carried out in the remainder of this paper.

Subsection 3.2 formalizes the key notions of sophistication and weak commitment, and establishes the main result of this section, the promised characterization of backward induction for decision trees.

The material in Subsections 3.2 and 3.1 is essentially orthogonal; readers who are more interested in the analysis of backward induction can safely read §3.2 first and §3.1 later.

Subsection 3.3 discusses alternatives to the weak commitment assumption, and draws connections with the "multi-selves" (or game-theoretic) approach to dynamic decision making in the absence of full commitment. [TO BE WRITTEN]

### 3.1 Eliciting Conditional Preferences

In the usual setting of choice among Savage acts, conditional preferences can equivalently be seen as either primitive or derived objects. In the former case, one can postulate the existence of a system of conditional preferences $\{\succcurlyeq_E\}_{\neq E \subset \Omega}$, related to one another by appropriate *consistency* conditions (cf. e.g. Myerson [12]). In the latter, conditional preferences are defined from unconditional preferences, provided the latter satify suitable *separability* assumptions (e.g. Savage's Sure Thing Principle). Loosely speaking, given a consistent system of conditional preferences, it can be shown that unconditional preferences must be separable, and that, moreover, conditional preferences can be derived from unconditional preferences by means of a uniquely defined updating rule. Conversely, if unconditional preferences are separable and all conditional preferences are defined via a suitable updating rule, the resulting system of conditional preferences will be consistent.[4] One convenient implication of these results is that, regardless of whether or not one wishes to treat them as primitive objects, conditional preferences can be retrieved from *un*conditional ones.

This subsection establishes a similar result in the setting of choice among trees. As argued in the Introduction. in order to account for non-neutral attitudes towards ambiguity, eventwise separability assumptions must be relaxed; this implies that standard updating rules no longer yield well-defined conditional preferences. Thus, a different approach will be adopted here. The key ingredients of the approach described in this section are *conditional certainty equivalents* and a weak form of *sophistication*. It should be emphasized that, while notions of sophistication play a key role throughout this paper, certainty equivalents merely provide one convenient way to elicit preferences—but not necessarily the only one. Alternative approaches may be viable, e.g. in a setting where the state space is "rich".

[4] One exposition of these results can be found in Ghirardato [5]. For a related result in the setting of choice under risk, see Karni and Schmeidler [7]. It should also be noted that this equivalence (which arguably has the status of a "folk theorem" in decision theory) is implicit in Savage's justification of his Postulate P2 (cf. [14], pp. 21–23).

### 3.1.1 Axioms on Conditional Preferences

To aid intuition, I begin by stating axioms on the preference system $\{\succcurlyeq_E\}_{\emptyset\neq E\subset\Omega}$. First, assume that preferences over prizes (constant trees) are unaffected by conditioning.[5]

**Axiom 3.1 (Stable Tastes)** *For all $x, x' \in X$, and all non-empty $E \subset \Omega$: $x \succcurlyeq_E x'$ if and only if $x \succcurlyeq x'$.*

Next, one must ensure that conditional certainty equivalents exist. Since $X$ is assumed to be a connected and separable topological space, standard dominance (or monotonicity) and continuity requirements suffice.

**Axiom 3.2 (Conditional Dominance)** *For all non-empty $E \subset \Omega$, $f = (E, H, x) \in F_E$, and all $x', x'' \in X$: if $x' \succcurlyeq x(z) \succcurlyeq x''$ for all terminal histories $z \in H$, then $x' \succcurlyeq_E f \succcurlyeq_E x''$.*

**Axiom 3.3 (Conditional Prize-Act Continuity)** *For all non-empty $E \subset \Omega$ and all $f \in F_E$, the sets $\{x \in X : x \succcurlyeq_E f\}$ and $\{x \in X : x \preccurlyeq_E f\}$ are closed in $X$.*

**Remark 3.1** Consider a non-empty event $E \subset \Omega$ and a complete and transitive relation $\succcurlyeq_E$ on $F_E$ that satisfies Axioms 3.1, 3.2 and 3.3. Then, for all $f \in F_E$, there exists $x \in X$ such that $x \sim_E f$.

**Proof:** [Standard] Denote by $U$ and $L$ the closed sets in Axiom 3.3. Since $\Omega$ is finite, there exist $x', x''$ such that $x' \succcurlyeq x(z) \succcurlyeq x''$ for all terminal histories $z$ of $f$. By Axioms 3.1, 3.2 and transitivity, $x' \in U$ and $x'' \in L$. By completeness, $U \cup L = X$. Since $X$ is separable, the non-empty, closed sets $U$ and $L$ must have non-empty intersection; any $x \in U \cap L$ clearly satisfies $x \sim_E f$. ■

Finally, conditional and unconditional preferences over trees must be axiomatically related: this is the objective of the following weak sophistication requirement. Suppose that, upon reaching the partial history $h = [(*, E)]$, the DM faces a choice between the tree $f \in F_E$ and the prize $x$. If $x$ is strictly preferred to $f$ conditional on $E$, then, a priori, the DM should recognize that the alternative $f$ is actually irrelevant: she will not choose it if the history $h$ is reached. Similar considerations hold in case $f$ is strictly preferred to $x$ conditional on $E$. This motivates the following axiom.

**Axiom 3.4 (Weak Sophistication)** *For all $E \subset \Omega$, all $g = (\Omega, H, x) \in F_\Omega$ such that $[(*, E)] \in H$, all $f \in F_E$, and all $\bar{x} \in X$:*

*(i) if $\bar{x} \succ_E f$, then $\{f, \bar{x}\}_h g \sim \{\bar{x}\}_h g$; and*

[5] For the present purposes, it would be sufficient to impose this requirement on a suitably rich subset of prizes. For instance, if $X$ consists of consumption streams, it would be enough to restrict Axiom 3.1 to constant streams.

*(ii) if $\bar{x} \prec_E f$, then $\{f, \bar{x}\}_h g \sim \{f\}_h g$.*

Notice that the stronger sophistication axiom considered in Section 3.2 implies Axiom 3.4. On the other hand, Axiom 3.4 is not sufficient to yield backward-induction decisions, except in very simple trees: it is "just enough" to ensure that conditional preferences can be retrieved from unconditional ones. Thus, the present approach makes it possible to address the distinct issues of elicitation and sophistication in a relatively independent way.

### 3.1.2 Axioms on Unconditional Preferences

Turn now to unconditional preferences; begin with a simple continuity requirement.

**Axiom 3.5 (Prize Continuity)** *For all $\bar{x} \in X$, the sets $\{x \in X : x \succcurlyeq \bar{x}\}$ and $\{x \in X : x \preccurlyeq \bar{x}\}$ are closed in $X$.*

Next, say that an event $E \subset \Omega$ is **immediately relevant** for the act $g \in F_\Omega$ if $[(*, E)]$ is a history of $g$; say that $E$ is **$\succcurlyeq$-essential** if, whenever $h = [(*, E)]$ is a history of $g \in F_\Omega$, and $x, x' \in X$ satisfy $x \succ x'$, $\{x\}_h g \succ \{x'\}_h g$. Denote the set of acts $g \in F_\Omega$ for which the non-empty event $E \subset \Omega$ is immediately relevant by $F_\Omega(E)$; denote the set of $\succcurlyeq$-essential events by $\mathrm{E}(\succcurlyeq)$. Observe that, if $g \in F_\Omega(E)$, then clearly $M_h g \in F_\Omega(E)$ for any finite set $M \subset F_E$.

It is now possible to state the key assumption on unconditional preferences. Consider the following situation: conditional upon the ($\succcurlyeq$-essential) event $E$, the certainty equivalent of the continuation tree $f \in F_E$ is $\bar{x}_{f|E}$; notice that, if conditional preferences are not directly observable, it is not possible to directly determine $\bar{x}_{f|E}$, so a restriction on *un*conditional preferences involving it would not be verifiable (i.e. "fully behavioral"). Thus, it is necessary to work around this observability restriction.

Consider an *arbitrary* prize $\bar{x}$. Suppose first that $\bar{x} \succcurlyeq \bar{x}_{f|E}$, and consider a prize $x \succ \bar{x}$ and a tree wherein, if the history $h = [(*, E)]$ is reached, the DM faces a choice between $f$ and $x$. *If* her ranking of prizes is not altered by conditioning, then it is also the case that $x \succ_E x \succeq_E \bar{x}_{f|E} \sim_E f$; thus, $x \succ_E f$. Furthermore, *if* the DM correctly anticipates her future preferences (or at least her ranking of prizes vs. trees), she should be able to conclude that, upon reaching $h$, she will choose $x$ rather than $f$. Hence, a priori, she should be indifferent between the tree under consideration, and a modified tree wherein the choice $f$ is removed at $h$. Notice that this conclusion must hold for *any* prize $x \succ \bar{x}$, and *any* tree wherein the choices available at $h$ are $f$ and $x$.

Now suppose instead that $\bar{x} \preccurlyeq x_{f|E}$, and consider an arbitrary prize $x \prec \bar{x}$. By analogous considerations, one can conclude that the DM should be indifferent between a tree wherein the choices available at $h$ are $f$ and $x$, and a modified tree wherein the choice $x$ is removed. Again, this must be the case regardless of what the tree specifies at histories that do not follow $h$ (i.e. "outside the event $E$"), and for any prize $x \prec \bar{x}$.

These arguments identify two alternatives, at least one of which *must* hold for any $\succcurlyeq$-essential event $E$, act $f$, and prize $\bar{x}$; neither alternative involves the (possibly unobservable) conditional certainty equivalent $\bar{x}_{f|E}$. Formally:

**Axiom 3.6 (Conjectural Separability)** *Consider $E \in \mathrm{E}(\succcurlyeq)$, $f \in F_E$, and $\bar{x} \in X$; let $h = [(*, E)]$. Then one (or both) of the following statements hold:*

*(i) for all $x \in X$ such that $x \succ \bar{x}$, and all $g \in F_\Omega(E)$, $\{f, x\}_h g \sim \{x\}_h g$;*

*(ii) for all $x \in X$ such that $x \prec \bar{x}$, and all $g \in F_\Omega(E)$, $\{f, x\}_h g \sim \{f\}_h g$.*

The reference to *separability* is justified by noting that, in accordance with the preceding intuitive discussion, both statements (i) and (ii) in Axiom 3.6 impose a restriction on unconditional preferences that must hold regardless of choices available at histories that do not follow $h$ (in the axiom, such choices are specified by the tree $g$). In this respect, the following axiom can be viewed as a counterpart to Savage's Postulate P2 in the present setting.

The third and final axiom on unconditional preferences serves as a counterpart to Axiom 3.2. Suppose that the prize $\bar{x}$ is strictly preferred to any prize that may be obtained if the continuation tree $f \in F_E$ is chosen; then, when the DM contemplates a tree that provides a choice between $f$ and $\bar{x}$ conditional upon the (essential and immediately relevant) event $E$, she should deem $f$ an irrelevant alternative. Similar cosiderations hold if $\bar{x}$ is strictly inferior to any prize that may be obtained in the continuation tree $f$.

**Axiom 3.7 (Conjectural Dominance)** *Consider $E \in \mathrm{E}(\succcurlyeq)$, $g \in F_\Omega(E)$, $f \in F_E$, and $\bar{x} \in X$; let $h = [(*, E)]$. Then*

*(i) if $x(z) \succ \bar{x}$ for all terminal histories $z$ of $f$, then $\{f, \bar{x}\}_h g \sim \{f\}_h g$;*

*(ii) if $x(z) \prec \bar{x}$ for all terminal histories $z$ of $f$, then $\{f, \bar{x}\}_h g \sim \{\bar{x}\}_h g$.*

#### 3.1.3 $E$-Certainty Equivalents, Equivalence Result and Updating

Consider the following definition:

**Definition 2** Consider $E \in \mathrm{E}(\succcurlyeq)$, $f \in F_E$, and $\bar{x} \in X$; let $h = [(*, E)]$. Then $\bar{x}$ is a **$E$-certainty equivalent of** $f$ iff Statements (i) and (ii) in Axiom 3.6 both hold.

Observe that $E$-certainty equivalents are defined solely in terms of the unconditional preference $\succcurlyeq$; as such, they are "conjectural" constructs—they reflect the conjectured evaluation of a continuation tree conditional upon $E$, from the perspective of "time 0". Proposition 2 below implies that this "conjectural" evaluation coincides with the actual "conditional" evaluation of the same act.

First of all, however, the existence of $E$-certainty equivalents must be established; the following Proposition (which, unlike Remark 3.1, does not follow from entirely standard arguments) provides the first main result of this subsection.

**Proposition 1** *Suppose that $\succcurlyeq$ is a complete and transitive relation that satisfies Axioms 3.5, 3.6 and 3.7. Then, for all $E \in \mathrm{E}(\succcurlyeq)$ and $f \in F_E$, the set of $E$-certainty equivalents of $f$ is non-empty, and forms an indifference class of prizes for $\succcurlyeq$.*

The generic $E$-certainty equivalent of the tree $f \in F_E$ will be denoted by $\bar{x}_{f|E}$.

**Proof:** Consider the sets

$$U(f|E) = \Big\{\bar{x} \in X \ : \ \forall x \in X \text{ s.t. } x \succ \bar{x},\ \forall g \in F_\Omega(E),\ \ \{f, x\}_h g \sim \{x\}_h g\Big\}, \tag{1}$$

$$L(f|E) = \Big\{\bar{x} \in X \ : \ \forall x \in X \text{ s.t. } x \prec \bar{x},\ \forall g \in F_\Omega(E),\ \ \{f, x\}_h g \sim \{f\}_h g\Big\}. \tag{2}$$

Clearly, the set of $E$-certainty equivalents of $f$ is $U(f|E) \cap L(f|E)$.

Note that, if $x \in U(f|E)$ [resp. $x \in L(f|E)$] and $x' \sim x$, then clearly $x' \in U(f|E)$ [resp. $x' \in L(f|E)$]. Also, suppose $x \in L(f|E)$, $x' \in U(f|E)$, and $x \succ x'$. By Axiom 3.5, there exist $x_1, x_2 \in X$ such that $x \succ x_1 \succ x_2 \succ x'$.[6] For $i = 1, 2$, $x \in L(f|E)$ and $x_i \prec x$ imply $\{f, x_i\}_h g \sim \{f\}_h g$, whereas $x' \in U(f|E)$ and $x_i \succ x'$ imply $\{f, x_i\}_h g \sim \{x_i\}_h g$. Thus,

$$\{x_1\}_h g \sim \{f, x_1\}_h g \sim \{f\}_h g \sim \{f, x_2\}_h g \sim \{x_2\}_h g,$$

which implies that $E$ cannot be essential. Hence, if $E$ is essential, then $x \in L(f|E)$ and $x' \in U(f|E)$ imply $x' \succcurlyeq x$; in particular, either $U(f|E) \cap L(f, E)$ is empty, or it is an indifference class of prizes for $\succcurlyeq$.

Moreover, it is clear that $x \in U(f|E)$ and $x' \succ x$ imply $x' \in U(f|E)$; similarly, $x \in L(f|E)$ and $x' \prec x$ imply $x' \in L(f|E)$. Now consider $U(f|E)$; since $\Omega$ is finite,[7] there is $\bar{x} \in X$ such that $\bar{x} \succcurlyeq x(z)$ for all terminal histories $z \geq h$. If there is no $x \in X$ with $x \succ \bar{x}$, then $\bar{x} \in U(f|E)$ holds vacuously; otherwise, Axiom 3.7 ensures that, for all $x \succ \bar{x}$, $\{f, x\}_h g \sim \{x\}_h g$, so again $\bar{x} \in U(f|E)$.

[6] Suppose that, for all $x''$, either $x'' \succcurlyeq x$ or $x'' \preccurlyeq x'$: by Axiom 3.5, $\{x'' : x'' \succcurlyeq x\}$ and $\{x'' : x'' \preccurlyeq x'\}$ are closed, so $X$ is the union of two disjoint closed sets, which contradicts the assumption that $X$ is connected. Thus, there is $x_1 \in X$ with $x \succ x_1 \succ x'$. Now repeat the argument with $x_1$ in lieu of $x$ to get a suitable $x_2$.

[7] If it is not, we assume trees are "bounded" in the usual sense, so the argument still goes through.

Similarly, there is $\bar{x}' \in X$ such that $\bar{x}' \preccurlyeq x(z)$ for all $z \geq h$, and again $\bar{x}' \in L(f|E)$. Hence, in particular, both $U(f|E)$ and $L(f|E$ are non-empty.

It follows that either $U(f|E) = \{x : x \succ x_U\}$ or $U(f|E) = \{x : x \succcurlyeq x_U\}$ for some $x_U \in X$. Also, in either case, for any $x \succ x_U$, $\{f, x\}_h g \sim \{x\}_h g$: to see that, recall that, as argued above, there exists $x'$ such that $x \succ x' \succ x_U$, so in either case $x' \in U(f|E)$, and the required conclusion holds. But then by definition $x_U \in U(f|E)$, i.e. it must be the case that $U(f|E) = \{x : x \succcurlyeq x_U\}$. Similarly, $L(f|E) = \{x : x \preccurlyeq x_L\}$ for some $x_L \in X$.

Finally, by Axiom 3.6, $U(f,E) \cup L(f,E) = X$. By Axiom 3.5, $U(f|E)$ and $L(F|E)$ are closed, so they cannot be disjoint because $X$ is connected. Clearly, any $x \in U(f|E) \cap L(f|E)$ is an $E$-certainty equivalent of $f$. ■

It is now possible to establish the equivalence of the two approaches to conditional preferences discussed in this subsection.

**Proposition 2** *Consider the conditional preference system $\{\succcurlyeq_E\}_{\emptyset \neq E \subset \Omega}$. Assume that $\succcurlyeq$ is a complete and transitive relation on $F_\Omega$. Then the following statements are equivalent.*

1. *$\succcurlyeq$ satisfies Axioms 3.5, 3.6 and 3.7; furthermore, for all $E \in \mathrm{E}(\succcurlyeq)$ and all $f, g \in F_E$, $f \succcurlyeq_E g$ if and only if $\bar{x}_{f|E} \succcurlyeq \bar{x}_{g|E}$.*

2. *$\{\succcurlyeq_E\}_{E \in \mathrm{E}(\succcurlyeq)}$ is a collection of complete and transitive relations that satisfies Axioms 3.1, 3.2, 3.3 and 3.4.*

*Furthermore, if either condition holds, then for every $E \in \mathrm{E}(\succcurlyeq)$, any prize $x \in X$, and any tree $f \in F_E$, $x \sim_E f$ if and only if $x$ is an $E$-certainty equivalent of $f$.*

Thus, if conditional preferences are defined as in 1 above, they satisfy the regularity and consistency axioms of subsection 3.1.1; conversely, if one assumes the existence of conditional preferences that satisfy the axioms in subsection 3.1.1, then *un*conditional preferences must satisfy the axioms in subsection 3.1.2 (in particular, Conjectural Separability), and it must be the case that the conditional ordering of any two acts coincides with the (unconditional) ordering of their respective $E$-certainty equivalents—which can be determined by observing *un*conditional preferences. The final claim in Proposition 2 confirms that, although $E$-certainty equivalents are a "conjectural" construct defined in terms of unconditional preferences alone, they do coincide with the intuitively more direct notion of conditional certainty equivalents.

**Proof:** Let $U(f|E)$ and $L(f|E)$ be as in Eqs. (1) and (2). Observe that $x \succcurlyeq \bar{x}_{f,E} \succcurlyeq x'$ for all $x \in U(f|E)$ and $x' \in L(f|E)$.

$\Rightarrow$: it is clear that, if $f \succcurlyeq_E g$ iff $\bar{x}_{f|E} \succcurlyeq \bar{x}_{g|E}$, then $\succcurlyeq_E$ is a weak order. Furthermore, consider a constant act $\bar{x}$. By Axiom 3.7, $\bar{x} \in U(\bar{x}|E) \cap L(\bar{x}|E)$; hence, $\bar{x}_{\bar{x},E} \sim \bar{x}$ for any constant $\bar{x}$; this implies that $\succcurlyeq_E$ satisfies Axiom 3.1.

To verify that Axiom 3.2 holds, consider $x \in X$ such that $x' \succcurlyeq x(z)$ for all terminal $z$. If there is no $x \succ x'$, then $x' \in U(f|E)$ holds vacuously; otherwise, Axiom 3.7 ensures that $\{f, x\}_h g \sim \{x\}_h g$ for all $x \succ x'$ and $g \in F_\Omega(E)$, so again $x' \in U(f|E)$. This implies that $x' \succcurlyeq \bar{x}_{f,E}$, and hence $x' \succcurlyeq_E f$, as required. The argument for $x'' \preccurlyeq x(z)$ is analogous.

As for Axiom 3.3, the first set in the statement can be written as $\{x \in X : x \succcurlyeq_E \bar{x}_{f,E}\}$, i.e. $\{x \in X : x \succcurlyeq \bar{x}_{f,E}\}$, which is closed by Axiom 3.5; similarly for the other set.

Finally, to see that Axiom 3.4 holds, note that $x \succ_E f$ implies $x \succ_E \bar{x}_{f,E}$, i.e. $x \succ \bar{x}_{f,E}$; since $\bar{x}_{f,E} \in U(f|E)$, this implies $\{f, x\}_h g \sim \{x\}_h g$ for all $g \in F_\Omega(E)$; similarly for (ii).

$\Leftarrow$: consider $E \in \mathrm{E}(\succcurlyeq)$. Since $\succcurlyeq_E$ satisfies Axiom 3.1 and 3.3, $\succcurlyeq$ must satisfy Axiom 3.5: just take $f = \bar{x}$ in Axiom 3.3. Next, for Axiom 3.6, let $x_{f|E}$ be such that $x_{f|E} \sim f$ (one such prize must exist by Remark 3.1). Suppose $\bar{x} \succcurlyeq x_{f,E}$: then $x \succ \bar{x}$ implies $x \succ_E f$, so Axiom 3.4 implies $\{f, x\}_h g \sim \{x\}_h g$ for any $g \in F_\Omega(E)$; if instead $\bar{x} \preccurlyeq_E x_{f,E}$, $x \prec \bar{x}$ implies $x \prec_E f$, and again Axiom 3.4 yields the required conclusion.

Now consider Axiom 3.7, case (i). Let $x_L$ be such that $x_L \preccurlyeq x(z)$ for all terminal histories $z$ of $f$, and $x_L = x(z_L)$ for some such $z_L$. Axiom 3.2 implies that $f \succcurlyeq_E x_L$, and by assumption $x_L \succ \bar{x}$. By Axiom 3.1 and transitivity, $f \succ_E \bar{x}$, so Axiom 3.4 implies $\{f, x\}_h g \sim \{f\}_h g$ for all $g \in F_\Omega(E)$, as required. Case (ii) is analogous.

Finally, it will be shown that $x \sim_E f$ iff $x \sim \bar{x}_{f|E}$, which is the last claim of Proposition 2; this also implies that $f \succcurlyeq_E g$ iff $\bar{x}_{f|E} \succcurlyeq \bar{x}_{g|E}$, which completes the proof that 2 implies 1. Suppose $x \sim_E f$: then $x' \succ x$ implies $x' \succ_E f$ by Axiom 3.1 and transitivity, and hence $\{f, x'\}_h g \sim \{x'\}_h g$ whenever $h = [(*, E)]$ and $g \in F_\Omega(E)$, by Axiom 3.4. Similarly, if $x' \prec x$, then $x' \prec_E f$, and hence $\{f, x'\}_h g \sim \{f\}_h g$. Thus, $x \sim_E f$ implies that $x$ is an $E$-certainty equivalent of $f$. Conversely, since it was just shown that $\succcurlyeq$ satisfies Axioms 3.5, 3.6 and 3.7, Proposition 1 can be invoked to conclude that the set of $E$-certainty equivalents is an indifference class for $\succcurlyeq$; by the argument just given it contains all $x \in X$ such that $x \sim_E f$: thus, $\bar{x}_{f|E} \sim x$, which implies $\bar{x}_{f|E} \sim_E f$ by Axiom 3.1 and transitivity, as required. ■

### 3.2 A decision-theoretic analysis of Backward Induction

This section focuses on the characterization of backward induction. The key ingredient is the assumption that the DM correctly anticipates her future preferences, and incorporates these expectations into her present choices. This assumption is discussed and formalized in §3.2.1; the characterization result is provided in §3.2.2.

### 3.2.1 Sophistication and Weak Commitment; Reduction of Compound Acts

It is useful to break this "correct expectations" assumption into two parts. The first, deemed *sophistication*, concerns future *strict* preferences: it prescribes that the DM be ex-ante indifferent between a tree that makes available a collection of actions at some history $h$, and another tree that differs from the first only in that one or more ex-post strictly dominated actions are removed. Intuitively, the DM can be sure that ex-post strictly dominated actions will never be chosen, so she should deem such actions irrelevant to her present decision problem.

The following axiom formalizes this intuition. Recall that, at any non-terminal history $h$, any action $a \in A_f(h)$ available in the tree $f$ at $h$, corresponds to a continuation tree $f(h,a) \in F_{E(h)}$; thus, the DM's conditional preference $\succcurlyeq_{E(h)}$ readily induces an ordering over actions available at $h$, which makes it possible to formalize the assumption that a subset $B$ of such actions are conditionally strictly dominated. The axiom then requires that, ex-ante, the DM deem actions in $B$ irrelevant.

**Axiom 3.8 (Sophistication)** *For all $f = (E,H,x) \in F_E$ and all $h \in H$ non-terminal: if, for some $B \subset A_f(h)$, $b \in B$ and $w \in A_f(h^+) \setminus B$ imply $f(h,b) \succ_{E(h)} f(h,w)$, then $f \sim_E \{f(h,b) : b \in B\}_h f$.*

The second part of the "correct expectations" assumption pertains to *tie breaking*: what if, upon reaching the history $h$, the DM is indifferent between two actions available there? As the example discussed in the Introduction demonstrates, the way indifference is resolved may affect her ex-ante evaluation of the tree under consideration. The approach taken in this subsection was introduced by Strotz [16], who refers to it as the strategy of *consistent planning*; according to this approach, the DM's problem is "to find the [ex-ante] best plan among those that he will actually follow [ex-post]" (cf. [16, p.173])

This may be viewed as a weak form of *commitment*. Ex-ante, the DM cannot commit to take actions later that, ex-post, she will deem inferior: this is ruled out by Axiom 3.8. However, she *can* commit to any one of the actions that, ex-post, she deems optimal; intuitively, her "future self" has no reason to object to her current self's commitment.

To formalize this assumption it is necessary to introduce additional notation. Consider an event $E \subset \Omega$, a tree $f \in F_E$, $h \in H$, and an action $a \in A_f(h)$. Recall that, after choosing $a$ at $h$, the DM learns that one of the events in the set $\mathcal{F}_f(h,a)$—say, $E'$—contains the true state. If $E'$ is a singleton, then the DM receives a prize and the decision problem terminates; otherwise, she will be able to choose any one of the actions available at $[h,(a,E')]$. I first define a continuation tree in $F_{E(h)}$ that differs from $f(h,a)$ only in that, following the choice of $a$, the DM commits to choosing one particular action at every non-terminal history of the form $[h,(a,E')]$, with $E' \in \mathcal{F}_f(h,a)$. Formally, assume that $E_1,\dots,E_n$ are the only non-singleton events in $\mathcal{F}_f(h,a)$, and consider $b_m \in A_f([h,(a,E_m)])$, for $m=1,\dots,n$; finally, let $f^0 = f(h,a)$ and $f^m = \{f([h,(a,E_m)],b_m)\}_{[h,(a,E_m)]} f^{m-1}$ for $m=1,\dots,n$. Then $f^n$ differs from $f$ only in that the choice $b_m$ is made at $[h,(a,E_m)]$, as required; to make this explicit, let $f(h,a;b_1,E_1;\dots;b_n,E_m) = f^n$, and denote by $F_f(h,a)$ the collection of all continuation trees thus obtained: that is, $F_f(h,a) = \{f(h,a;b_1,E_1;\dots;b_n,E_n) \,:\, \forall m,\, b_m \in A_f([h,(a,E_m)]\}$. Finally, let $M_f(h,a)$ be the set of $\succcurlyeq_{E(h)}$–maximal elements of $F_f(h,a)$: that is,

$$M_f(h,a) = \{g \in F_f(h,a) \,:\, \forall g' \in F_f(h,a),\; g \succcurlyeq_{E(h)} g'\}; \tag{3}$$

Intuitively, the continuation trees in $M_f(h,a)$ represent courses of actions the DM *would* want to follow at history $h$, *if* she could commit to appropriate history-$[h,(a,E_m)]$ choice for every $m$.

The appropriate notion of commitment can now be formalized. Consider the following situation: at every history of the form $[h,(a,E')]$, with $E' \in \mathcal{F}_f(h,a)$ as above, (a) all actions available to the DM correspond to *plans*, and (b) the DM is indifferent among all such actions. That is: if there is any commitment problem at any history following the DM's choice of $a$ at $h$, then it must pertain to choices at histories *immediately* following $a$; furthermore, there is in fact *no* commitment problem, because the DM is actually indifferent among all actions she has available at such histories. Then, it makes sense to assume that (i) the DM is able to commit at $h$ to the choice(s) she likes best at histories that immediately follow $a$. Furthermore, (ii) ex-ante, the DM should understand that her "history-$h$ self" *will* have this capability to commit, and will take advantage of it; in particular, the DM's time-0 (empty history) self cannot hope to "commit" to a continuation tree at $h$ that is not an element of $M_f(h,a)$: her history-$h$ self has the last say as to which actions can be prescribed to the history-$[h,(a,E_m)]$ selves. This is formalized by requiring that the DM be indifferent ex-ante between $f$ and another tree that differs from it in that the continuation tree $f(h,a)$ at $h$ is replaced by the collection $M_f(h,a)$ of $\succcurlyeq_{E(h)}$–maximal "one-period-commitment" trees (whereas all other continuation trees $f(h,a')$, with $a \neq a' \in A_f(h)$, are unchanged).

**Axiom 3.9 (Weak Commitment)** *For all $f=(E,H,x)\in F_E$ and all histories $h \in H$: if, for some $a \in A_f(h)$, and for every non-singleton $E' \in \mathcal{F}_f(h,a)$, (a) $f([h,(a,E')],a')$ is a plan for every $a' \in A_f([h,(a,E')])$, and (b) $f([h,(a,E')],a') \sim_{E'} f([h,(a,E')],b')$ for all $a',b' \in A_f([h,(a,E')])$, then (i) $f(h,a) \sim_{E(h)} g$ for all $g \in M_f(h,a)$, and (ii)*

$$f \sim_E \{\{f(h,a') \,:\, a' \in A_f(h)\setminus\{a\}\} \cup M_f(h,a)\}_h f.$$

As noted in the Introduction, all the axioms in this subsection have the flavor of (and have been motivated as) restrictions on the DM's beliefs about the preferences and behavior of her future selves. Yet, they are fully behavioral assumptions that can be tested, at least in principle.

Hence, the approach taken in this subsection is motivated by game-theoretic considerations, but its substance is decision-theoretic.

### 3.2.2 Formulation and Characterization of Backward Induction

Backward induction in decision trees can now be characterized as the iterative application of Axioms 3.8 and 3.9; upon each iteration, a suitably smaller and simpler equivalent tree is obtained. The procedure terminates when the reduced tree is a *plan*, i.e. a decision tree wherein a single action is available at every non-terminal history. Furthermore, in order to carry out the procedure, it is enough to specify the DM's conditional preferences over plans. One way to do so is to assume that the DM evaluates a plan according to the assignment of prizes to states that it determines; hence, *any system of conditional preferences over Savage acts can be uniquely extended to a system of conditional preferences over trees.* This approach will be exemplified in the next section.[8]

The main result of this section also shows that, if backward induction is used to extend preferences from plans to arbitrary trees, then the resulting conditional preference system satisfies Axioms 3.8 and 3.9. Hence, these axioms fully characterize a complete behavioral theory of dynamic choice in the presence of ambiguity.

Consider a tree $f = (E, H, x)$; say that an action $a$ at a history $h \in H$ is **reducible** if $f(h, a)$ is *not* a plan and, for every non-singleton element $E'$ of $\mathcal{F}_f(h, a)$, and every $b \in A_f([h, (a, E')])$, $f([h, (a, E')], b)$ is a plan. That is, the action $a$ is reducible if the only choices the DM needs to make are at histories immediately following $h$ and $a$. Notice that, since $f(h, a)$ is not a plan, it is a fortiori not an act, which implies that one or more of the elements of $\mathcal{F}_f(h, a)$ *must* be non-singleton.

**Algorithm 1 (Backward Induction)** Let $E \subset \Omega$ be nonempty and $f \in F_E$.

1. Find a history $h$ and a reducible action $a \in A_f(h)$. IF there is none, then STOP and RETURN the set $\{f\}$.

2. Denote the non-singleton elements of $\mathcal{F}_f(h, a)$ by $E_1, \ldots, E_n$. Inductively construct an act $f' \in F_E$ as follows. Let $f^0 = f$ and, for $m = 1, \ldots, n$,

    (a) $h^m = [h, (a, E_m)]$,

    (b) $M_m = \{b' \in A_f(h^m) : \forall a' \in A_f(h^m),\ f(h^m, b') \succcurlyeq_{E_m} f(h^m, a')\}$,

    (c) $f^m = \{f(h^m, b') : b' \in M_m\}_{h^m} f^{m-1}$.

[8] It should be noted, however, that the results in this section are independent of the actual way the DM evaluates plans: they do *not* require that plans be evaluated by reducing them to acts. In particular, they can accommodate a preference for early or late resolution of uncertainty.

The required act $f'$ is then $f^n$.

3. IF $h = \emptyset$, then STOP and RETURN the set $M_{f'}(h,a) = M_{f'}(\emptyset, *)$.
4. Let $f'' = \Big\{ \{ f(h,a') \,:\, a' \in A_f(h) \setminus \{a\} \} \cup M_{f'}(h,a) \Big\}_h f$.
5. REPLACE $f$ WITH $f''$ and GO TO step 1.

Thus, the Backward induction algorithm (BI henceforth) produces a sequence of trees $f^0, \ldots, f^K$ (a **run**) followed by a set of trees as output. For every $k = 1, \ldots, K$, $f^k$ differs from $f^{k-1}$ in that some reducible action is replaced with a set of (actions corresponding to) *plans*. To see this, observe that, by definition, actions at histories following a reducible actions are themselves plans; hence, the elements of the set $M_{f'}(h,a)$ referred to in Step 4 are necessarily plans. This also implies that the output of the algorithm consists of plans.

Notice that, in general, an action $a$ at a history $h$ that is not reducible in the initial tree $f$ may[9] become reducible after a number of iterations have been performed. But this requires that actions at subsequent histories be reduced first. Thus, the "backward" qualifier is justified: in general, iterations in BI begin close to terminal histories, and move up to the root (but see below).

It is also useful to notice that, in any given run $f^0, \ldots, f^K$ of BI, every history $h$ of the original tree $f$ is also a history of $f^0, \ldots, f^{k^*}$, where $k^*$ is the iteration at which the last reducible action at $h$ is replaced.[10]

These considerations immediately imply that:

1. BI terminates in a finite number of steps;
2. In each iteration, BI only requires comparisons among *plans*, not general trees;
3. The output of BI consists of a (possibly singleton) set of *plans* that belong to the same indifference class for the DM.

For BI to be operational, conditional preferences over plans must clearly be specified; however, notice that the above conclusions do not rely on any of the axioms mentioned in the preceding subsection: they follow from the definition of the algorithm. As will be clear momentarily, the role of these axioms is to ensure that the output of the BI procedure consists of plans that the DM deems indifferent to the input tree $f$.

[9] It *need not* become reducible, however, because the tree $f$ could be such that, after a suitable number of iterations, there is a unique action available at histories following $h$ and $a$: that is, $f^k(h,a)$ may be a plan for $k$ suitably large. In this case, no further reduction is necessary.

[10] It may also be a history of subsequent elements of the run, in case $a(h)$ is not immediately replaced in the $(k+1)$-th iteration. This, however, is unimportant for the present purposes.

One last issue must be addressed before the main result can be stated. The specification of BI does not constrain the choice of a reducible action among those available at Step 1; hence, more than one run may be consistent with the specification provided above. However, this is immaterial as far as the output of BI is concerned. Again, this conclusion does *not* rely on the axioms in the preceding subsection. It is useful to state the result in a more general (but equivalent) form.

For future reference, it is useful to define the **height** $\eta(h)$ of a history $h$ as follows. Fix a tree $f = (E, H, x) \in F_E$; if $h$ is terminal, then $\eta(h) = 0$; otherwise, $\eta(h) = \max\{\eta(h') : h < h'\} + 1$.

**Remark 3.2** Consider two runs $f^0, \ldots, f^K$ and $\bar{f}^0, \ldots, \bar{f}^{\bar{K}}$ of BI, possibly originating from different inputs. If there exist $k, \bar{k}$ such that $f^k = \bar{f}^{\bar{k}}$, then both runs yield the same output.

**Proof:** It is clear that $f^k, \ldots, f^K$ and $\bar{f}^{\bar{k}}, \ldots, \bar{f}^{\bar{K}}$ are also runs of BI, so it is wlog to take $k = \bar{k} = 0$, so $f^0 = \bar{f}^0 = f$. Arguing by contradiction, if the two runs have different outputs, there must be some history and some action available at that history that is replaced with different (sets of) plans in the two runs; consider the *lowest-height* history $h$ where such an action $a$ can be found. Formally, let $\ell$ and $\bar{\ell}$ be such that (1) $h$ is a history of $f$ and $a$ is reducible in $f^\ell$ and $\bar{f}^{\bar{\ell}}$ (hence, as argued above, $h$ is also a history of $f^\ell$ and $\bar{f}^{\bar{\ell}}$); (2) $a$ is replaced in the $\ell$-th and $\bar{\ell}$-th iterations of the two runs respectively, with *different* sets of plans; and (3) at all histories of $f$ of height less than $\eta(h)$, the two runs replace reducible actions with the same set of plans.

Since $a$ is reducible, for all $E' \in \mathcal{F}_{f^\ell}(h, a) = F_{\bar{f}^{\bar{\ell}}}(h, a) = F_f(h, a)$, $f^\ell([h, (a, E')], a')$ is a plan for all $a' \in A_{f^\ell}([h, (a, E')])$, and $\bar{f}^{\bar{\ell}}([h, (a, E')], a')$ is a plan for all $a' \in A_{\bar{f}^{\bar{\ell}}}([h, (a, E')])$.

Furthermore, there must be $E^*$ such that $\{f^\ell([h, (a, E^*)]), a') : a' \in A_{f^\ell}([h, (a, E^*)])\} \neq \{\bar{f}^{\bar{\ell}}([h, (a, E^*)]), a') : a' \in A_{\bar{f}^{\bar{\ell}}}([h, (a, E^*)])\}$, or the algorithm would prescribe the same replacement for $a$ in both runs. But this implies that the two runs of the algorithm perform different replacements for some action $a^* \in A_f([h, (a, E^*)])$, which contradicts the assumption that equal replacements are performed at all histories of height less than $\eta(h)$. ■

**Corollary 3** *Suppose that BI outputs the collection $\{g_1, \ldots, g_N\}$ for the continuation act $f(h, a)$; then BI produces the same output for both $f$ and $\{\{f(h, b) : b \in A_f(h) \setminus \{a\}\} \cup \{g_1, \ldots, g_N\}\}_h f$.*

**Proof:** The claim is obvious if $f(h, a)$ is itself a plan; thus, assume it is not. Consider a run $f^0, \ldots, f^K$ with input $f^0 = f$, and a run $\bar{f}^0, \ldots, \bar{f}^{\bar{K}}$ with the other act under consideration as input. Suppose that the runs have the following features: in $f^0, \ldots, f^K$, only histories that follow $h$ and $a$ are considered (i.e. histories $h'$ such that $h_t = h$ and $a(h_{t+1}) = a$ for some $t < \lambda(h')$), until, for some $\ell > 0$, either $f^\ell(h, a)$ is a plan, or $a$ is reducible at $h$ in $f^\ell$. Now observe that $f^0(h, a), \ldots, f^\ell(h, a) \in F_{E(h)}$ constitute a run of BI with input $f(h, a)$. If $f^\ell(h, a)$ is a plan, then BI must output $\{f^\ell(h,a)\}$ for the input $f(h,a)$, so in fact $f^\ell = \bar{f}^0$, and the Remark yields the result. In the other case, an additional iteration of BI for the input $f(h,a)$ yields $\{g_1,\ldots,g_N\}$; wlog, assume that the run $f^0,\ldots,f^K$ with input $f$ also replaces the reducible action $a$ at the $(\ell+1)$-st iteration; the set of plans substituted for $f^\ell(h,a)$ is clearly $\{g_1,\ldots,g_N\}$. But then $f^{\ell+1} = \bar{f}^0$, and again the claim follows from the Remark. ■

The main result of this section follows.

**Theorem 4** *Consider a system of preferences $\{\succcurlyeq_E\}_{\emptyset \neq E \subset \Omega}$ such that, for every non-empty $E \subset \Omega$, $\succcurlyeq_E$ is a complete and transitive binary relation on $F_E$. Then the following statements are equivalent.*

1. *Axioms 3.8 and 3.9 hold;*
2. *For every non-empty $E \subset \Omega$ and every $f \in F_E$, if BI outputs the collection $\{g_1,\ldots,g_n\}$ of plans, then $f \sim_E g_1 \sim_E \ldots \sim_E g_n$.*

**Proof:** (1) $\Rightarrow$ (2): It must first be shown that, in any run $f^0,\ldots,f^K$ of BI, $f^k \sim_E f^{k-1}$ for all $k > 0$; then, it will be argued that, starting with $f^K$, an additional iteration of BI yields a collection of plans $\{g_1,\ldots,g_N\}$ that satisfies the required property.

Consider an arbitrary non-terminal iteration $k \in \{1,\ldots,K\}$ (if BI terminates immediately, there is no such iteration, of course). By convention, the input to this iteration is $f^{k-1}$, and the output is $f^k$. In the description of Algorithm 1, at every iteration, $f^{k-1}$ is $f$ and $f^k$ is $f''$; thus, it must be shown that $f \sim_E f''$ if the algorithm does not terminate in Steps 1 or 3.

Since the algorithm does not terminate, there must be a reducible action $a$ at a history $h$ in Step 1. Now consider Step 2. Clearly, for $m = 1,\ldots,n$, Axiom 3.8 implies that $f^m \sim_E f^{m-1}$ (take $B = M^m$), so $f \sim f'$. Since BI does not terminate at the $k$-th iteration, consider Step 4 next. Then (ii) in Axiom 3.9 ensures that $f'' \sim_E f' \sim_E f$ (note that $f'$ and $f$ coincide at histories that do not weakly follow $h$). This proves the first claim.

We thus have $f^0 \sim_E f^K$. If now, at the $(K+1)$-th iteration, BI terminates in Step 1, the output is $\{f^K\}$ and the result follows immediately. Otherwise, BI must terminate in Step 3; this means that the only reducible action is $*$ at $\emptyset$. In this case, again denote the input $f^K$ to the last iteration of BI by $f$; as above, the tree $f'$ constructed in Step 2 satisfies $f \sim_E f'$. Furthermore, (i) in Axiom 3.9 implies that, for any $g \in M_{f'}(\emptyset, *)$, $f' = f'(\emptyset,*) \sim_E g$; therefore, $f \sim_E g$ as well, and the proof of this direction is complete.

(2) $\Rightarrow$ (1): Consider Axiom 3.8 first, and let $f$ and $h$ be as in the statement of the latter. For every $a \in A_f(h)$, denote by $G(h,a)$ the collection of plans in $F_{E(h)}$ that BI produces for the input $f(h,a)$. Corollary 3 implies that BI yields the same output for $f$ and for the tree where each $f(h,a)$

is replaced with $G(h,a)$. The same is true of the act $\{f(h,b) : b \in B\}_h f$. Moreover, by assumption, $f(h,b) \succ_{E(h)} f(h,a)$ iff $g_{h,b} \succ_{E(h)} g_{h,a}$ for any (hence all) $g_{h,b} \in G(h,b)$ and $g_{h,a} \in G(h,a)$. Hence, it is sufficient (and notationally simpler) to prove the result under the assumption that every $f(h,a)$ is a plan.

By Remark 3.2, it is wlog to assume that, for both trees, BI begins by replacing reducible actions at histories that follow $h_{\lambda(h)-1}$; clearly, no reduction will occur at $h$ and at histories that follow $h$, because every action therein corresponds to a plan by assumption. Notice that $f$ and $\{f(h,b) : b \in B\}_h f$ are identical at histories that follow $h_{\lambda(h)-1}$, but do not weakly follow $h$: hence, after $\ell$ iterations, the two runs will produce trees $f^\ell$ and $\bar{f}^\ell = \{f(h,b) : b \in B\}_h f^\ell$; furthermore, in such trees, $a(h)$ is reducible.

In the $(\ell+1)$-th iteration, it is wlog to assume that BI replaces $a(h)$; but notice that the *same* modified acts will be constructed in Step 2 (denoted $f'$ there) for both $f^\ell$ and $\bar{f}^\ell$: in particular, the continuation plans selected at $h = [h_{\lambda(h)-1}, (a(h), E(h))]$ will correspond to one or more actions in the set $B$. Hence, either BI terminates with the same output, or $f^{\ell+1} = \bar{f}^{\ell+1}$, in which case, again, Remark 3.2 ensures that the output for the two runs will be identical. Hence, $f \sim_E \{f(h,b) : b \in B\}_h f$.

Next, consider Axiom 3.9, and let $f, h, a$ be as in the statement of the latter. Note that, by assumption, for every non-singleton $E' \in \mathcal{F}_f(h,a)$ and $a' \in A_f([h,(a,E')])$, $f([h,(a,E')],a')$ is a plan, so either $f(h,a)$ is a plan, or $a$ is reducible.

If $f(h,a)$ is a plan, then (i) and (ii) hold because $M_f(h,a) = \{f(h,a)\}$, so the two acts under consideration in (i) and (ii) actually coincide. Otherwise, for (i), apply BI to $f(h,a)$. Since the DM is indifferent among all continuation plans at histories following $h$, the act $f'$ constructed in Step 2 coincides with $f(h,a)$; moreover, BI will terminate in Step 3 and output $M_{f(h,a)}(\emptyset,*) = M_f(h,a)$. The required indifference then follows from the assumption that the DM is indifferent between a tree and any element of the BI output for that tree (the "extension assumption"). For (ii), apply BI to $f$; since $a$ is reducible, it is wlog to assume that the algorithm begins by replacing $a$ at $h$. Again, the act $f'$ constructed in Step 2 coincides with $f$; in Step 4, the act $\{\{f(h,a') : a' \in A_f(h) \setminus \{a\}\} \cup M_f(h,a)\}_h f$ is obtained from $f$; the next iteration then begins. But this is precisely the act appearing in the r.h.s of the indifference in (ii), so Remark 3.2 implies that this run of BI for $f$ and any run of BI for the modified act will yield the same output. Again, by the extension assumption, this implies that the required indifference will hold. ■

### 3.3 Alternatives to Weak Commitment

[TO BE WRITTEN]

## 4 An application: Full Bayesian Updating for MEU preferences

This section considers preferences consistent with the MEU decision model (Gilboa and Schmeidler [6]) and provides a characterization of prior-by-prior, or "full" Bayesian updating. By Theorem 4, it is sufficient to provide a characterization of this updating rule for preferences over plans. For simplicity, it will be assumed here that the DM is indifferent between a plan and the corresponding Savage act, so the analysis can be carried out entirely in the "standard" setting of conditional preferences over maps from states to prizes. As was noted in the preceding section, the results on observability and backward induction do *not* depend upon this assumption in any way.

It must also be emphasized that the results in this section are merely meant to exemplify the approach suggested in this paper. It is straightforward to adapt the analysis to different representations of preferences (e.g. Choquet-expected utility) and different updating rules (e.g. the Dempster-Shafer rule).

### 4.1 Setup

#### 4.1.1 Acts and Plans

Recall that, in the present setting, a tree $f = (E, H, x) \in F_E$ is an act if $H$ consists of the empty history, and of histories of the form $[(*, \{\omega\})]$, for all $\omega \in \Omega$. To simplify the notation, the prize assigned by $f$ to the terminal history $[(*, \{\omega\})]$, corresponding to the event that $\omega$ is the prevailing state, will be denoted simply by $f(\omega)$, insted of $x([(*, \{\omega\})])$. This is in accordance with standard notation for Savage acts.

Also recall that a plan is an act wherein a single action is available at every non-terminal history. Clearly, every act is a plan, but the converse is not true; however, there is a "natural" map from plans to acts. It is easy to verify that, in any plan, a unique prize can be assigned to every state $\omega \in \Omega$ by following the unique choices made by the DM at each history, and Nature's choices corresponding to $\omega$: formally, if $f = (E, H, x) \in F_E$ is a plan, then for every $\omega \in \Omega$ there is a unique terminal history $z_\omega \in H$ such that $E(z_\omega) = \{\omega\}$.[11] Hence, one can define an act, denoted $\hat{f} = \{E, \hat{H}, \hat{x}\}$, characterized by $\hat{H} = \{\emptyset\} \cup \{[(*, \{\omega\})] : \omega \in E\}$, and $\hat{x}([(*, \{\omega\})]) = x(z_\omega)$.

It is convenient to denote the set of Savage acts and plans in $F_E$ by $F_E^a$ and $F_E^p$ respectively. Also, it is convenient to introduce simplified notation for combinations of acts. Specifically, given

[11] Suppose there are two, say $z, z'$; let $h$ be the longest history such that $h \le z$, $h \le z'$ (perhaps $h = \emptyset$). Suppose that in fact $h < z, z'$. Since $f$ is a plan, $A_f(h) = \{a_h\}$ is a singleton, so there must be $E, E' \subset E(h)$ such that $E \cap E' = \emptyset$and $[h, (a_h, E)] \le z$, $[h, (a_h, E')] \le z'$. But this implies that $E(z) \subset E$ and $E(z') \subset E'$, so $E \cap E' \neq \emptyset$: hence, it cannot be that $h < z, z'$. The argument for $h < z$, $h = z'$ is analogous.

$E \subset \Omega$, $f \in F^a_E$ and $g \in F^a_{\Omega\setminus E}$, denote by $fEg$ the act in $F_\Omega$ such that

$$\forall \omega \in E, \qquad fEg(\omega) = \begin{cases} f(\omega) & \omega \in E; \\ g(\omega) & \omega \in \Omega \setminus E. \end{cases}$$

### 4.1.2 MEU representation; Reduction

It will now be assumed that unconditional preferences over acts have a MEU representation.

**Assumption 4.1 (MEU)** There exists a closed,[12] convex set of probability measures $C \subset \Delta(\Omega)$ and a continuous function $u : X \to \mathbb{R}$ such that

$$\forall f, g \in F^a_\Omega, \qquad f \succcurlyeq g \quad \Leftrightarrow \quad \min_{q\in C} \int_S u(f(s))q(d\omega) \geq \min_{q\in C} \int_S u(g(s))q(d\omega).$$

Moreover, there exist $f, g \in F^a_\Omega$ such that $f \succ g$.

Notice that no assumption is made concerning the representation of conditional preferences; it will be enough to require that they be complete and transitive. The last part of Assumption 4.1 is a non-triviality requirement.

For simplicity, it will also be assumed that all non-empty events are $\succcurlyeq$-essential. The following requirement is slightly weaker than the notion of essentiality introduced in the preceding section, because it only pertains to acts; however, if Axiom 4.2 holds, together with Axioms 3.8 and 3.9, the two notions turn out to coincide.

**Assumption 4.2** For all non-empty $E \subset \Omega$, and for all $x, x' \in X$ such that $x \succ x'$, $x \succ xEx' \succ x'$.

It is easy to see that Assumption 4.2 holds if and only if $\min_{q\in C} q(E) > 0$ for all non-empty events $E \subset \Omega$.

Finally, the assumption that the DM evaluates plans by reducing them to acts is explicitly formulated, which employs the notation introduced above.

**Assumption 4.3** For all non-empty $E \subset \Omega$, and all $f \in F^p_E$: $f \sim_E \hat{f}$.

## 4.2 Dynamic Consistency and Constant-Act Dynamic Consistency

It is useful to consider the standard Dynamic Consistency axiom as a starting point.

[12]If the analysis is extended to infinite state spaces $\Omega$, then $C$ must be assumed to be weak*-closed.

**Axiom 4.1 (Dynamic Consistency)** *For all non-empty $E \subsetneq \Omega$, all $f, g \in F^a_E$, and all $f', g' \in F^a_\Omega$:*

$$f \succcurlyeq_E g,\ f' \succcurlyeq_{\Omega\setminus E} g' \quad \Longrightarrow \quad fEg \succcurlyeq f'Eg',$$

*and the unconditional preference is strict if either one of the conditional preferences is.*

As was noted in the Introduction, Dynamic Consistency may be inconsistent with full Bayesian updating and non-neutral attitudes towards ambiguity. It is useful to revisit one particularly simple example, based on Ellsberg's three-color urn model. Let $\Omega = \{\rho, \upsilon, \beta\}$, representing the draw of a red, yellow, or blue ball respectively, and assume that $X$ contains the prizes $x = 0$ and $x = 10$; recall that the urn is assumed to contain 30 red balls and 60 non-red balls, in unspecified proportions. Let $E = \{\rho, \upsilon\}$, and consider the acts $f, f' \in F^a_E$ and $g, g' \in F^a_{\Omega\setminus E}$ such that

$$f(\rho) = 10, \quad f(\upsilon) = 0; \qquad f'(\rho) = 0, \quad f'(\upsilon) = 10; \qquad g(\beta) = g'(\beta) = 10.$$

If unconditional preferences have a MEU representation with $u(10) > u(0)$ and $C = \{q \in \Delta(\Omega) : q(\rho) = \frac{1}{3},\ q(\upsilon) \geq \epsilon,\ q(\beta) \geq \epsilon\}$ for some $\epsilon \in (0, \frac{1}{3})$, then it is easy to verify that $fEg \prec f'Eg'$. If $\succcurlyeq_E$ and $\succcurlyeq_{\Omega\setminus E}$ are also MEU preferences, characterized by the set of posteriors derived from $C$ by full Bayesian updating, i.e.

$$C_E = \left\{ q \in \Delta(S) \ :\ \frac{\frac{1}{3}}{1-\epsilon} \leq q(\rho) = 1 - q(\upsilon) \leq \frac{\frac{1}{3}}{\frac{1}{3}+\epsilon} \right\},$$

and $C_{\Omega\setminus E} = \{q^*\}$, with $q^*(\beta) = 1$, then $f \succ_E g$ and of course $g \sim_{\Omega\setminus E} g'$. Together with $fEg \prec f'Eg'$, this constitutes a violation of Dynamic Consistency.

One interpretation of this pattern of preferences builds on the observation that the elementary events $\upsilon$ and $\beta$ are "complementary", Intuitively, the DM knows that there are 60 balls that are either yellow or blue, so $g$ represents an unambiguous choice for her. On the other hand, $fEg$ wins if a red or a blue ball is drawn, but the DM may suspect that all non-red balls are actually yellow (except for a "small" fraction $\epsilon$). In this respect, a promise of a favorable outcome if $\upsilon$ or $\beta$ obtain is more attractive to the DM than the promise of the same favorable outcome in case $\rho$ or $\beta$ obtain; in loose but suggestive terminology, $\beta$ is "more complementary" with $\upsilon$ than with $\rho$, so "attaching" $g' = g$ to $f'$ is worth more to the DM than attaching $g$ to $f$.

However, this complementarity plays no role in the evaluation of the acts $f$ and $g$ conditional on $E$, because $\beta$ is known not to have occurred. Now the consideration that the urn is *known* to contain 30 red balls looms larger, which justifies the ranking $f \succ_E g$.

This discussion suggests that Dynamic Consistency should hold in circumstances where these complementarities are less likely to play a decisive role. Consider replacing $f'$ and $g'$ with a constant act $x$ such that $f \succcurlyeq_E x$ and $g(\beta) \succcurlyeq x$ in this example. Since ambiguity plays no role in the evaluation of constant acts, there is no reason to expect that the *un*conditional evaluation be "boosted" by complementarities, as was the case for $f'Eg'$ in the preceding example; thus, one should expect that $fEg \succcurlyeq x$.

More interestingly, suppose that $f \preccurlyeq_E x$ and $g(\beta) \preccurlyeq x$. Now the evaluation of the composite act $fEg$ *may* receive a boost from complementarities, so one might be worried that $fEg \succ x$. However, the following argument suggests that this may not be the case. It is easiest to focus on the Ellsberg example. As was noted above, the elementary events $\upsilon$ and $\beta$ *are* complementary, and $fEg$ yields a favorable outcome (the prize 10) if one of these states obtains, *the prize $x$ is by assumption at least as good as this favorable outcome.* Indeed, in the example, $x$ must be at least as good as $fEg$ in every state, so one must expect that $fEg \preccurlyeq x$.

In other decision settings, there may be states $\omega \in E$ such that $f(\omega) \succ x$; but, if it is assumed that $g(\omega) \preccurlyeq x$ for all $\omega \notin E$ [and not just that $g \preccurlyeq_{\Omega \setminus E} x$], then the states that receive a "boost" due to complementarities are guaranteed *not* to be those where $fEg$ delivers the most desirable prizes; in fact, at these states, $fEg$ delivers prizes that are at most as good as $x$.

In light of these considerations, it is not unreasonable to expect that $fEg \preccurlyeq x$ under such circumstances. This is the content of the following axiom.

**Axiom 4.2 (Constant-Act Dynamic Consistency)** *For every non-empty $E \subset \Omega$, and for all prizes $x \in X$ and acts $f \in F^a_E$ and $f' \in F^a_{\Omega \setminus E}$:*

$$f \succcurlyeq_E x, \quad f'(\omega) \succcurlyeq x \; \forall \omega \notin E \implies fEf' \succcurlyeq_\Omega x;$$
$$f \preccurlyeq_E x, \quad f'(\omega) \preccurlyeq x \; \forall \omega \in E^c \implies fEf' \preccurlyeq_\Omega x,$$

*and the unconditional preference is strict if the conditional preference is.*

The main result of this section shows that Axiom 4.2 characterizes full Bayesian updating for acts.

**Proposition 5** *Under Assumptions 4.1 and 4.2, the following statements are equivalent:*

1. *Every $\succcurlyeq_E$ ($E$ non-empty) is complete and transitive on $F^a_E$, and Axiom 4.2 holds;*
2. *For every non-empty $E \subset \Omega$, and for every $f, g \in F^a_E$,*

$$f \succcurlyeq_E g \quad \text{if and only if} \quad \min_{q \in C} \int_E u \circ f \, q(d\omega|E) \geq \min_{q \in C} \int_E u \circ g \, q(d\omega|E) \tag{4}$$

   *where $u$ and $C$ are as in Assumption 4.1.*

**Proof:** [TO BE ADDED: the proof is contained in the working paper Siniscalchi [15].] ■

Notice that Proposition 5 does *not* provide a recursive MEU representation of preferences over acts (unlike the main result in Epstein and Schneider [4]). However, together with Assumption 4.3, it guarantees that *every decision problem can be solved using the BI algorithm, invoking the conditional MEU representation in Eq. 4.*

The straightforward details are as follows. First, BI must be modified to employ the conditional MEU representation; notice that this is possible because, in each step and at each iteration, BI relies solely upon comparisons of *plans*, which can be reduced to acts by Assumption 4.3.

**Algorithm 2 (Backward Induction for MEU preferences)** Let $E \subset \Omega$ be nonempty and $f \in F_E$. Let $u$ and $C$ be as in Assumption 4.1.

1. Find a history $h$ and a reducible action $a \in A_f(h)$. IF there is none, then STOP and RETURN the set $\{f\}$.

2. Denote the non-singleton elements of $\mathcal{F}_f(h,a)$ by $E_1, \ldots, E_n$. Inductively construct an act $f' \in F_E$ as follows. Let $f^0 = f$ and, for $m = 1, \ldots, n$,

   (a) $h^m = [h,(a,E_m)]$,

   (b) $M_m$ is the collection of $b' \in A_f(h^m)$ such that

   $$\forall a' \in A_f(h^m),\ \min_{q\in C} \int_{E_m} u \circ \hat{f}(h^m,b')dq(\omega|E_m) \geq \min_{q\in C} \int_{E_m} u \circ \hat{f}(h^m,a')dq(\omega|E_m),$$

   (c) $f^m = \{f(h^m,b') : b' \in M_m\}_{h^m} f^{m-1}$.

   The required act $f'$ is then $f^n$.

3. Define the set

   $$M_{f'}(h,a) = \Big\{g \in F_{f'}(h,a) : \forall g' \in F_{f'}(h,a),\ \min_{q\in C} \int_E u \circ \hat{g}dq(\omega|E) \geq \min_{q\in C} \int_E u \circ \hat{g}'dq(\omega|E)\Big\}.$$

4. IF $h = \emptyset$, then STOP and RETURN the set $M_{f'}(h,a) = M_{f'}(\emptyset,*)$.

5. Let $f'' = \Big\{\{f(h,a') : a' \in A_f(h) \setminus \{a\}\} \cup M_{f'}(h,a)\Big\}_h f$.

6. REPLACE $f$ WITH $f''$ and GO TO step 1.

The counterpart to Theorem 4 can then be stated.

**Theorem 6** *Consider a system of preferences $\{\succcurlyeq_E\}_{\emptyset\neq E\subset\Omega}$ and suppose that Assumptions 4.1, 4.2 and 4.3 hold. Then the following statements are equivalent.*

1. *Every $\succcurlyeq_E$ is complete and transitive, and Axioms 3.8, 3.9 and 4.2 hold;*
2. *For every non-empty $E \subset \Omega$ and every $f \in F_E$, if Algorithm 2 outputs the collection $\{g_1, \ldots, g_n\}$ of plans, then $f \sim_E g_1 \sim_E \ldots \sim_E g_n \sim x$, where*

$$u(x) = \min_{q\in C} \int_E u \circ g_1 q(d\omega|E).$$

**Proof:** (1) $\Rightarrow$ (2): since every $\succcurlyeq_E$ is complete and transitive, Assumptions 4.1, 4.2 and 4.3 hold, and Axiom 4.2 hold, Algorithm 2 coincides with Algorithm 1, i.e. BI; in particular, the inequality in Step 2 is $\hat{f}(h^m, b') \succcurlyeq_{E_m} \hat{f}(h^m, a')$, which by Assumption 4.3 is equivalent to $f(h^m, b') \succcurlyeq_{E_m} f(h^m, a')$; similarly, the definition of $M_{f'}(h, a)$ in Step 3 coincides with the preference-based one provided in the preceding section. Theorem 4 then implies that $f \sim_E g$ whenever $g$ is an ouptut of Algorithm 2; Assumption 4.3 implies that $g \sim_E \hat{g}$, and Proposition 5 yields the expression for $u(x)$.

(2) $\Rightarrow$ (1): note first that, if $f$ is an act, then Algorithm 2 terminates in Step 1 with output $f$ itself; the assumption in (2) then implies that $f \sim x$, where $u(x)$ is the conditional MEU evaluation of $f$. Thus, conditional preferences restricted to $F_E^a$ are satisfy (2) in Proposition 5, so they are complete and transitive; furthermore, Axiom 4.2 holds. Now, by Assumption 4.3, Algorithm 2 coincides with BI, so Theorem 4 ensures that Axioms 3.8 and 3.9 also hold. ■

# References


[1] Daniel Ellsberg. Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75:643–669, 1961.

[2] Larry Epstein. An axiomatic model of non-bayesian updating. University of Rochester, March 2004.

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

[4] Larry G. Epstein and Martin Schneider. Recursive multiple-priors. *Journal of Economic Theory*, 113:1–31, 2001.

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

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

[7] Edi Karni and David Schmeidler. Atemporal dynamic consistency and expectded utility theory. *Journal of Economic Theory*, 54:401–408, 1991.

[8] Peter Klibanoff. Dynamic choice with uncertainty aversion. March 1995. mimeo, Kellogg School of Management.

[9] Peter Klibanoff and Emre Ozdenoren. Subjective recursive expected utility. mimeo, Kellogg School of Management, 2004.

[10] David M. Kreps. A representation theorem for "preference for flexibility". *Econometrica*, 47(3):565–577, May 1979.

[11] David M. Kreps and Evan L. Porteus. Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 46:185–200, 1978.

[12] 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 1986.

[13] Martin J. Osborne and A. Rubinstein. *A Course on Game Theory.* MIT Press, Cambridge, MA, 1994.

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

[15] Marciano Siniscalchi. Bayesian updating for general maxmin expected utility preferences. Discussion Paper n. 1366, CMS-EMS, Northwestern University: `http://kellogg.northwestern.edu/research/math`, 2001.

[16] R.H. Strotz. Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, 23(3):165–180, 1955-1956.