Prerationality in Finite Decision Trees with Chains of Consequence Nodes*

Peter J. Hammond[†]
University of Warwick

Agustín Troccoli Moretti[‡]
UPF and BSE

June 17, 2026
typeset from `timedConseqs_BCN.tex`

Abstract

Given a consequence domain, following an idea due to Vickrey (1964), Hammond (1976) showed that a behaviour rule defined at every decision node in an unrestricted domain of finite decision trees, including continuation decision trees, will have a set of consequences that depends only on the feasible set of consequences if and only if behaviour maximises a complete and transitive weak preference relation. This result was extended in Hammond (1998a, b) to trees that include chance and event nodes, and to preferences that satisfy the independence axiom with risk, and Anscombe and Aumann's (1963) sure thing principle if there is also uncertainty. If preferences are continuous on Marschak triangles, behaviour must maximise the expectation of each utility function in a unique cardinal equivalence class. This previous work followed Raiffa (1968) by limiting attention to trees with consequences that accrue only at a terminal node. Here, for the case of trees without event nodes, we extend these earlier results to trees that admit finite chains of consequence nodes that accrue along any path, for which intertemporal utility can be expressed as the sum of history-dependent temporal utility functions. A subsequent paper considers a further extension allowing menu consequences. `197 words]`

Keywords: *prerationality, consequentialist decision theory, decision trees, consequence nodes, normal form invariance, history-dependent preferences*

JEL Classification: D81, D91, D11, D63, E21, E71, G41

*Earlier versions of this paper were included in Chapter 2 of Agustín's PhD dissertation. They were also presented to internal Economic Theory workshops at both Warwick and UPF, as well as to the SAET 2024 conference in Manchester, and to the IX Hurwicz Workshop in Warsaw in 2024. For their helpful discussion we would like to thank Agustín's thesis adviser Herakles Polemarchakis and his thesis examiners Pablo Beker and Georgios Gerasimou, as well as audience members at the different seminars.

[†]CRETA, Department of Economics, University of Warwick, Coventry CV4 7AL, UK. E-mail: p.j.hammond@warwick.ac.uk

[‡]Departament d'Economia i Empresa, UPF, Jaume I Building, Ramon Trias Fargas, 25-27, 08005 Barcelona and Barcelona School of Economics. E-mail: agustin.troccoli@upf.edu

1 Introduction and Outline

1.1 Motivation

Bayesian rationality has become a standard paradigm in normative decision theory, requiring the maximisation of the subjective expected value of each Bernoulli utility function within a unique cardinal equivalence class of utility functions. Previous work, such as Hammond (1988b, 2022), has characterised Bayesian rationality by examining the consequences of behaviour in an unrestricted domain of finite decision trees, with a consequence attached to the terminal node of each path through the tree. The domain of trees includes all possible continuation subtrees.

The key result of this prior work is that behaviour defined over this domain is prerational (in effect, consequentialist) and suitably continuous if and only if it maximises an underlying Bayesian rational base preference relation defined on the domain of lotteries over consequences at terminal nodes.[1] This establishes a fundamental connection between preferences over risky or uncertain consequences and behaviour in dynamic environments.

This past work considered finite decision trees that have four different kinds of node. Depending on what explains the transition, if there is one, from a given node of the tree to the next immediately succeeding node, these are:

1. *decision nodes*, where the transition is controlled by the decision maker;

2. *chance nodes*, where what Anscombe and Aumann (1963) described as a "roulette lottery" with known or hypothetical but specified probabilities determines the next state;

3. *event nodes*, where an uncertain move to the next state is determined by what Anscombe and Aumann (1963) described as a "horse lottery" without assigned probabilities;[2]

4. *terminal nodes*, which have no successors and determine a final consequence.

[1] Previously Hammond (1988a,b, 1998a) had worked directly with the notion of *consequentialist* behaviour. Later, starting with Hammond (2022), this concept mutated to the notion of *prerationality*. While both notions can be used interchangeably, as we will at times in the present paper, the notion of prerationality emphasises how the assumed restriction on behaviour comes *before* any assumption on neoclassical rationality. Put differently, neoclassical rationality assumes either the existence of a complete and transitive binary relation describing an agent's preference, or a well-behaved choice correspondence that generates it. Our primitive object in this paper, as well as in previous work, is behaviour defined on an unrestricted domain of finite decision trees. This behaviour need not conform, prima facie, to a choice correspondence. Prerationality then implies the existence of both a choice correspondence and an underlying complete and transitive preference base relation which rationalises that correspondence.

[2] In Hammond (1988b), event nodes were referred to as "natural nodes." The revised terminology was suggested by the work of Debreu (1959), where decision trees are described as "event trees."

In particular, this prior work has assumed that a consequence is assigned to a node if and only if it is a terminal node. In this paper, we allow trees in which a consequence can be assigned to non-terminal nodes as well. Thus, we introduce a novel fifth category of node, which we call an *intermediate timed consequence node*.

Like terminal nodes, these nodes do not induce branching. Instead, they are associated with a specified timed consequence that is relevant to the decision maker. Unlike a terminal node, however, an intermediate consequence node does admit a non-empty set of immediate successors, although that set must consist of just one member. By extending the consequentialist decision-theoretic framework developed in Hammond (1988a,b, 2022) to include trees with intermediate timed consequence nodes, our theory can account for some critical yet previously overlooked aspects of dynamic choice, including intermediate consumption and, in a subsequent paper, menu choice.

Regarding these new aspects, note first how, in many economic applications, decision problems are structured in a way which makes it natural to regard the decision maker as accruing timed consequences at intermediate nodes in a decision tree. For instance, in a standard macroeconomic consumption–savings model, at each decision node the DM chooses a consumption level. This counts as an intermediate timed consequence, which in turn determines savings. A succeeding chance or event node then leads to a potentially random wealth level in the next period, thus inducing the continuation subtree in which future consumption choices will be made. Incorporating such intermediate consequence nodes broadens the scope of consequentialist decision theory to accommodate these common economic scenarios. Moreover, it lays the foundations for a subsequent analysis of infinite-horizon decision problems where backward induction methods like those developed in Hammond (1988a,b, 2022) and in this paper are not directly applicable. In these infinite-horizon settings, recursive decision structures emerge naturally, making it essential to include intermediate consequences.

It is natural next to extend the unrestricted domain assumption of the earlier work in order to admit all logically possible finite decision trees that now include nodes of all five kinds. Then, as discussed in Section 5, our results show that the previous characterisation in Hammond (2022) of continuous prerational behaviour as maximizing expected utility still holds in our enriched domain. Each Bernoulli utility function in the relevant cardinal equivalence class, however, no longer has just one terminal consequence as its only argument. Instead, the relevant arguments together form a chain of successive timed consequences which accumulate along a path through the entire tree.

Second, following Koopmans (1964) and Kreps (1979) for the case of deterministic decision trees, as well as Kreps and Porteus (1978) for decision trees that admit chance nodes, decision theory has long recognised that menus can themselves count as consequences. Finite decision trees with menu consequences are the subject of our ensuing paper, Hammond and Troccoli-Moretti (2026).

1.2 Two Early Attempts to Justify Ordinality

The ordinal revolution in demand theory originated in the work of economists like Pareto, Hicks and Allen. Together with an appendix in later editions of the treatise by von Neumann and Morgenstern (1944), which provided a system of axioms for expected utility, these early ideas have been developed into an extensive literature on alternative axiom systems in decision theory.[3] The key axiom of this ordinal revolution, of course, is *ordinality*. This requires that a consumer's demand behaviour can be modelled as the choice of a consumption bundle that maximizes a (complete and transitive) *preference ordering* over a feasible set, such as a budget set, of feasible consumption bundles.

In due course, subsequent work considered whether ordinality is reasonable in more general decision problems. One early noteworthy contribution is the well known "money-pump" argument due to Davidson et al. (1955)(pp. 145–146). This and the later literature on this topic are the subject of, inter alia, Cubitt and Sugden (2001), as well as the monograph by Gustafsson (2022). A second noteworthy contribution is based on the notion of "path independence". This was inspired by the concluding paragraphs of the second edition of Arrow's *Social Choice and Individual Values*. The exploration by Plott (1973) showed, however, that path independence was only necessary, but not sufficient, for ordinality.[4]

1.3 Definition of Normal Form Invariance

The present paper uses a different approach to justifying ordinality which was first used by Hammond (1977). Motivated by work such as that by Strotz (1956) and by Phelps and Pollak (1968), this relies on considering dynamic behaviour in decision trees. Earlier, in their works on intertemporal economics, including macroeconomic theory, economists such as Irving Fisher, Hicks, Arrow, and Debreu in particular considered the implications of an approach which Vickrey (1964) described as "metastatics". This prompted the rather unusual title of Hammond (1977). In a competitive market economy, Vickrey's approach assumes that every economic agent chooses a single trading strategy for all time in order to maximize a single intertemporal preference

[3] Famously, though von Neumann and Morgenstern's system had many axioms, it did not explicitly state the key independence axiom. This was realized following the subsequent discussion by Marschak (1950), Samuelson (1952), and Malinvaud (1952).

[4] For subsequent refinements of Plott's concept which do fully characterize ordinality, see Campbell (1978) for work that focuses on computational procedures and their stability, as well as Bandyopadhyay (1988) for an alternative refined definition of "sequential" path independence.

ordering — or in the case of a producer, the expected present discounted value of an intertemporal profit stream. Thus, it is as if economic agents were participating in a "metastatic" economy where, as in Debreu (1959)'s model of complete markets for all dated contingent commodities, markets need to open only once.

In the specific context of an intertemporal economy, this "metastatic" approach is equivalent to the claim in von Neumann (1928), elaborated later in von Neumann and Morgenstern (1944), that a game in extensive form can be reduced to an associated game in strategic or normal form in which each player has only a single decision to make. This single decision is the choice of a *normal form strategy* specifying what action to take at every possible information set that the player may face anywhere in the whole extensive form game. Then *normal form invariance* is the claim that, instead of having to consider all the details of a game in extensive form, it loses no generality to consider only the strategic form of the game.

It should be noted that modern game theorists usually reject this claim, though only outside two prominent special cases. The first special case arises when the game in extensive form is a strictly competitive game, otherwise commonly known as a two-person zero-sum game. The second case is a single-person game with only one player, which of course is exactly the right setting when trying to justify ordinal preferences for a decision maker.

Indeed, in its textbook treatment of dynamic choice, (Kreps, 2013, p. 149) states

> *"The standard approach to dynamic decision making (and to decision making within a larger dynamic context) is to regard the decision problem as one of choosing, at the outset, an optimal overall* strategy *and then implementing flawlessly."*

Kreps continues by attaching to the standard approach a consequentialist flavour

> *"Optimality is determined according to the* outcome *engendered by the strategy; outcomes are (within the model) what matters to the decision maker (...)".*

The decision trees we consider are related to the "decision-flow diagrams" defined in Raiffa (1968) (p. 10) as finite trees in which a real-valued pecuniary payoff is attached to each terminal node. Unlike Raiffa, however, in a *deterministic decision tree* there are no *chance nodes* at which a lottery determines the next move; instead, all nodes except terminal nodes are *decision nodes* at which the decision maker determines the next move. Also, rather than having a real-valued pecuniary payoff attached to each terminal node n, instead we specify a *consequence domain* in the form of an arbitrary non-empty set Y, and then we attach a specified *terminal consequence* $\eta(n) \in Y$ to each terminal node n.

1.4 Implications of Normal Form Invariance

In Hammond (1977) the concept we are now calling normal form invariance allowed a novel justification of preference maximization. Specifically, consider behaviour in deterministic finite decision trees, where all nodes except terminal nodes are decision nodes. The normal form invariance hypothesis is that behaviour in any such tree has consequences that are identical to those that would result if a normal form decision strategy were used to choose a set of appropriate consequences that is entirely determined by the feasible set of consequences. As will be discussed in Section 4.1, it turns out that behaviour satisfies this hypothesis if and only if the appropriately chosen consequences are those that maximize a (complete and transitive) preference ordering over the feasible set.

Next, subsequent work such as Hammond (1988a, 1998a) initially considered risky decision trees with chance nodes in addition to decision and terminal nodes. In effect, the results of that work showed that normal form invariance could be satisfied if and only if there was a preference ordering over risky consequences that satisfied the well known and contentious independence axiom of expected utility theory. Provided a further continuity axiom due to Jensen (1967) is imposed, the preference ordering over risky consequences must have an expected utility representation.

Further extensions set out in Hammond (1988b, 1998b) considered uncertain decision trees with event nodes in addition to decision, chance, and terminal nodes. Then, at least in the case when the set of possible uncertain states of the world is finite, normal form invariance along with continuity is satisfied if and only if every state of the world can be given a positive subjective probability having the property that the preference ordering over risky consequences must have an expected utility representation.

Though it would be fairly easy to include decision trees with event nodes in our analysis, we have chosen not to do in order to avoid notational complexity.

1.5 Outline

After this introduction, the next Section 2 begins our formal discussion of decision trees which, because they may have chance nodes but lack event nodes, may have risky consequences but no consequences that are uncertain. Compared to earlier work where consequences occur only at terminal nodes, the key novel feature of the decision trees we consider in this paper is that in addition there may be timed consequence nodes that have consequences that are also of concern to the decision maker.

Following these essential preliminaries, Section 2 sets out a formal model of finite decision trees which allow timed consequence nodes. In particular, it is shown formally how any such tree can be reduced to an equivalent decision tree in which the only consequence nodes are terminal nodes, but where each terminal consequence is built up from a sequence of successive timed consequences that arise along the unique path in the decision tree which reaches that terminal node. For this reduction to work, however, the preferences which apply to consequence chains that arise in any continuation decision tree must generally be "history dependent", in the sense that they depend on the part of a consequence chain that precedes the initial node of the continuation decision tree.

Next, Section 3 turns to behaviour rules. In past work such rules were defined so that, at any decision node of a relevant decision tree, they specify a non-empty set of the moves which the decision maker can select at that node. Here we also consider how a behaviour rule also specifies, at any decision node, a non-empty set of path strategies.

Definitions of consequentialism abound in philosophy and economics, and it is not always clear that authors employing the same concept mean the same thing, and how each definition relates to each other. Some intuitive interpretations of consequentialism go along the lines "the structure of a decision tree doesn't matter" or "behaviour is explained solely by its consequences". In the present paper we make a novel formal link between the concept of consequentialism developed by Hammond (1988b), the concept of prerationality developed by Hammond (2022), and the concept of *normal form invariance* developed by von Neumann (1928). This key notion of consequentialist normal form invariance is the subject of Section 4.

After extending the previous discussion of deterministic decision trees to allow trees which include chance nodes, Section 5 is concerned with how prerationality implies the independence axiom of expected utility theory. In particular, introducing a familiar extra continuity axiom due to Jensen (1967) gives a justification for choosing lotteries over consequence chains that maximize the expected value of each Bernoulli utility function in a unique cardinal equivalence class of functions defined on the domain of deterministic consequence chains. The same section also considers the dynamic programming properties of any optimal path strategy.

Finally, Section 6 offers a concluding summary.

2 Decision Trees with Chains of Timed Consequences

2.1 Beyond Terminal Consequentialism

In this section we begin the process of extending the consequentialist decision theory developed in Hammond (1988a,b, 1998a, 2022) to the domain of decision trees that include chains of timed consequence nodes, of which only the last is a terminal node. In the existing literature (in addition to the already cited papers by Hammond, we also have the work by Karni and Schmeidler (1991)) the only nodes in a decision tree that have consequences attached are terminal nodes. This leads to a terminal or *"deathbed"* form of consequentialism in which, for example, an individual's entire life history is evaluated only at a time when life is about to end.

In this paper, we go beyond terminal consequentialism. For example, we might want to accommodate more directly consumers' decision trees in macroeconomics, where consequences are *intertemporal consumption chains*. Accordingly, each **consequence node** will not only have an *elementary consequence* attached to it, but also a (potentially non-terminal) time of occurrence.

2.2 Trees, Succeeding Nodes, and Continuation Subtrees

In mathematics, a **graph** (N, E) is a set N of vertices or **nodes** together with a set $E \subseteq N \times N$ of **edges** $e = (n, n')$ with $n \neq n'$ that consist of linked pairs of nodes. The graph (N, E) is **directed** just in case there is an antisymmetric and complete binary relation $>_{+1}$ of **immediate succession** on E — i.e., for each directed edge $(n, n') \in E$, either $n >_{+1} n'$, or $n' >_{+1} n$, but not both. We also let $n \to n'$ denote the directed edge defined by the ordered pair (n, n'). Then the set of **immediate successors** of n is

$$N^{+1}(n) := \{\, n' \in N : (n, n') \in E \,\} \;=\; \{\, n' \in N : n' >_{+1} n \,\} \;=\; \{\, n' \in N : n \to n' \in E \,\}.$$

A **path** of length $k \in \mathbb{N}$ is a chain $\langle n_j \rangle_{j=1}^k = \langle n_1, n_2, \dots, n_k \rangle$ of k successive nodes in N such that $n_{j+1} >_{+1} n_j$ for all $j = 1, \dots, k-1$. A node $n' \in N$ **ultimately succeeds** n just in case either $n' = n$, or else there is a unique path $n_1 \to n_2 \to \dots \to n_k$ with $k > 1$ such that $n_1 = n$ and $n_k = n'$.

A directed graph (N, E) is a **tree** just in case there exists a unique node $n \in N$ such that, for every other node $n' \in N \setminus \{n\}$, there exists a unique path $n \to \cdots \to n'$. We denote this unique **initial node** of the tree $T = (N, E)$ by $n_0(T) = n$.[5] Given a finite tree $T = (N, E)$ and a node $n \in N$, the **continuation subtree** $T_{\geq n} = (N_{\geq n}, E_{\geq n})$ at n is the unique subtree of T with initial node n in which

$$N_{\geq n} := \{\, n' \in N : n' \text{ ultimately succeeds } n \,\}, \quad \text{and} \quad E_{\geq n} := E \cap (N_{\geq n} \times N_{\geq n}).$$

2.3 Risky Decision Trees with Timed Consequence Nodes

Let Y be any non-empty set representing the elementary **consequence** domain. We denote by $\Delta(Y)$ the set of all probability mass function or simple **roulette lottery** on

[5] Mathematicians and computer scientists often call the initial node of a tree its "root". This is botanically incorrect because almost all trees have entire root systems. Other more appropriate terms include "seed" or "entry point".

Y that are mappings $Y \ni y \mapsto \pi(y) \in [0,1] \subset \mathbb{R}$ for which there exists a finite **support** $\operatorname{supp}(\pi) \subseteq Y$ such that

$$\pi(y) > 0 \iff y \in \operatorname{supp}(\pi) \quad \text{and} \quad \sum_{y \in Y} \pi(y) = \sum_{y \in \operatorname{supp}(\pi)} \pi(y) = 1.$$

Definition 1. *For an elementary consequence domain Y, a* ***finite decision tree with timed consequences*** *T is a tree with a finite set $N(T)$ of nodes such that there exists:*

1. *a non-empty set $N^d(T)$ of* ***decision nodes*** *n, at each of which there exists a non-empty set $N^{+1}(n)$ of immediate successors;*
2. *a non-empty set $N^c(T)$ of* ***chance nodes*** *n at each of which a specified roulette lottery $\pi(\cdot|n) \in \Delta(N^{+1}(n))$ that satisfies $\pi(n'|n) > 0$ for all $n' \in N^{+1}(n)$ determines randomly which node n' will succeed n;*
3. *a non-empty set $N^y(T)$ of* ***timed consequence nodes*** *n such that:*
   - *$\#N^{+1}(n) \in \{0,1\}$ where, if $\#N^{+1}(n) = 0$ and so $N^{+1}(n) = \emptyset$, then n is a* ***terminal node****, whereas if $\#N^{+1}(n) = 1$, then n is an* ***intermediate consequence node****;*
   - *there exists a* ***consequence mapping***

     $$N^y(T) \ni n \mapsto \boldsymbol{y}(n) = (t(n), y(n)) \in \mathbb{R} \times Y$$

     specifying the timed consequence to the DM, in the form of a pair consisting of an elementary consequence $y(n) \in Y$ and a time $t(n) \in \mathbb{R}$ of its occurrence, of reaching any timed consequence node.

Thus, a finite decision tree T with timed consequences has a partition

$$N(T) = N^d(T) \cup N^c(T) \cup N^y(T)$$

of the set of nodes $N(T)$ into the sets:

1. $N^d(T)$ of decision nodes at which the DM makes a move;
2. $N^c(T)$ of chance nodes at which a lottery is played;
3. $N^y(T) = N^t(T) \cup N^i(T)$ of timed consequence nodes that can be:
   - either terminal nodes $n \in N^t(T) = \{n' \in N^y(T) : N_{+1}(n') = \emptyset\}$;
   - or intermediate timed consequence nodes $n \in N^i(T) = N^y(T) \setminus N^t(T)$.

Note that at any consequence node, whether terminal or not, there is no branching. Also, any terminal node is a consequence node.

We denote by $\mathcal{T}(\mathbb{R} \times Y)$ the unrestricted domain of all finite risky decision trees with decision, chance, and timed consequence nodes, where each consequence is an element of $\mathbb{R} \times Y$. This is the general domain studied in the present paper. The domain of terminal consequences considered in previous work can be identified as a structural subdomain of $\mathcal{T}(\mathbb{R} \times Y)$ satisfying the following restrictions. First, consequence nodes must be terminal nodes, so that $N^y(T) = N^t(T)$ (or equivalently, $N^i(T) = \emptyset$). Second, time must play no role. So we can fix any terminal time $t \in \mathbb{R}$ and restrict attention to trees whose consequence nodes are all elements of $\{t\} \times Y$. Then the resulting class

$$\mathcal{T}^{\mathsf{term}}(Y) := \big\{ T \in \mathcal{T}(\mathbb{R} \times Y) : \ N^i(T) = \emptyset, \ \boldsymbol{y}(n) \in \{t\} \times Y \ \forall n \in N^y(T) \big\},$$

of trees can be naturally identified with the earlier domain of trees where all consequence nodes are terminal.[6]

For a decision tree $T \in \mathcal{T}(\mathbb{R}\times Y)$, any path $(n_1, n_2, \ldots, n_k)$ in the tree includes a subsequence of $m \leq k$ consequence nodes, and so a finite sequence or **chain** $\mathbf{y} = (\boldsymbol{y}_j)_{j=1}^m = (t_j, y_j)_{j=1}^m$ of **timed consequences** $(t_j, y_j) \in \mathbb{R} \times Y$ whose successive times satisfy $t_{j+1} > t_j$ for $j = 1, 2, \ldots, m-1$. Given the timed consequence domain $\mathbb{R}\times Y$, for each $\tau \in \mathbb{N}$, let

$$\mathbf{Y}^\tau_< \ := \ \Big\{(t_j, y_j)_{j=1}^\tau \in (\mathbb{R} \times Y)^\tau : \ t_{j+1} > t_j \text{ for all } j = 1, \ldots, \tau - 1\Big\}.$$

denote the domain of all possible timed consequence chains $\mathbf{y} = (t_j, y_j)_{j=1}^\tau$ of length τ. Finally, we define $\mathbf{Y}(\mathbb{R}\times Y) \ := \ \bigcup_{\tau\in\mathbb{N}} \mathbf{Y}^\tau_<$ as the set of all consequence chains of any finite length.

We also regard a history as a finite chain of timed consequence. Define the domain of possible histories as

$$H := \{\varnothing\} \cup \mathbf{Y}(\mathbb{R} \times Y),$$

where $\varnothing$ denotes the empty history. For $h \in H$, let $\bar{t}(h)$ denote the last date appearing in h, with $\bar{t}(\varnothing) = -\infty$.

2.4 Decision Strategies and Path Strategies

Central to any consequentialist decision theory is the requirement that, for every tree T, the theory provides a way of identifying what are feasible consequences for the decision maker in that tree, provided that an appropriate decision is made at each decision node.

Each decision node $n \in N^d(T)$ has a finite set of immediate successors $N^{+1}(n)$. Each of these immediately succeeding nodes corresponds to a possible move that the

[6] Hammond often works directly with $\mathcal{T}^{\mathsf{term}}(\Delta(Y_0))$, so that terminal consequence nodes may carry lotteries over an underlying outcome space Y_0. In the present framework this enrichment is not essential: one can equivalently keep the consequence domain fixed at Y_0 and represent any such terminal randomisation by inserting an explicit (possibly degenerate) chance node immediately before the terminal consequence node.

DM might make at n. Because the tree is finite, the obvious counterpart of a strategy in an extensive form game is a **decision strategy** specifying what move the DM makes at every decision node $n \in N^d(T)$.

A decision strategy must specify what move is made even at decision nodes that are irrelevant because they cannot ever be reached after moves that the decision strategy has specified at previous decision nodes. For this reason, we prefer to consider instead a **path strategy** which specifies what sequence of nodes emerges from following each step of a decision strategy all the way through to the end of a continuation subtree. Indeed, a path strategy takes the form of a complete contingent plan which can be specified by describing, for every decision node that can be reached, both which successor is chosen and also how play continues after each possible realisation at any chance node.

A more formal definition of a path strategy must involve recursion. To describe this requires some new notation. First, for any node n, let $\langle n \rangle$ denote the singleton node chain whose only member is n. Then, given any $k \in \mathbb{N}$ and any chain $\mathbf{p} = \langle m_1, \dots, m_k \rangle$ consisting of k consecutive nodes, let

$$\langle n \rangle \parallel \mathbf{p} := \langle n, m_1, \dots, m_k \rangle \tag{1}$$

denote the **concatenation** of n with $\mathbf{p}$. That is, $\langle n \rangle \parallel \mathbf{p}$ is the chain with $k+1$ nodes that results by letting the chain $\langle n \rangle$ precede the chain $\mathbf{p}$.

Now a path strategy in a finite tree T is a complete contingent plan which specifies, for each decision node $n \in N^d(T)$ and associated continuation decision tree $T_{\geq n}$, what random path through $T_{\geq n}$ emerges from the decision process. In particular, a path strategy prescribes exactly one immediate successor at every decision node that can be reached; it also specifies what is to be done *after every possible realisation* at any encountered chance node. Crucially, a path strategy lists only moves chosen at decision nodes. At each timed consequence node there is only one forced move. And at any chance node there is also no choice; all that happens is that a specified roulette lottery expands a relevant path into a randomized path whose continuations are then fully specified.

Definition 2 (Path strategy)**.** *Given a finite decision tree $T \in \mathcal{T}(\mathbb{R} \times Y)$, a* path strategy *on T is a family $\{\mathbf{p}(T_{\geq n})\}_{n \in N^d(T)}$ of contingent paths $\mathbf{p}(T_{\geq n})$, following each decision node $n \in N^d(T)$, that pass through the continuation subtree $T_{\geq n}$. Moreover, the family $\{\mathbf{p}(T_{\geq n})\}_{n \in N^d(T)}$ of contingent paths is generated by the following recursive* *rule:*

$$\mathbf{p}(T_{\geq n}) = \begin{cases} \langle\,\rangle, & \textit{if } n \in N^t(T), \quad \textit{so there are no further decisions;} \\ \mathbf{p}(T_{\geq n'}), & \textit{if } n \in N^y(T) \textit{ with } N^{+1}(n) = \{n'\}; \\ \langle n^* \rangle, & \textit{if } n \in N^d(T),\ n^* \in N^{+1}(n) \textit{ and } N^d(T_{\geq n^*}) = \emptyset; \\ \langle n^* \rangle \parallel \mathbf{p}(T_{\geq n^*}), & \textit{if } n \in N^d(T),\ n^* \in N^{+1}(n) \textit{ and } N^d(T_{\geq n^*}) \neq \emptyset; \\ \mathbb{E}_{\pi(\cdot|n)} \mathbf{p}(T_{\geq n'}), & \textit{if } n \textit{ is a chance node with kernel } \pi(\cdot \mid n). \end{cases}$$

where $\mathbb{E}_{\pi(\cdot|n)}\mathbf{p}(T_{\geq n'})$ *denotes the expectation* $\sum_{n' \in N^{+1}(n)} \pi(n' \mid n)\ \mathbf{p}(T_{\geq n'})$ *of the path strategy* $\mathbf{p}(T_{\geq n'})$*.*

Remark 1. We adopt the mixture preservation rule requiring that, for any node n, for any finite family $\{\mathbf{q}_i\}_{i\in I}$ of paths, and for any set $\{\alpha_i\}_{i\in I}$ of *convex weights* with $\alpha_i \geq 0$ for all $i \in I$ and $\sum_{i\in I} \alpha_i = 1$, one has

$$\langle n \rangle \parallel \left(\sum\nolimits_{i\in I} \alpha_i\, \mathbf{q}_i \right) \ := \ \sum\nolimits_{i\in I} \alpha_i \left(\langle n \rangle \| \mathbf{q}_i \right) \tag{2}$$

This definition ensures that the recursion in Definition 2 is well defined when concatenation applies to chains of consequence nodes that appear along a random branch of the tree that appears downstream.

Definition 3 (Set of path strategies)**.** *For any decision tree* T*, let* $\mathbf{P}(T)$ *denote the set of all path strategies* $\{\mathbf{p}(T_{\geq n})\}_{n\in N^d(T)}$ *that satisfy the recursion specified in Definition 2.*

The recursive construction in Definition 2 is intended to capture formally the intuitive notion of a contingent plan: a strategy should prescribe actions only along branches that can in fact be realised, and it should specify what happens after every possible chance realisation. The following proposition confirms that the recursive definition indeed has these properties. It ensures that the syntactic object $\mathbf{p}(T)$ corresponds to the semantic idea of an on-path, fully specified contingent plan.

Proposition 1. *Every path strategy* $\{\mathbf{p}(T_{\geq n})\}_{n\in N^d(T)} \in \mathbf{P}(T)$ *satisfies:*

1. Reachability: *it prescribes a move only at decision nodes that can be reached under earlier prescribed choices and chance realisations;*

2. Inclusiveness at chance nodes: *whenever a chance node* n *is reached with positive probability under* $\{\mathbf{p}(T_{\geq n})\}_{n\in N^d(T)}$*, the path strategy unfolds into a mixture that includes the full continuation* $\mathbf{p}(T_{\geq n'})$ *for every* $n' \in N^{+1}(n)$*. In particular, no successor of* n *is omitted from the induced contingent plan.*

Proposition 1 establishes that the recursive construction expands into precisely the kind of object one intuitively understands as a contingent plan. The path strategy never prescribes actions at nodes that cannot be reached under its own prior choices. Also, whenever chance intervenes, it expands into all possible branches that emanate after the roulette lottery at the relevant chance node has been resolved, so each branch carries within itself the full continuation plan in each case. Thus the recursively defined path strategy provides a complete specification of behaviour along every possible history that may be realised, without omitting any feasible contingencies, or introducing any prescriptions concerning behaviour at off-path decision nodes.

Example. Consider a decision tree T in which the set of immediate successors of the initial decision node n_0 satisfies $N^{+1}(n_0) = \{n^c\}$ for a unique chance node n^c. Suppose that two chance edges $n^c \to n_1$ and $n^c \to n_2$ emanate from n^c and lead to decision nodes $n_1, n_2 \in N^{+1}(n^c)$ with respective probabilities α and $1 - \alpha$. Finally, assume that the immediate successors of these decision nodes in $N^{+1}(n_1) \cup N^{+1}(n_2)$ are all terminal nodes.

Consider a decision strategy for which the chosen move at n_1 is $m_1 \in N^{+1}(n_1)$ and at n_2 it is $m_2 \in N^{+1}(n_2)$. Then, using (2) for the second equality, one has

$$\mathbf{p}(T) \;=\; \langle n^c \rangle \,\|\, \big(\alpha \,\langle m_1 \rangle + (1-\alpha)\,\langle m_2 \rangle\big) \;=\; \alpha\,\big(\langle n^c \rangle \| \langle m_1 \rangle\big) \;+\; (1-\alpha)\,\big(\langle n^c \rangle \| \langle m_2 \rangle\big)$$

This explicitly shows the random path as a mixture over the two contingencies created by the chance node, while recording exactly one move at each decision node that is reachable by following $\mathbf{p}(T)$ along each branch. ♣

Remark 2. Observe that if $N^c(T) = \emptyset$, so that the tree T contains no chance nodes, then every path strategy specifies a single path through each continuation subtree of the tree.

2.5 Path Strategies and their Consequences

Fix a finite decision tree $T \in \mathcal{T}(\mathbb{R} \times Y)$ and a path strategy $\{\mathbf{p}(T_{\geq n})\}_{n \in N^d(T)} \in \mathbf{P}(T)$. Executing the path strategy from any node n determines a unique continuation path strategy for the continuation subtree $T_{\geq n}$. Indeed:

1. at each decision node, a single successor path strategy is selected;
2. at each chance node, the path strategy unfolds into a lottery over all succeeding subpaths with their assigned probabilities;
3. at each timed consequence node that is not a terminal node, the timed consequence is experienced and play continues;
4. at each terminal node, the process stops.

For every continuation subtree $T_{\geq n}$, the path strategy therefore induces a *unique* random path in the form of a lottery over the finite set of terminating paths of successive nodes in the decision tree. The random path is obtained by specifying the different generated paths along all reachable branches and then weighting them by the relevant chance probabilities. Formally, this lottery can be constructed using a backward recursion that mirrors that used to construct $\mathbf{p}$.

In order to determine what consequences result from a path strategy, we will use two prefixing concatenation operators. The first prefixing operator $\|$, which applies to chains of arbitrary nodes, is the one defined by Equation (1) that simply prepends one node to a succeeding node chain. The second prefixing operator $\circ$, which applies only to chains of timed consequence nodes, prepends a timed consequence to a subsequent consequence chain. Concretely, given a consequence chain $\mathbf{y} \in \mathbf{Y}(\mathbb{R} \times Y)$ and a timed consequence $y(n) \in \mathbb{R} \times Y$ at consequence node n, the concatenation $y(n) \circ \mathbf{y} \in \mathbf{Y}(\mathbb{R} \times Y)$ is the consequence chain obtained by prepending $y(n)$ to $\mathbf{y}$.

Thus, the two operators $\|$ and $\circ$ are formally distinct but play analogous roles: operator $\|$ prefixes nodes in the strategy representation, whereas $\circ$ prefixes consequences in the induced consequence chain.[7]

Theorem 2.1 (Chain consequence mapping)**.** *For any decision tree $T \in \mathcal{T}(\mathbb{R} \times Y)$, there exists a unique* ***chain consequence mapping***

$$\mathbf{P}(T) \ni \mathbf{p} \mapsto \boldsymbol{\eta}(T, \mathbf{p}) \in \Delta\big(\mathbf{Y}(\mathbb{R} \times Y)\big), \tag{3}$$

such that for every path strategy $\mathbf{p} \in \mathbf{P}(T)$ and every node n, the chain consequence $\boldsymbol{\eta}(T_{\geq n}, \mathbf{p})$ is the unique solution of the following backward recursion:

1. *$\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) = \delta_{\langle \mathbf{y}(n) \rangle}$ in case $n \in N^t(T)$ with timed consequence $\mathbf{y}(n)$;*

2. *$\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) = \mathbf{y}(n) \circ \boldsymbol{\eta}(T_{\geq n'}, \mathbf{p})$ in case $n \in N^i(T)$ whose unique immediate successor satisfies $N^{+1}(n) = \{n'\}$;*

3. *$\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) = \boldsymbol{\eta}(T_{\geq n^*}, \mathbf{p})$ if $n \in N^d(T)$ and n^* is uniquely chosen by path strategy $\mathbf{p}$ at n;*

4. *$\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) = \sum_{n' \in N^{+1}(n)} \pi(n'|n)\, \boldsymbol{\eta}(T_{\geq n'}, \mathbf{p})$ if $n \in N^c(T)$ with probability kernel given by $\pi(\cdot|n) \in \Delta^0\left(N^{+1}(n)\right)$.*

Moreover, $\boldsymbol{\eta}(T, \mathbf{p}) = \boldsymbol{\eta}(T_{\geq n_0(T)}, \mathbf{p})$ and the same characterisation holds for any continuation subtree $T_{\geq n}$.

[7] Formally, $\circ$ maps $(\mathbf{y}, y(n))$ to a consequence chain $\mathbf{y} \circ y(n)$, whereas $\|$ maps $(\langle n \rangle, \mathbf{p})$ to a node chain $\langle n \rangle \| \mathbf{p}$.

Theorem 2.1 provides a precise link between behaviour and consequences. It links a path strategy to the lottery over consequence chains that is generated when the strategy is executed. The support of this lottery consists exactly of those chains that can arise under the strategy itself. Along any realised history, the strategy follows the unique successor it prescribes at each decision node that is reached, while chance nodes generate the only source of randomisation through their kernels. Thus every chain receiving positive probability corresponds to a path that is consistent with the prescribed behaviour and with the probabilistic structure of the tree, with timed consequence nodes contributing only by prefixing their dated consequences.

Example. Let $n_0 \in N^d(T)$ and suppose choosing path $\mathbf{p}$ at n_0 leads to a chance node $n^c \in N^c(T)$. At n^c the tree allows a random move: either to n_1 with probability α; or to n_2 with probability $1-\alpha$. Assume that $n_1, n_2 \in N^d(T)$, and that their chosen successors under $\mathbf{p}$ are terminal nodes $n_1^t, n_2^t \in N^t(T)$ with respective consequences $y_1, y_2 \in Y$. In addition, suppose a timed intermediate consequence node n^m with consequence y^m lies on the path between n^c and n_1, whereas no such node lies on the path between n^c and n_2. Then the recursion specified in part 4 of Theorem 2.1 results in

$$\boldsymbol{\eta}(T, \mathbf{p}) = \alpha\, \delta_{\langle y^m, y_1 \rangle} + (1-\alpha)\, \delta_{\langle y_2 \rangle}.$$

♣

2.6 Feasible Sets of Consequence Chains

So far we have established that, for a fixed tree T, every path strategy $\mathbf{p} \in \mathbf{P}(T)$ induces, via the mapping $\boldsymbol{\eta}$, a unique lottery over consequence chains. It is therefore natural to ask, for the same tree T, which lotteries over consequence chains can be attained by some admissible path strategy. This leads to the notion of the **feasible set** $F(T)$ of random consequence chains at T. It is the collection of all lotteries over consequence chains which arise from the finitely many path strategies in T that are available to the decision maker.

Definition 4 (Feasible set of random consequence chains)**.** *For any decision tree $T \in \mathcal{T}(\mathbb{R} \times Y)$, the **feasible consequence set** is the range*

$$F(T) \;:=\; \big\{\, \boldsymbol{\eta}(T, \mathbf{p}) \;:\; \mathbf{p} \in \mathbf{P}(T) \,\big\} \;\subseteq\; \Delta\big(\mathbf{Y}(\mathbb{R} \times Y)\big) \tag{4}$$

of the consequence chain mapping defined by (3).

In words, the feasible set $F(T)$ is the range of the consequence mapping $\boldsymbol{\eta}$ as $\mathbf{p}$ varies over $\mathbf{P}(T)$. The set $F(T)$ collects every possible lottery over consequence chains that can arise from executing some possible path strategy in the tree T starting at the initial node $n_0(T)$.

For any lottery $\boldsymbol{\lambda} \in \Delta\big(\mathbf{Y}(\mathbb{R} \times Y)\big)$ whose support consists of chains of timed consequences that occur strictly after time $t(n)$, define

$$\boldsymbol{y}(n) \circ \boldsymbol{\lambda} := \sum\nolimits_{\mathbf{y} \in \operatorname{supp}(\boldsymbol{\lambda})} \boldsymbol{\lambda}(\mathbf{y})\, \delta_{(\boldsymbol{y}(n), \mathbf{y})} \;\in\; \Delta\big(\mathbf{Y}(\mathbb{R} \times Y)\big)$$

That is, at an intermediate consequence node n, the prepending operator attaches before the timed consequence chain $\mathbf{y}$ the additional timed consequence $\boldsymbol{y}(n)$ in order to form the lottery over expanded consequence chains $(\boldsymbol{y}(n), \mathbf{y})$ that is now feasible conditional on having reached node n.

Then, for any finite set $F \subset \Delta\big(\mathbf{Y}(\mathbb{R} \times Y)\big)$ of lotteries, write

$$\boldsymbol{y}(n) \circ F := \{\, \boldsymbol{y}(n) \circ \boldsymbol{\lambda} \;:\; \boldsymbol{\lambda} \in F \,\}$$

Then, for the finite list $(\alpha_i)_{i \in I}$ of weights, with $\alpha_i \geq 0$ for all $i \in I$ and $\sum_i \alpha_i = 1$, and of corresponding sets $F_i \subseteq \Delta\big(\mathbf{Y}(\mathbb{R} \times Y)\big)$, define the mixture set

$$\sum\nolimits_{i \in I} \alpha_i F_i := \Big\{ \sum\nolimits_{i \in I} \alpha_i \boldsymbol{\lambda}_i \;:\; \boldsymbol{\lambda}_i \in F_i \text{ for all } i \in I \Big\}.$$

Theorem 2.2. *Given any finite decision tree $T \in \mathcal{T}(\mathbb{R} \times Y)$, define for each node $n \in N(T)$ the feasible set*

$$F(T_{\geq n}) := \big\{ \boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) \;:\; \mathbf{p} \in \mathbf{P}(T_{\geq n}) \big\} \;\subseteq\; \Delta\big(\mathbf{Y}(\mathbb{R} \times Y)\big).$$

Then the collection $\{F(T_{\geq n}) : n \in N(T)\}$ is the unique family of subsets of $\Delta\big(\mathbf{Y}(\mathbb{R} \times Y)\big)$ that satisfies the backward recurrence relation

$$F(T_{\geq n}) = \begin{cases} \{\delta_{\boldsymbol{y}(n)}\} & \text{if } n \in N^t(T), \\ \boldsymbol{y}(n) \circ F(T_{\geq n'}) & \text{if } n \in N^i(T) \text{ and } N^{+1}(n) = \{n'\}, \\ \bigcup_{n' \in N^{+1}(n)} F(T_{\geq n'}) & \text{if } n \in N^d(T), \\ \sum_{n' \in N^{+1}(n)} \pi(n' \mid n)\, F(T_{\geq n'}) & \text{if } n \in N^c(T). \end{cases}$$

In particular,

$$F(T) = F(T_{\geq n_0(T)}).$$

Now Definition 4 of the feasible set $F(T)$ is conceptually natural: it is simply the range of the consequence mapping as $\mathbf{p}$ varies over $\mathbf{P}(T)$. Taken on its own, however, this characterisation is abstract. It does not explain how to determine $F(T_{\geq n})$ for a given continuation subtree, nor does it provide a practical way to compute feasible sets node by node. Theorem 2.2 fills this gap. It replaces the abstract range definition by one based on a unique corresponding backward recursion that can be calculated directly within the tree, thereby yielding the feasible set $F(T_{\geq n})$ at every node n. We note that if $N^i(T) = \emptyset$, then the domain $\mathcal{T}(\mathbb{R} \times Y)$ collapses to the terminal consequentialism domain $\mathcal{T}(Y)$ considered in Hammond (1988a,b, 1998a, 2022), and the recursion reduces accordingly.

2.7 Consequence Chains as Histories

The feasible set $F(T_{\geq n})$ constructed in Theorem 2.2 of Section 2.6 is generated entirely by the possible decisions and chance moves that remain available in the continuation subtree $T_{\geq n}$ whose initial node is n. Timed consequences that have already accrued before n form a fixed prefix which cannot be varied by decisions within $T_{\geq n}$. Nevertheless this prefix can still matter for the full consequence chain generated by any continuation from n.

Definition 5. *Fix a finite risky decision tree $T \in \mathcal{T}(\mathbb{R} \times Y)$. For any node $n \in N(T)$, let $h(n) \in H$ denote the history at n. This is the finite chain of timed consequences attached to the timed consequence nodes on the unique path from the initial node $n_0(T)$ to, but not including, n, listed in their path order. If no timed consequence node strictly precedes n, then $h(n) = \varnothing$.*

Equivalently, if $m_1, \ldots, m_{k(n)}$ are the timed consequence nodes on that path, ordered from $n_0(T)$ towards n, then $h(n) := \big(\boldsymbol{y}(m_1), \ldots, \boldsymbol{y}(m_{k(n)})\big) \in H$, where $\boldsymbol{y}(m_j) = (t(m_j), y(m_j))$ for each $j = 1, \ldots, k(n)$. The number $k(n)$ is the consequence depth of n. In particular, $h(n_0(T)) = \varnothing$.

If $n \in N^i(T)$ is an intermediate timed consequence node with unique successor n', then $h(n') = (h(n); \boldsymbol{y}(n))$. Thus the timed consequence at n is appended to the history before the successor n' is reached. Conversely, when computing continuation consequences by backward recursion, $\boldsymbol{y}(n)$ is prepended to the consequence chains generated by the successor subtree $T_{\geq n'}$.

For any $h \in H$, let $\mathbf{Y}_{>h}(\mathbb{R} \times Y)$ denote the set of all finite timed consequence chains whose first date, when the chain is non-empty, is strictly greater than $\bar{t}(h)$. For any lottery $\boldsymbol{\lambda} \in \Delta\big(\mathbf{Y}_{>h}(\mathbb{R} \times Y)\big)$, define the history-prefixed lottery

$$h \circ \boldsymbol{\lambda} := \sum_{\mathbf{y} \in \operatorname{supp}(\boldsymbol{\lambda})} \boldsymbol{\lambda}(\mathbf{y})\, \delta_{(h;\mathbf{y})} \in \Delta\big(\mathbf{Y}(\mathbb{R} \times Y)\big),$$

with the convention that $\varnothing \circ \boldsymbol{\lambda} = \boldsymbol{\lambda}$. For any finite set $F \subseteq \Delta(\mathbf{Y}_{>h}(\mathbb{R} \times Y))$, write $h \circ F := \{h \circ \boldsymbol{\lambda} : \boldsymbol{\lambda} \in F\}$. At any node $n \in N(T)$, the continuation feasible set $F(T_{\geq n})$ is the history-free set generated by the continuation tree. The corresponding history-prefixed feasible set is

$$F_{h(n)}(T_{\geq n}) := h(n) \circ F(T_{\geq n}).$$

All lotteries in $F_{h(n)}(T_{\geq n})$ share the same deterministic prefix $h(n)$. The only choice-dependent part is the continuation lottery generated from n onward. More generally, suppose a tree T is faced after a prior history $h \in H$. For any node $n \in N(T)$, the accrued history before reaching n is $(h; h(n))$, and the associated history-prefixed feasible set is

$$F_{(h;h(n))}(T_{\geq n}) := (h; h(n)) \circ F(T_{\geq n}).$$

For each $h \in H$, let

$$\mathcal{D}_h := \Big\{T \in \mathcal{T}(\mathbb{R} \times Y) : n_0(T) \in N^d(T) \text{ and } F(T) \subseteq \Delta\big(\mathbf{Y}_{>h}(\mathbb{R} \times Y)\big)\Big\},$$

denote the domain of decision trees that begin with a decision node, and whose feasible consequence chains, following the history h, are lotteries over the domain $\mathbf{Y}_{>h}(\mathbb{R} \times Y$ of consequence chains that was defined in Section 2.3. Then let

$$\mathcal{D} := \Big\{(h, T) : h \in H,\ T \in \mathcal{D}_h\Big\}$$

The condition $n_0(T) \in N^d(T)$ ensures that behaviour after h selects admissible immediate successors at the initial node of T. The restriction $F(T) \subseteq \Delta\big(\mathbf{Y}_{>h}(\mathbb{R} \times Y)\big)$ ensures that T generates only continuation lotteries whose timed consequence chains can follow the history h. Moreover, whenever $(h, T) \in \mathcal{D}$ and $n \in N^d(T)$, the continuation subtree $T_{\geq n}$ is admissible after the accrued history $(h; h(n))$, so

$$\big((h; h(n)), T_{\geq n}\big) \in \mathcal{D}.$$

Thus the same global behaviour rule can be evaluated at later decision nodes by replacing the fixed history h with the accrued history $(h; h(n))$.

3 Behaviour Rules and Their Consequences

3.1 Defining a Behaviour Rule

A behaviour rule specifies which immediate moves are admissible at decision nodes. As in classical decision and game theory, behaviour may be multi-valued. This allows for the possibility that more than one immediate move is admissible at a decision node.

In earlier work, a behaviour rule was defined as a correspondence specifying, at each decision node n of a tree T, a non-empty set of admissible immediate successors. There is, however, some redundancy in specifying behaviour at all decision nodes of T from the perspective of the initial node. If a later decision node n is actually reached, the decision maker then faces the continuation subtree $T_{\geq n}$, after the history accumulated before n. Thus actual behaviour at n should be described by applying the same rule to the history–tree pair $(h, T_{\geq n}$ consisting of that accrued history followed by the continuation tree starting at n. For this reason, we define behaviour on the domain $\mathcal{D}$ introduced in Section 2.7. Recall that $(h, T) \in \mathcal{D}$ means that T is a history-compatible tree faced after h, with $n_0(T) \in N^d(T)$.

Definition 6. *A behaviour rule is a correspondence*

$$\mathcal{D} \ni (h, T) \mapsto \beta_h(T) \in 2^{N^{+1}(n_0(T))} \setminus \{\emptyset\}.$$

Thus, for each admissible pair $(h, T) \in \mathcal{D}$, the set $\beta_h(T) \subseteq N^{+1}(n_0(T))$ is the non-empty set of immediate successors of the initial decision node $n_0(T)$ that are admissible after history h.

The notation $\beta_h(T)$ is shorthand for $\beta(h, T)$. Thus β_h should not be read as an arbitrary behaviour rule assigned independently to history h. The primitive object is one global correspondence β, whose h-section is evaluated at the continuation tree T. Consequently, if $(h, T) \in \mathcal{D}$ and $n \in N^d(T)$, behaviour at n inside the original tree is described by applying the same rule to the pair $\big((h; h(n)), T_{\geq n}\big) \in \mathcal{D}$. That is, the admissible immediate successors at n are the elements of $\beta_{(h;h(n))}(T_{\geq n}) \subseteq N^{+1}(n)$. At the initial node $n_0(T)$, since $h(n_0(T)) = \varnothing$, this reduces to $\beta_h(T)$.

Thus, given any pair (h, T) in the domain $\mathcal{D}$, a behaviour rule determines a **behaviour set** $\beta_h(T)$ as the non-empty subset of all those decisions or moves in $N_{+1}(n_0(T))$ that are normatively acceptable or *admissible* at the initial node $n_0(T)$ of the tree T.

3.2 Why Behaviour Rules *are* Dynamically Consistent

Consider a DM who started off at the initial node $n_0(T)$ of the decision tree $T \in \mathcal{T}(\mathbb{R}{\times}Y)$, but whose decisions combined with chance events have subsequently led to the decision node $n \in N^d(T)$ of T. At this subsequent decision node n, the DM faces the continuation subtree $T_{\geq n}$ whose initial node is n. Possible *actual* behaviour at the decision node n is, by definition, specified by the non-empty behaviour set $\beta(T_{\geq n}) \subseteq N_{+1}(n)$. This choice set overrules any previous plan or intention to:

1. include unacceptable nodes $\tilde{n} \in N_{+1}(n) \setminus \beta(T_{\geq n})$;

2. exclude acceptable nodes $\tilde{n} \in \beta(T_{\geq n})$.

It follows that, whereas dynamic inconsistency is possible for preferences (Strotz (1956)), or for plans or intentions (Bratman (1987)), it is *impossible* for *actual* behaviour. This is because, in any continuation decision tree that is so far unreached, plans are at best unrealized commitments. Note too that if plans or intentions are dynamically inconsistent, then actual behaviour cannot realize all these plans. We defer a formal treatment of dynamic consistency until Section 5.2.

3.3 The Consequences of a Behaviour Rule

Starting from an admissible history–tree pair $(h, T) \in \mathcal{D}$, consistently following the behaviour rule β restricts which path strategies through T are admissible. In this section we characterise the subset $\Phi_h^\beta(T) \subseteq F(T)$ of continuation lotteries over timed consequence chains that can result from following β, given the chain consequence mapping $\boldsymbol{\eta}$ specified by (3) in Theorem 2.1.

Definition 7 (β-admissible path strategies)**.** *Fix $(h, T) \in \mathcal{D}$ and a behaviour rule β. A path strategy $\mathbf{p} \in \mathbf{P}(T)$ is* β-admissible *from (h, T) if, for every decision node $n \in N^d(T)$ reached with strictly positive probability under $\mathbf{p}$, the immediate successor selected by $\mathbf{p}$ at n, denoted by $\mathbf{p}(n) \in N^{+1}(n)$, satisfies*

$$\mathbf{p}(n) \in \beta_{(h;h(n))}(T_{\geq n}).$$

Let $\mathbf{P}_h^\beta(T) \subseteq \mathbf{P}(T)$ denote the set of all β-admissible path strategies from (h, T).

The definition applies the same global behaviour rule at every decision node that can be reached. The initial history is h. If decision node n is later reached inside T, the history relevant for behaviour at n is the accrued history $(h; h(n))$, while the remaining decision problem is the continuation tree $T_{\geq n}$.

Definition 8 (β-consequence set)**.** *For each $(h, T) \in \mathcal{D}$, define the consequence set generated by β from (h, T) as*

$$\Phi_h^\beta(T) := \left\{ \boldsymbol{\eta}(T, \mathbf{p}) : \mathbf{p} \in \mathbf{P}_h^\beta(T) \right\} \subseteq F(T).$$

When full consequence chains after the already realised history h are needed, we write

$$h \circ \Phi_h^\beta(T) := \left\{ h \circ \boldsymbol{\lambda} : \boldsymbol{\lambda} \in \Phi_h^\beta(T) \right\} \subseteq h \circ F(T).$$

For the recursion below, we extend the notation $\mathbf{P}_h^\beta(T)$ to continuation subtrees whose initial node need not be a decision node. Fix $(h, T) \in \mathcal{D}$. For any node $n \in N(T)$, a path strategy $\mathbf{p} \in \mathbf{P}(T_{\geq n})$ belongs to $\mathbf{P}_{(h;h(n))}^\beta(T_{\geq n})$ if, for every decision node $m \in N^d(T_{\geq n})$ reached with strictly positive probability under $\mathbf{p}$, the immediate successor selected by $\mathbf{p}$ at m satisfies

$$\mathbf{p}(m) \in \beta_{(h;h(m))}(T_{\geq m}).$$

Thus the admissibility restriction is imposed only at decision nodes, using the accrued history from the original tree.

Theorem 3.1. *Fix $(h,T) \in \mathcal{D}$ and a behaviour rule β. For each node $n \in N(T)$, define*

$$\Phi^{\beta}_{(h;h(n))}(T_{\geq n}) := \Big\{ \boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) : \mathbf{p} \in \mathbf{P}^{\beta}_{(h;h(n))}(T_{\geq n}) \Big\} \subseteq F(T_{\geq n}).$$

Then the collection

$$\Big\{ \Phi^{\beta}_{(h;h(n))}(T_{\geq n}) : n \in N(T) \Big\}$$

is the unique family of subsets of $\Delta\big(\mathbf{Y}(\mathbb{R} \times Y)\big)$ satisfying the backward recurrence

$$\Phi^{\beta}_{(h;h(n))}(T_{\geq n}) = \begin{cases} \{\delta_{\boldsymbol{y}(n)}\} & \text{if } n \in N^t(T), \\ \boldsymbol{y}(n) \circ \Phi^{\beta}_{(h;h(n'))}(T_{\geq n'}) & \text{if } n \in N^i(T) \text{ and } N^{+1}(n) = \{n'\}, \\ \displaystyle\bigcup_{n' \in \beta_{(h;h(n))}(T_{\geq n})} \Phi^{\beta}_{(h;h(n'))}(T_{\geq n'}) & \text{if } n \in N^d(T), \\ \displaystyle\sum_{n' \in N^{+1}(n)} \pi(n'|n)\, \Phi^{\beta}_{(h;h(n'))}(T_{\geq n'}) & \text{if } n \in N^c(T). \end{cases}$$

In particular, since $h(n_0(T)) = \varnothing$, one has

$$\Phi^{\beta}_{h}(T) = \Phi^{\beta}_{(h;h(n_0(T)))}(T_{\geq n_0(T)}).$$

Theorem 3.1 plays the same role for the correspondence $(h,T) \mapsto \Phi^{\beta}_{h}(T)$ that describes the consequences of β-admissible behaviour as Theorem 2.2 does for the correspondence $T \mapsto F(T)$ that describes the set of consequences that are feasible. Whereas Definition 8 characterises $\Phi^{\beta}_{h}(T)$ abstractly as the range of the consequence mapping restricted to $\mathbf{P}^{\beta}_{h}(T)$, the theorem provides a backward recursion that allows these sets to be computed node by node. The definition of the recursion for $\Phi^{\beta}_{h}(T)$ differs from that for F only at decision nodes, where the union is taken only over successors that are β-admissible given the accrued history $(h; h(n))$ rather than over all available successors. Since the admissible successors are always among the feasible successors, the admissible consequence sets are contained in the corresponding feasible sets.

Corollary 1 (β-consequences are feasible)**.** *For every $(h,T) \in \mathcal{D}$ and every node $n \in N(T)$, one has*

$$\Phi^{\beta}_{(h;h(n))}(T_{\geq n}) \subseteq F(T_{\geq n}).$$

In particular, $\Phi^{\beta}_{h}(T) \subseteq F(T)$.

4 Normal Form Invariance

4.1 Normal Form Reduction of Risky Decision Trees

Following von Neumann (1928) and von Neumann and Morgenstern (1944), in the normal or reduced form of a game each player has only a single opportunity to make a move. Moreover, this single move must select a complete plan of action specifying what to do at every information set specified in the extensive form of the game. In a single–person decision tree $T \in \mathcal{T}(\mathbb{R} \times Y)$, such a complete plan is represented by a contingent terminating chain $\mathbf{p} \in \mathbf{P}(T)$ of nodes constructed recursively as in Section 2.4.

In a deterministic decision tree, our definitions imply that there is no distinction between a path strategy and the path that it realises. Indeed, a path strategy then specifies, for each prior history, each decision node, and each ensuing continuation decision tree, *precisely* the unique terminating chain of consequence nodes that results from the choices made at each decision node that the decision maker reaches.

In a decision tree that has chance nodes, the random result of the roulette lottery at each chance node replaces the deterministic outcome of a path strategy with a lottery over terminating chains of successive consequence nodes which are specified for each continuation subtree that starts at a decision node. The consequence mapping $\boldsymbol{\eta}(T, \mathbf{p})$ therefore already attaches to each path strategy $\mathbf{p} \in \mathbf{P}(T)$ its induced lottery over terminating consequence chains.

To work with a reduced (normal) form, we collapse the extensive form of a decision tree into a single static choice at the initial node, in the same spirit as the reduction to normal form of an extensive form game. Concretely, we replace any finite decision tree T by a uniquely defined tree T^{R} in normal form where the initial node is the only decision node, at which the DM selects a path strategy $\mathbf{p} \in \mathbf{P}(T)$. The terminal consequences attached to this path strategy form a lottery over terminating consequence chains through T that are induced by $\mathbf{p}$, which are specified by $\boldsymbol{\eta}(T, \mathbf{p})$. In this reduction, all sequential aspects of decision making are absorbed into the initial choice of the path strategy $\mathbf{p}$, and all intermediate resolutions of uncertainty are folded into a single lottery over terminal nodes, which are consequence nodes. This motivates the following formal definition.

Definition 9 (Normal form reduction)**.** *Let $T \in \mathcal{T}(\mathbb{R} \times Y)$ be any risky finite decision tree with timed consequences. Define the* ***normal form reduction*** *as the decision tree $T^{\mathrm{R}} \in \mathcal{T}(\mathbb{R} \times Y)$ with:*

(i) a single initial node $n_0^{\mathrm{R}}(T) := n_0(T^{\mathrm{R}})$;

(ii) one terminal node $n_{\mathbf{p}}^t \in N^t(T^{\mathrm{R}})$ for each path strategy $\mathbf{p} \in \mathbf{P}(T)$;

(iii) *set of edges* $E^{\mathrm{R}} := \{n_0^{\mathrm{R}}(T) \to n_{\mathbf{p}}^t : \mathbf{p} \in \mathbf{P}(T)\}$*, which all end in a terminal node;*

(iv) *for each path strategy* $\mathbf{p} \in \mathbf{P}(T)$ *and corresponding terminal node* $n_{\mathbf{p}}^t$*, a corresponding terminal lottery given by*

$$\eta(T^{\mathrm{R}}, n_{\mathbf{p}}^t) := \boldsymbol{\eta}(T, \mathbf{p}) \in \Delta\big(\mathbf{Y}(\mathbb{R}\times Y)\big).$$

Observe that T^{R} lies in the terminal consequentialism subdomain $\Delta\big(\mathbf{Y}(\mathbb{R}\times Y)\big)$ of decision trees with consequence space. Indeed, the only decision node of T^{R} is its initial node n_0^{R}, whose immediate successors are all terminal nodes, so the only intermediate node is the initial node n_0^{R}. Compared to Hammond's original formulation where the only consequence nodes are terminal nodes, what changes is only the terminal consequence space: Hammond works with lotteries over Y, whereas here terminal nodes are labelled by lotteries over timed-consequence chains, namely elements of $\Delta\big(\mathbf{Y}(\mathbb{R}\times Y)\big)$.

Definition 10 (Consequentialist normal form reduction)**.** *Let* $T \in \mathcal{T}(\mathbb{R}\times Y)$ *be any risky finite decision tree with timed consequences, and let* T^{R} *be its normal form reduction as specified in Definition 9. Define the tree* $T^{\mathrm{C}} \in \mathcal{T}\big(\Delta\big(\mathbf{Y}(\mathbb{R}\times Y)\big)\big)$ *to be the* ***consequentialist normal form reduction*** *of* T *with:*

(i) *a single initial node* $n_0^{\mathrm{C}}(T) := n_0(T^{\mathrm{C}})$*;*

(ii) *one terminal node* $n_{\boldsymbol{\lambda}}^t \in N^t(T^{\mathrm{C}})$ *for each lottery* $\boldsymbol{\lambda} \in F(T)$*;*

(iii) *one terminal consequence lottery given by* $\eta(T^{\mathrm{C}}, n_{\boldsymbol{\lambda}}^t) = \boldsymbol{\lambda}$ *for each* $\boldsymbol{\lambda} \in F(T)$*.*

A minor difference between the normal form reduction T^{R} of a decision tree $T \in \mathcal{T}(\mathbb{R} \times Y)$ and the consequentialist normal form reduction T^{C} of the same tree lies in the labelling of the terminal nodes, each of which is an immediate successor in T^{C} of the initial node $n_0^{\mathrm{C}}(T)$. In the tree T^{R} these nodes are indexed by path strategies $\mathbf{p} \in \mathbf{P}(T)$, whereas in T^{C} they are indexed directly by consequence lotteries $\boldsymbol{\lambda} \in F(T)$.

A more substantive difference arises whenever there are two distinct paths $\mathbf{p}, \mathbf{p}' \in \mathbf{P}(T)$ satisfying $\boldsymbol{\eta}(T, \mathbf{p}) = \boldsymbol{\eta}(T, \mathbf{p}')$. In this case one has $\#N_{+1}\big(n_0^{\mathrm{C}}(T)\big) = \#F(T) < \#N_{+1}\big(n_0^{\mathrm{R}}(T)\big) = \#\mathbf{P}(T)$. Thus T^{C} merges dynamically distinct path strategies which induce the same consequence lottery. So in the end the difference between the two reductions does not matter; they both generate exactly the same set of lotteries over consequence chains as are feasible in the original tree.

Proposition 2. *Let* $T \in \mathcal{T}(\mathbb{R}\times Y)$ *be any risky finite decision tree with timed consequences. Let* T^{R} *and* T^{C} *be, respectively, its normal-form reduction and consequentialist normal-form reduction, as specified in Definitions 9 and 10. Then*

$$F\big(T^{\mathrm{R}}\big) \;=\; F\big(T^{\mathrm{C}}\big) \;=\; F(T) \;\subseteq\; \Delta\big(\mathbf{Y}(\mathbb{R}\times Y)\big).$$

In view of the discussion in Section 2.3 on the embedding of terminal consequentialism within $\mathcal{T}(\mathbb{R} \times Y)$, the feasible consequence sets $F\big(T^{\mathrm{R}}\big)$ and $F\big(T^{\mathrm{C}}\big)$ are well defined. In particular, even when terminal nodes are labelled by consequence lotteries, this can be understood either by allowing a mixture-space consequence domain or, equivalently, by representing the relevant randomisation explicitly via chance nodes immediately upstream. In either interpretation, the feasible sets for T^{R} and T^{C} are constructed by the same backward recurrence as in Theorem 2.2 which, when specialised to trees where all consequence nodes are terminal, coincides with the recurrence relations in Hammond (1988a,b, 2022).

4.2 Consequentially Normal Form Invariant Behaviour

The process of normal form reduction compresses an original "dynamic" decision tree with multiple decision nodes into a single "static" decision tree with only one initial decision node. Moreover, it does so while preserving the set of feasible lotteries over consequence chains. In this section we study the behavioural implication of this observation. We do so following the guiding principle that behaviour should be determined solely by the consequentialist content of the problem — namely, the finite menu $F(T)$ of lotteries over consequence chains — rather than by the particular dynamic presentation of the tree. In short: once $F(T)$ is fixed, the dynamic structure of the decision problem should be irrelevant to the admissible set of consequences that behaviour in the tree makes possible.

The next definition formalises this irrelevance.

Definition 11. *A behaviour rule β on $\mathcal{D}$ is consequentially normal-form invariant just in case, for every history $h \in H$ and all trees $T, \widetilde{T} \in \mathcal{D}_h$, one has*

$$F(T) = F(\widetilde{T}) \quad \Longrightarrow \quad \Phi_h^\beta(T) = \Phi_h^\beta(\widetilde{T}).$$

Consequential normal-form invariance is imposed history by history. For a fixed history h, two continuation trees $T, \widetilde{T} \in \mathcal{D}_h$ that generate the same feasible set $F(T) = F(\widetilde{T})$ must generate the same admissible consequence set under β. The comparison is made after the same already realised history h, so the invariance requirement does not compare behaviour across different histories.

Definition 12. *Let $\mathcal{F}_{\setminus\emptyset}$ denote the family of all non-empty finite subsets or menus of $\Delta\big(\mathbf{Y}(\mathbb{R}\times Y)\big)$. Then*

- *A **choice correspondence** on $\Delta\big(\mathbf{Y}(\mathbb{R}\times Y)\big)$ is a mapping $\mathcal{F}_{\setminus\emptyset} \ni F \mapsto \mathcal{C}(F) \in \mathcal{F}_{\setminus\emptyset}$ satisfying the restriction $\mathcal{C}(F) \subseteq F$ for all F in $\mathcal{F}_{\setminus\emptyset}$.*
- *Given any such $\mathcal{C}$, the associated **base preference relation** $\succsim$ on $\Delta\big(\mathbf{Y}(\mathbb{R}\times Y)\big)$ is the revealed preference relation defined, for all pairs of lotteries $\boldsymbol{\lambda}, \boldsymbol{\lambda}' \in \Delta\big(\mathbf{Y}(\mathbb{R}\times Y)\big)$,*

by

$$\boldsymbol{\lambda} \succsim \boldsymbol{\lambda}' \iff \boldsymbol{\lambda} \in \mathcal{C}\big(\{\boldsymbol{\lambda}, \boldsymbol{\lambda}'\}\big) \tag{5}$$

Theorem 4.1. *A behaviour rule β on $\mathcal{D}$ is consequentially normal-form invariant if and only if, for each history $h \in H$, there exists a choice correspondence $\mathcal{C}_h^\beta$ on finite subsets of $\Delta(\mathbf{Y}_{>h}(\mathbb{R} \times Y))$ such that, for every tree T in the domain $\mathcal{D}_h$, one has*

$$\Phi_h^\beta(T) = \mathcal{C}_h^\beta(F(T)). \tag{6}$$

For each $h \in H$, the induced base preference relation $\succsim_h^\beta$ is defined by

$$\boldsymbol{\lambda} \succsim_h^\beta \boldsymbol{\mu} \quad \iff \quad \boldsymbol{\lambda} \in \mathcal{C}_h^\beta\big(\{\boldsymbol{\lambda}, \boldsymbol{\mu}\}\big).$$

In earlier work Hammond (1988b, 1998a, 2022), condition (6) was taken as the *definition* of consequentialist behaviour. By contrast, Theorem 4.1 shows that this condition is not an assumption but rather a *characterisation* that follows from the primitive concept of consequentialist normal-form invariance. In particular, a behaviour rule β is consequentially normal-form invariant if and only if, for each history $h \in H$, there exists a choice correspondence $\mathcal{C}_h^\beta$ on finite menus of lotteries over timed consequence chains such that, for every tree $T \in \mathcal{D}_h$, behaviour after history h coincides with the choice from the menu $F(T)$.

Proposition 2 then links this invariance to von Neumann's normal-form perspective: every tree T has the same feasible set of lotteries over timed consequence chains as its normal-form and consequentialist normal-form reductions. Combining the proposition with Theorem 4.1 yields, for every history $h \in H$ and every tree T in the domain $\mathcal{D}_h$, the equation

$$\Phi_h^\beta(T) = \Phi_h^\beta\big(T^{\mathrm{R}}\big) = \Phi_h^\beta\big(T^{\mathrm{C}}\big).$$

Thus, the dictum often attributed to (Kreps, 1988, p. 189–190) —*"dynamic choice equals the static choice of a strategy"*— is not merely a heuristic: under normal-form invariance it becomes *formally equivalent* to Hammond's notion of consequentialism, history by history. In this way, normal-form invariance provides a foundational bridge between the classical representation of dynamic problems and the purely consequentialist content encoded by the feasible set $F(T)$.

5 Prerational Behaviour and Expected Utility

5.1 Definition of Prerationality

The preceding section associates with any consequentially normal form invariant behaviour rule β a history-indexed family $\{C_h^\beta\}_{h \in H}$ of choice correspondences, and so a history-indexed family $\{\succsim_h^\beta\}_{h\in H}$ of base preference relations that depend on choices from pair sets. For each fixed history $h \in H$, the relation $\succsim_h^\beta$ ranks continuation lotteries over chains of timed consequences that can follow h. The following definition applies the notion of prerationality history by history.

Definition 13. *Fix $h \in H$. The binary relation $\succsim_h$ on the domain $\Delta\big(\mathbf{Y}_{>h}(\mathbb{R}\times Y)\big)$ of lotteries over consequence chains is* prerational *just in case there exists a consequentially normal-form invariant behaviour rule β on $\mathcal{D}$ such that $\succsim_h = \succsim_h^\beta$, where $\succsim_h^\beta$ is the base preference relation induced by $\mathcal{C}_h^\beta$.*

Definition 14. *Fix $h \in H$. A binary relation $\succsim_h$ on $\Delta\big(\mathbf{Y}_{>h}(\mathbb{R}\times Y)\big)$ is* mixture continuous *just in case, for all triples $\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu} \in \Delta\big(\mathbf{Y}_{>h}(\mathbb{R}\times Y)\big)$ such that $\boldsymbol{\lambda} \succ_h \boldsymbol{\mu}$ and $\boldsymbol{\mu} \succ_h \boldsymbol{\nu}$, the two sets*

$$\{\alpha \in [0,1] : \alpha\boldsymbol{\lambda} + (1-\alpha)\boldsymbol{\nu} \succsim_h \boldsymbol{\mu}\} \quad \textit{and} \quad \{\alpha \in [0,1] : \boldsymbol{\mu} \succsim_h \alpha\boldsymbol{\lambda} + (1-\alpha)\boldsymbol{\nu}\}$$

are both closed in $[0,1]$.

Theorem 5.1. *Fix $h \in H$. A base preference relation $\succsim_h$ on $\Delta\big(\mathbf{Y}_{>h}(\mathbb{R}\times Y)\big)$ is prerational and mixture continuous if and only if there exists a history-dependent Bernoulli utility index $\mathbf{Y}_{>h}(\mathbb{R}\times Y) \ni \mathbf{y} \mapsto u_h(\mathbf{y}) \in \mathbb{R}$ such that the expected utility functional $\Delta\big(\mathbf{Y}_{>h}(\mathbb{R}\times Y)\big) \ni \boldsymbol{\lambda} \mapsto U_h(\boldsymbol{\lambda}) \in \mathbb{R}$ satisfies, for all $\boldsymbol{\lambda}, \boldsymbol{\lambda}' \in \Delta(\mathbf{Y}_{>h}(\mathbb{R}\times Y))$, the equivalence*

$$\boldsymbol{\lambda} \succsim_h \boldsymbol{\lambda}' \iff U_h(\boldsymbol{\lambda}) = \sum_{\mathbf{y}\in \mathrm{supp}(\boldsymbol{\lambda})} \boldsymbol{\lambda}(\mathbf{y}) u_h(\mathbf{y}) \geq \sum_{\mathbf{y}\in \mathrm{supp}(\boldsymbol{\lambda}')} \boldsymbol{\lambda}'(\mathbf{y}) u_h(\mathbf{y}) = U_h(\boldsymbol{\lambda}') \tag{7}$$

Moreover, the Bernoulli utility index $\mathbf{y} \mapsto u_h(\mathbf{y})$ belongs to a unique cardinal equivalence class

$$\mathsf{A}_{u_h} := \{v_h \in \mathbb{R}^{\mathbf{Y}_{>h}(\mathbb{R}\times Y)} : \exists (a,b) \in \mathbb{R}_{>0}\times\mathbb{R} \textit{ such that } u_h = a v_h + b\}$$

The proof proceeds in two distinct steps. First, fix any history $h \in H$ and consider any risky decision tree $T \in \mathcal{D}_h$. We operate with its consequentialist normal-form reduction T^{C}. By Proposition 2, this reduction preserves the feasible set of consequences, which is

$$F\big(T^{\mathrm{C}}\big) = F(T) \subseteq \Delta\big(\mathbf{Y}_{>h}(\mathbb{R}\times Y)\big).$$

Moreover, because the behaviour rule β is consequentially normal-form invariant, the induced admissible consequence set $\Phi_h^\beta(T)$ after history h depends only on $F(T)$. So restricting attention to the smallest consequence domain that includes the terminal consequences of any consequentialist normal form tree T^{C} leaves both feasible consequences and behaviour after h unchanged. The dynamic structure of T can therefore be collapsed without loss of consequential content.

Second, on the terminal domain with consequence space $\mathbf{Y}_{>h}(\mathbb{R} \times Y)$, Hammond's representation theorem applies. A base preference relation $\succsim_h$ on $\Delta(\mathbf{Y}_{>h}(\mathbb{R} \times Y))$ is prerational and mixture continuous if and only if there exists a Bernoulli index $u_h : \mathbf{Y}_{>h}(\mathbb{R} \times Y) \to \mathbb{R}$ such that $\succsim_h$ is represented by the expected utility function

$$\Delta(\mathbf{Y}_{>h}(\mathbb{R} \times Y)) \ni \boldsymbol{\lambda} \mapsto U_h(\boldsymbol{\lambda}) = \sum_{\mathbf{y} \in \operatorname{supp}(\boldsymbol{\lambda})} \boldsymbol{\lambda}(\mathbf{y})\, u_h(\mathbf{y}).$$

Because T and T^{C} share the same feasible set $F(\cdot)$ and the same invariant behaviour after history h, the same expected utility function $\boldsymbol{\lambda} \mapsto U_h(\boldsymbol{\lambda})$ also represents the original base preference relation $\succsim_h$ on $\Delta\big(\mathbf{Y}_{>h}(\mathbb{R} \times Y)\big)$.

5.2 Prerationality and Dynamic Consistency

Suppose that a behaviour rule β is defined on the doamin $\mathcal{D}$ and is consequentially normal-form invariant. Then Theorem 5.1 shows that, for each fixed history $h \in H$, the induced base preference relation $\succsim_h^\beta$ is mixture continuous if and only if it admits an expected utility representation. In other words, once a history h and a continuation tree $T \in \mathcal{D}_h$ are given, the dynamic choice problem reduces to the static choice of a path strategy $\mathbf{p} \in \mathbf{P}(T)$, with each path strategy inducing a continuation lottery $\boldsymbol{\eta}(T, \mathbf{p}) \in \Delta(\mathbf{Y}_{>h}(\mathbb{R} \times Y))$ over timed consequence chains. Prerationality and mixture continuity then guarantee the existence of a Bernoulli utility index u_h on $\mathbf{Y}_{>h}(\mathbb{R} \times Y)$ such that the strategies prescribed by β maximize expected utility after history h.

Because β is defined on pairs (h, T), it induces not only the base relation at the initial history h, but also a family of local, history-dependent base preference relations at all the decision nodes of any tree T in the domain $\mathcal{D}$. Fix any pair $(h, T) \in \mathcal{D}$. For each decision node $n \in N^d(T)$, the accrued history before reaching n is $h_n := (h; h(n))$. Since $(h_n, T_{\geq n}) \in \mathcal{D}$, the same global behaviour rule β induces a base preference relation $\succsim_{h_n}^\beta$ on the full continuation-lottery space $\Delta\big(\mathbf{Y}_{>h_n}(\mathbb{R} \times Y)\big)$ even though, in the continuation subtree $T_{\geq n}$, only the feasible set $F(T_{\geq n})$ is available. Equivalently, if full consequence chains after the accrued history are used, the relevant feasible set is the history-prefixed set $F_{h_n}(T_{\geq n}) = h_n \circ F(T_{\geq n})$. The resulting family

$$\big(\succsim_{h_n}^\beta \big)_{n \in N^d(T)} \tag{8}$$

provides a complete explanation for the decision maker's behaviour throughout the tree. This is the family fow which we can now define dynamic consistency.

It is important that this family is not understood as an arbitrary collection of preference relations. The construction first fixes one behaviour rule β on $\mathcal{D}$. The family in (8) is then derived by evaluating that same rule at the successive pairs $(h_n, T_{\geq n})$. Without this common behavioural source, there would be no reason to expect the relations at different histories to satisfy any dynamic consistency restriction.

Given two decision nodes $n, n' \in N^d(T)$, say that n' is a *proximate successor* decision node of n just in case if n' succeeds n and the unique path from n to n' contains no decision node other than n and n'. Let $\mathbf{y}$ denote the deterministic chain of timed consequences attached to the timed consequence nodes lying strictly between n and n'. Then

$$h(n') = (h(n); \mathbf{y}) \quad \text{and} \quad h_{n'} = (h_n; \mathbf{y})$$

The prepending notation introduced in Section 2.7 applies in particular to the deterministic segment $\mathbf{y}$. Thus, for any $\boldsymbol{\lambda} \in \Delta(\mathbf{Y}_{>h_{n'}}(\mathbb{R} \times Y))$, the lottery $\mathbf{y} \circ \boldsymbol{\lambda}$ belongs to $\Delta(\mathbf{Y}_{>h_n}(\mathbb{R} \times Y))$.

Definition 15. *For a fixed consequentially normal-form invariant behaviour rule β on the domain $\mathcal{D}$, the family $\left(\succsim^{\beta}_{h_n} \right)_{n \in N^d(T)}$ induced on a tree T after history h is* dynamically consistent *just in case, for any pair of decision nodes $n, n' \in N^d(T)$ such that n' is a proximate successor decision node of n with $h_{n'} = (h_n; \mathbf{y})$, for all $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \Delta\big(\mathbf{Y}_{>h_{n'}}(\mathbb{R} \times Y)\big)$, one has*

$$\mathbf{y} \circ \boldsymbol{\lambda} \succsim^{\beta}_{h_n} \mathbf{y} \circ \boldsymbol{\mu} \quad \Longleftrightarrow \quad \boldsymbol{\lambda} \succsim^{\beta}_{h_{n'}} \boldsymbol{\mu}.$$

That is, the ranking at the later history $h_{n'}$ coincides with the ranking induced at the earlier history h_n after the deterministic segment $\mathbf{y}$ has been fixed. Dynamic consistency therefore compares two rankings of the same tail lotteries: (i) directly after $h_{n'}$; and (ii) after prefixing $\mathbf{y}$ and then evaluating them from h_n.

Next, we show that prerationality of the base preference relation induced at the initial node propagates to prerationality of the base preference relations induced at all continuation subtrees rooted at decision nodes.

Proposition 3 (Propagation of prerationality)**.** *Let β be a consequentially normal-form invariant behaviour rule on $\mathcal{D}$. Fix $(h, T) \in \mathcal{D}$, and write $h_n = (h; h(n))$ for each decision node $n \in N^d(T)$. Then, for every $n \in N^d(T)$, the induced base preference relation $\succsim^{\beta}_{h_n}$ on $\Delta\big(\mathbf{Y}_{>h_n}(\mathbb{R} \times Y)\big)$ is prerational.*

Proposition 3 shows that, if β is consequentially normal-form invariant on $\mathcal{D}$, then on each continuation subtree $T_{\geq n}$, evaluated after its accrued history h_n, there is an inherited prerational base preference relation $\succsim^{\beta}_{h_n}$. Dynamic consistency within a tree therefore concerns the coherence of the family induced by the same rule β across different decision nodes, rather than being an implication of prerationality at the initial node n_0 alone.

Proposition 4 (Dynamic consistency of induced base preferences)**.** *Let β be a consequentially normal-form invariant behaviour rule on $\mathcal{D}$. For every $(h, T) \in \mathcal{D}$, writing $h_n = (h; h(n))$, the induced family $\left(\succsim^{\beta}_{h_n} \right)_{n \in N^d(T)}$ is dynamically consistent.*

Theorem 5.1 provides a static characterisation of base preference relations. It states that, for each fixed history h, the base preference relation $\succsim_h$ is prerational and mixture continuous if and only if it admits an expected utility representation on $\Delta(\mathbf{Y}_{>h}(\mathbb{R}\times Y))$. In particular, prerationality in Definition 13 is a property of a single relation $\succsim_h$, and refers to the existence of a consequentially normal-form invariant behaviour rule that induces it after history h.

By contrast, Propositions 3 and 4 operate at the level of a fixed behaviour rule β on $\mathcal{D}$. Proposition 3 shows that if β is consequentially normal-form invariant, then every continuation base preference relation $\succsim^{\beta}_{h_n}$ induced by β satisfies the prerationality hypothesis of Theorem 5.1. Thus, once mixture continuity is imposed, Theorem 5.1 can be applied history by history to obtain expected utility representations for the entire induced family.

Theorem 5.1 therefore does not by itself impose dynamic consistency. Dynamic consistency arises only after one fixes a single behaviour rule β, evaluates that same rule at each history-tree pair $(h_n, T_{\geq n})$, and studies the coherence of the continuation preferences that it induces.

5.3 Intertemporal Aggregation of Utility Functions

The preceding subsection shows that a single consequentially normal-form invariant behaviour rule β induces a dynamically consistent family of base preference relations along any tree. We now ask how the expected-utility representations of these relations are linked across successive decision histories.

Fix $(h,T)\in\mathcal{D}$, and write $h_n := (h; h(n))$ for the accrued history before a decision node $n\in N^d(T)$. If n' is a proximate successor decision node of n, let $\mathbf{y}$ denote the deterministic chain of timed consequences attached to the timed consequence nodes lying strictly between n and n'. Then $h_{n'} = (h_n; \mathbf{y})$. For any tail consequence chain $\mathbf{s}\in\mathbf{Y}_{>h_{n'}}(\mathbb{R}\times Y)$, we write $(\mathbf{y},\mathbf{s})$ for the concatenation of the realised segment $\mathbf{y}$ with $\mathbf{s}$. When $\succsim^{\beta}_{h_n}$ admits an expected-utility representation, let u_{h_n} denote the associated Bernoulli utility index and U_{h_n} the corresponding expected utility functional.

Theorem 5.2 (History-dependent weak intertemporal aggregation)**.** *Let β be a consequentially normal-form invariant behaviour rule on $\mathcal{D}$. Fix $(h,T)\in\mathcal{D}$, and suppose that, for every $n\in N^d(T)$, the induced base preference relation $\succsim^{\beta}_{h_n}$ is mixture continuous and non-trivial. For each $n\in N^d(T)$, let u_{h_n} be a non-constant Bernoulli utility index representing $\succsim^{\beta}_{h_n}$, as in Theorem 5.1. If n' is a proximate successor decision node of n, with $h_{n'} = (h_n;\mathbf{y})$, then there exist unique numbers $v_{h_n}(\mathbf{y})\in\mathbb{R}$ and $\xi_{h_n}(\mathbf{y})>0$ such that, for every $\mathbf{s}\in\mathbf{Y}_{>h_{n'}}(\mathbb{R}\times Y)$, one has*

$$u_{h_n}(\mathbf{y},\mathbf{s}) = v_{h_n}(\mathbf{y}) + \xi_{h_n}(\mathbf{y})u_{h_{n'}}(\mathbf{s}). \tag{9}$$

By the unrestricted-domain assumption, these coefficients extend uniquely to functions

$$\mathbf{Y}_{>h_n}(\mathbb{R}\times Y)\ni \mathbf{y}\mapsto v_{h_n}(\mathbf{y})\in\mathbb{R},\qquad \mathbf{Y}_{>h_n}(\mathbb{R}\times Y)\ni \mathbf{y}\mapsto \xi_{h_n}(\mathbf{y})\in\mathbb{R}_{>0}.$$

After a realised segment $\mathbf{y}$ has been fixed, the earlier preference at h_n induces a preference over the same tail lotteries that are ranked directly at $h_{n'}$. Dynamic consistency identifies these two rankings. Since both are represented by non-trivial expected-utility functionals on the same mixture space, cardinal uniqueness gives the affine aggregator in (9).

The dependence of v_{h_n} and ξ_{h_n} on $\mathbf{y}$ is essential. The continuation utilities for different segments that are free of decision nodes can be measured on different cardinal scales. Thus the theorem does not impose a single history-specific discount factor linking h_n to every successor history. Instead it imposes an affine link only after the realised segment connecting the two histories has been passed. The corresponding lottery version follows immediately. For any $\boldsymbol{\lambda}\in\Delta\big(\mathbf{Y}_{>h_{n'}}(\mathbb{R}\times Y)\big)$, the bridge (9) implies

$$U_{h_n}(\mathbf{y}\circ\boldsymbol{\lambda}) = v_{h_n}(\mathbf{y}) + \xi_{h_n}(\mathbf{y})U_{h_{n'}}(\boldsymbol{\lambda}). \tag{10}$$

This is the form needed for the finite-tree dynamic programme below. Each realised segment in a transition from one decision node to the next carries its own affine conversion of future utility into the current history's units.

5.4 Dynamic Programming

The preceding subsection links the Bernoulli utility indices at adjacent decision histories. Theorem 5.2 applies after the decision-node-free segment between two decision nodes has been realised. The Bellman recursion must therefore evaluate each possible segment separately.

Fix $(h,T)\in\mathcal{D}$, and write $h_n=(h;h(n))$ for each $n\in N^d(T)$. The state at a decision node is the history-tree pair $(h_n,T_{\geq n})$, and the feasible controls are the immediate successors $a\in N^{+1}(n)$.

For $n\in N^d(T)$ and $a\in N^{+1}(n)$, let $\mathcal{L}(n,a)$ denote the finite set of transition links generated after choosing a. A link $\ell\in\mathcal{L}(n,a)$ consists of a path segment starting at a, ending at the first node $m(\ell)\in N^d(T)\cup N^t(T)$ reached along that segment, while containing no decision node before $m(\ell)$. Let $\mathbf{y}(\ell)$ be the timed consequence chain realised along this segment, and let $\alpha_{n,a}(\ell)>0$ be the product of any chance probabilities along it, with product equal to 1 if no chance node lies on the segment. The probabilities satisfy

$$\sum_{\ell\in\mathcal{L}(n,a)}\alpha_{n,a}(\ell)=1.$$

If $m(\ell) \in N^d(T)$, then $h_{m(\ell)} = (h_n; \mathbf{y}(\ell))$. If $m(\ell) \in N^t(T)$, then $\mathbf{y}(\ell)$ is a complete continuation chain after h_n. The set $\mathcal{L}(n,a)$ is only a transition record generated by the tree; it is not a feasible set, and no preference relation is defined on it.

Theorem 5.3 (Bellman equation)**.** *Let β be a consequentially normal-form invariant behaviour rule on $\mathcal{D}$. Fix $(h,T) \in \mathcal{D}$, and suppose that, for every $n \in N^d(T)$, the induced base preference relation $\succsim^{\beta}_{h_n}$ is mixture continuous and non-trivial. Let u_{h_n}, v_{h_n}, and ξ_{h_n} be the functions specified in Theorem 5.2. Then for each decision node $n \in N^d(T)$ there is a unique value $V(n) \in \mathbb{R}$ which, at each $a \in N^{+1}(n)$, satisfies the following Bellman equation. For, define*

$$Q_{h_n}(n,a) := \sum_{\ell \in \mathcal{L}(n,a)} \alpha_{n,a}(\ell) R_{h_n}(\ell), \tag{11}$$

where

$$R_{h_n}(\ell) := \begin{cases} v_{h_n}(\mathbf{y}(\ell)) + \xi_{h_n}(\mathbf{y}(\ell)) V(m(\ell)), & \text{if } m(\ell) \in N^d(T), \\ u_{h_n}(\mathbf{y}(\ell)), & \text{if } m(\ell) \in N^t(T). \end{cases} \tag{12}$$

Then

$$V(n) = \max_{a \in N^{+1}(n)} Q_{h_n}(n,a). \tag{13}$$

Moreover, for every $n \in N^d(T)$,

$$\beta_{h_n}(T_{\geq n}) \subseteq \operatorname*{arg\,max}_{a \in N^{+1}(n)} Q_{h_n}(n,a). \tag{14}$$

The inclusion in (14) is the strongest node-level implication delivered by consequential normal-form invariance alone. Lemma 1 implies that $\Phi^{\beta}_{h_n}(T_{\geq n})$ contains all $\succsim^{\beta}_{h_n}$-maximal feasible continuation lotteries. It does not require $\beta_{h_n}(T_{\geq n})$ to contain every immediate successor that can implement one of those maximal lotteries. A move may therefore maximise $Q_{h_n}(n,\cdot)$ and still be excluded from $\beta_{h_n}(T_{\geq n})$ when the same maximal consequence lottery is already implemented through another immediate successor.

To state the exact consequential content of the Bellman recursion, define, for each $a \in N^{+1}(n)$,

$$F_a(T_{\geq n}) := \{\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) : \mathbf{p} \in \mathbf{P}(T_{\geq n}),\ \mathbf{p}(n) = a\}.$$

Thus $F_a(T_{\geq n})$ is the set of continuation lotteries that can be generated after choosing the immediate successor a, before imposing β-admissibility. Define an equivalence relation on $N^{+1}(n)$ by

$$a \sim_n b \iff F_a(T_{\geq n}) = F_b(T_{\geq n}).$$

Let

$$\bar{N}^{+1}(n) := N^{+1}(n)/\sim_n,$$

and write $[a]_n$ for the equivalence class of a. For $[a]_n \in \bar{N}^{+1}(n)$, define

$$Q_{h_n}(n, [a]_n) := Q_{h_n}(n, a),$$

which is well defined because all representatives of $[a]_n$ generate the same feasible continuation lotteries. Finally, define

$$\widehat{\beta}_{h_n}(T_{\geq n}) := \left\{ [a]_n \in \bar{N}^{+1}(n) : F_a(T_{\geq n}) \cap \Phi^{\beta}_{h_n}(T_{\geq n}) \neq \varnothing \right\}.$$

Corollary 2 (Consequential Bellman characterisation)**.** *Under the assumptions of Theorem 5.3, for every $n \in N^d(T)$,*

$$\widehat{\beta}_{h_n}(T_{\geq n}) = \underset{[a]_n \in \bar{N}^{+1}(n)}{\arg\max}\ Q_{h_n}(n, [a]_n).$$

The Bellman equation evaluates a control by rolling the value forward to the first subsequent decision node or terminal node. If a link reaches another decision node, the continuation value $V(m(\ell))$ is converted into h_n-units using the affine coefficient associated with the realised segment $\mathbf{y}(\ell)$. If the link terminates, $\mathbf{y}(\ell)$ is already a complete continuation chain and is evaluated directly by u_{h_n}. Thus chance nodes between decision nodes are handled by averaging over transition links, while choices are handled by maximisation over consequentially distinct immediate controls.

6 Concluding Summary

In the traditional definition of a non-cooperative game, any player's objectives are assumed to be represented by the expected value of a real-valued payoff function. Following Raiffa (1968) in particular, one-person games in extensive form, when regarded as finite decision trees, also have a pecuniary payoff attached to each terminal node. In decision theory, payoffs are typically replaced by consequences in a specified domain, as in the key pioneering works of Arrow (1959), Savage (1954), and Anscombe and Aumann (1963). In Hammond (2022) and earlier papers, consequences rather than payoffs were attached to terminal nodes of decision trees.

This paper has explored the implications of allowing intermediate consequences that occur before the end of a decision tree. Specifically, we consider decision trees that contain a new class of "timed consequence" nodes at which no branching occurs, but there is a timed consequence which matter to the decision maker, and so affects the decisions that one should expect to be taken.

Formally, decision trees with intermediate consequence nodes are equivalent to trees where all such nodes have been removed, but where each terminal consequence becomes elongated into a consequence chain or history of successive consequences that have arisen along the entire path in the tree which leads to that terminal node.

Appendix A Proofs

Recall that a timed-consequence chain is $\mathbf{y} = (\boldsymbol{y}_j)_{j=1}^m \in \mathbf{Y}(\mathbb{R} \times Y)$, where $\boldsymbol{y}_j = (t_j, y_j)$ and the times are strictly increasing, $t_{j+1} > t_j$. For a history $h(n) = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k)$ and a chain $\mathbf{s}$, the concatenation $(h(n); \mathbf{s})$ is a valid chain whenever $\mathbf{s}$ is empty or, when non-empty, its first time is strictly greater than the last time in $h(n)$. At a timed node n, let $\boldsymbol{y}(n) = (t(n), y(n))$ be its consequence; if n is intermediate with unique successor n', then $h(n') = (h(n); \boldsymbol{y}(n))$. When a tree is evaluated after an external history h, we write $h_n = (h; h(n))$. We start with some preliminary results.

Lemma 1. *Fix a consequentially normal-form invariant behaviour rule β on $\mathcal{D}$. Fix $(h, T) \in \mathcal{D}$, and let $n \in N^d(T)$. Write $h_n = (h; h(n))$. Suppose that the choice correspondence $\mathcal{C}^\beta_{h_n}$ is rationalised by the base preference relation $\succsim^\beta_{h_n}$ on finite feasible sets. Then*

$$\Phi^\beta_{h_n}(T_{\geq n}) = \max\left(F(T_{\geq n}), \succsim^\beta_{h_n}\right),$$

where, for any finite set A, one has

$$\max(A, \succsim) := \{\boldsymbol{\lambda} \in A : \boldsymbol{\lambda}' \in A \Longrightarrow \boldsymbol{\lambda} \succsim \boldsymbol{\lambda}'\}.$$

Proof. Since $(h_n, T_{\geq n}) \in \mathcal{D}$, Theorem 4.1 gives $\Phi^\beta_{h_n}(T_{\geq n}) = \mathcal{C}^\beta_{h_n}(F(T_{\geq n}))$. By hypothesis, $\mathcal{C}^\beta_{h_n}$ is rationalised by $\succsim^\beta_{h_n}$ on finite feasible sets. Hence, for every finite $A \subseteq \Delta(\mathbf{Y}_{>h_n}(\mathbb{R} \times Y))$, $\mathcal{C}^\beta_{h_n}(A) = \max(A, \succsim^\beta_{h_n})$. Applying this equality to $A = F(T_{\geq n})$ gives

$$\Phi^\beta_{h_n}(T_{\geq n}) = \max\left(F(T_{\geq n}), \succsim^\beta_{h_n}\right),$$

as required. □

A.1 Proof of Proposition 1

For each node $n \in N(T)$, define the depth $d(n)$ of the continuation subtree $T_{\geq n}$ recursively by:

$$d(n) = \begin{cases} 0, & \text{if } n \in N^t(T), \\ 1 + \max_{n' \in N^{+1}(n)} d(n'), & \text{otherwise.} \end{cases} \tag{15}$$

Since T is finite, $h(n)$ is well defined and finite for every node n. Fix T and prove both claims simultaneously by induction on the finite depth $d(n)$ of $T_{\geq n}$.

Base cases. If $n \in N^t(T)$, then $\mathbf{p}(T_{\geq n}) = \langle\,\rangle$, so no decision can change any consequence, implying that both properties are immediate. If n is a timed consequence node with unique successor n', then $\mathbf{p}(T_{\geq n}) = \mathbf{p}(T_{\geq n'})$. Since no decision at n can change any consequence, both properties for $\mathbf{p}(T_{\geq n})$ follow directly from those for

$\mathbf{p}(T_{\geq n'})$.

Inductive step. Assume the claims hold for all m with $d(m) < d(n)$.

(1) *Decision node $n \in N^d(T)$.* Fix $n^* \in N^{+1}(n)$. If $N^d(T_{\geq n^*}) = \emptyset$, the recursion yields $\mathbf{p}(T_{\geq n}) = \langle n^* \rangle$, which records exactly one move at n and none elsewhere. Reachability holds trivially, and chance-exhaustion is vacuous. If $N^d(T_{\geq n^*}) \neq \emptyset$, then

$$\mathbf{p}(T_{\geq n}) = \langle n^* \rangle \| \mathbf{p}(T_{\geq n^*}).$$

By the induction hypothesis, $\mathbf{p}(T_{\geq n^*})$ records moves only at decision nodes reachable within $T_{\geq n^*}$ and is exhaustive at every reached chance node downstream. Prefixing with $\langle n^* \rangle$ introduces a move only at n and preserves the structure of all downstream mixtures. By Remark 1, concatenation is compatible with mixtures. Hence both reachability and chance-exhaustion hold for $\mathbf{p}(T_{\geq n})$.

(2) *Chance node $n \in N^c(T)$.* The recursion gives

$$\mathbf{p}(T_{\geq n}) = \sum_{n' \in N^{+1}(n)} \pi(n' \mid n)\, \mathbf{p}(T_{\geq n'}).$$

By the induction hypothesis, for each $n' \in N^{+1}(n)$ the strategy $\mathbf{p}(T_{\geq n'})$ satisfies reachability and is exhaustive at all reached chance nodes within $T_{\geq n'}$. Since the recursion includes every successor $n' \in N^{+1}(n)$, the strategy unfolds at n into all successor branches, with their full continuation plans. Thus chance-exhaustion holds at n, and reachability and exhaustiveness downstream follow from the induction hypothesis. This completes the induction.

A.2 Proof of Theorem 2.1

Existence. Fix $\mathbf{p} \in \mathbf{P}(T)$. We construct $\boldsymbol{\eta}(T_{\geq n}, \mathbf{p})$ for all nodes n by induction on the finite depth $d(n)$ of $T_{\geq n}$.

Base (depth 0). If $n \in N^t(T)$, set $\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) := \delta_{\boldsymbol{y}(n)}$, which is a probability measure.

Inductive step. As the induction hypothesis, suppose that $\boldsymbol{\eta}(T_{\geq m}, \mathbf{p})$ is defined for all m with $d(m) < d(n)$.

(a) If n is an intermediate timed consequence node with unique successor n', define $\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) := \boldsymbol{y}(n) \circ \boldsymbol{\eta}(T_{\geq n'}, \mathbf{p})$. By Remark 1, this is a probability measure.

(b) If $n \in N^d(T)$, let n^* denote the unique successor prescribed by $\mathbf{p}$ at n, equivalently, the first element of the contingent node–sequence $\mathbf{p}(T_{\geq n})$. By the induction hypothesis, $\boldsymbol{\eta}(T_{\geq n^*}, \mathbf{p})$ is defined. Set $\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) := \boldsymbol{\eta}(T_{\geq n^*}, \mathbf{p})$.

(c) If n is a chance node with kernel $\pi(\cdot \mid n)$, define the mixture

$$\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) := \sum_{n' \in N^{+1}(n)} \pi(n' \mid n)\, \boldsymbol{\eta}(T_{\geq n'}, \mathbf{p})$$

of probability measures. This is well defined by the induction hypothesis.

Since T is finite, the depth $d(n)$ is finite for every node n, so backward recursion succeeds in defining $\boldsymbol{\eta}(T_{\geq n}, \mathbf{p})$ for all n. Set $\boldsymbol{\eta}(T, \mathbf{p}) := \boldsymbol{\eta}(T_{\geq n_0(T)}, \mathbf{p})$. By construction, the recursive characterisation stated in Theorem 2.1 holds.

Uniqueness. Suppose $\widehat{\boldsymbol{\eta}}(T_{\geq n}, \mathbf{p})$ is another mapping that satisfies the same recursive characterisation. We show that $\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) = \widehat{\boldsymbol{\eta}}(T_{\geq n}, \mathbf{p})$ by induction on depth. At any terminal node n, both equal the degenerate lottery $\delta_{\boldsymbol{y}(n)}$. At an intermediate timed consequence node n with successor n', the induction hypothesis implies that the two mappings coincide at n', and therefore their push-forwards under $x \mapsto \boldsymbol{y}(n) \circ x$ coincide at n. At a decision node n, both equal the value at the uniquely selected successor n^*, which coincides by the induction hypothesis. At a chance node, both equal the same convex combination of values at successors, which coincide by the induction hypothesis. Hence $\boldsymbol{\eta} = \widehat{\boldsymbol{\eta}}$ pointwise on all subtrees, in particular on T.

Truncations. Since the construction is by continuation subtrees $T_{\geq n}$, the same argument applies verbatim to any truncation.

A.3 Proof of Theorem 2.2

Fix $T \in \mathcal{T}(\mathbb{R} \times Y)$. For each node $n \in N(T)$, define

$$F(T_{\geq n}) := \big\{\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) \ : \ \mathbf{p} \in \mathbf{P}(T_{\geq n})\big\} \subseteq \Delta\big(\mathbf{Y}(\mathbb{R} \times Y)\big).$$

We show that the family $\{F(T_{\geq n}) : n \in N(T)\}$ satisfies the backward recurrence, and that it is the unique family of subsets of $\Delta(\mathbf{Y}(\mathbb{R} \times Y))$ that does so.

Step 1 (Existence). Fix $n \in N(T)$. We verify each clause of the recurrence.

(a) Terminal node $n \in N^t(T)$. Then $T_{\geq n}$ consists only of the terminal timed consequence node n. Hence $\mathbf{P}(T_{\geq n})$ contains only the trivial path strategy that ends at n, and by definition of $\boldsymbol{\eta}$ at a terminal timed node,

$$\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) = \delta_{\boldsymbol{y}(n)}.$$

Therefore $F(T_{\geq n}) = \{\delta_{\boldsymbol{y}(n)}\}$.

(b) Intermediate timed node $n \in N^i(T)$ with $N^{+1}(n) = \{n'\}$. Let $\mathbf{p} \in \mathbf{P}(T_{\geq n})$. By the recursive definition of path strategies, every $\mathbf{p}$ passes from n to its unique successor n' and then continues as some $\mathbf{p}' \in \mathbf{P}(T_{\geq n'})$. By the recursion for the consequence mapping $\boldsymbol{\eta}$ at an intermediate timed node,

$$\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) = \boldsymbol{y}(n) \circ \boldsymbol{\eta}(T_{\geq n'}, \mathbf{p}').$$

Taking the range over all possible $\mathbf{p} \in \mathbf{P}(T_{\geq n})$ is therefore equivalent to taking the range over all $\mathbf{p}' \in \mathbf{P}(T_{\geq n'})$ and then prefixing by $\boldsymbol{y}(n)$. Hence $F(T_{\geq n}) = \boldsymbol{y}(n) \circ F(T_{\geq n'})$.

(c) Decision node $n \in N^d(T)$. Let $\mathbf{p} \in \mathbf{P}(T_{\geq n})$. Because of the recursive construction, the path strategy $\mathbf{p}$ chooses exactly one successor $n' \in N^{+1}(n)$ and then continues as some $\mathbf{p}' \in \mathbf{P}(T_{\geq n'})$. Conversely, for each $n' \in N^{+1}(n)$ and each $\mathbf{p}' \in \mathbf{P}(T_{\geq n'})$, there exists a path strategy in $\mathbf{P}(T_{\geq n})$ that chooses node n' at n and then follows $\mathbf{p}'$ thereafter. Since $\boldsymbol{\eta}$ depends only on the realised contingent path, one has

$$\big\{\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) : \mathbf{p} \in \mathbf{P}(T_{\geq n})\big\} = \bigcup_{n' \in N^{+1}(n)} \big\{\boldsymbol{\eta}(T_{\geq n'}, \mathbf{p}') : \mathbf{p}' \in \mathbf{P}(T_{\geq n'})\big\}.$$

Thus $F(T_{\geq n}) = \bigcup_{n' \in N^{+1}(n)} F(T_{\geq n'})$.

(d) Chance node $n \in N^c(T)$. Let $\mathbf{p} \in \mathbf{P}(T_{\geq n})$. By the structure of $T_{\geq n}$, chance selects $n' \in N^{+1}(n)$ with probability $\pi(n' \mid n)$, and the continuation thereafter is specified by a path strategy $\mathbf{p}_{n'} \in \mathbf{P}(T_{\geq n'})$. Because of the recursive construction of $\boldsymbol{\eta}$, at a chance node one has

$$\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) = \sum_{n' \in N^{+1}(n)} \pi(n' \mid n)\, \boldsymbol{\eta}(T_{\geq n'}, \mathbf{p}_{n'}).$$

Hence every element of $F(T_{\geq n})$ is a mixture $\sum_{n' \in N^{+1}(n)} \pi(n' \mid n)\lambda_{n'}$ with $\lambda_{n'} \in F(T_{\geq n'})$ for each $n' \in N^{+1}(n)$.

Conversely, given any selection satisfying $\lambda_{n'} \in F(T_{\geq n'})$ for each $n' \in N^{+1}(n)$, choose $\mathbf{p}_{n'} \in \mathbf{P}(T_{\geq n'})$ for each $n' \in N^{+1}(n)$, so that $\boldsymbol{\eta}(T_{\geq n'}, \mathbf{p}_{n'}) = \lambda_{n'}$. Then assemble all these into a contingent path strategy in $\mathbf{P}(T_{\geq n})$. The induced consequence lottery is then exactly $\sum_{n' \in N^{+1}(n)} \pi(n' \mid n)\lambda_{n'}$. Therefore

$$F(T_{\geq n}) = \sum_{n' \in N^{+1}(n)} \pi(n' \mid n)\, F(T_{\geq n'}).$$

This verifies the backward recurrence.

Step 2 (Uniqueness). Let $\{G(T_{\geq n}) : n \in N(T)\}$ be any family of subsets of $\Delta(\mathbf{Y}(\mathbb{R} \times Y))$ satisfying the same recurrence as in Step 1. We show $G(T_{\geq n}) = F(T_{\geq n})$ for every n by induction on the depth $d(n)$ of n defined by (15).

Base case (depth 0). If $n \in N^t(T)$, the backward recursion in the first case implies that $G(T_{\geq n}) = \{\delta_{\boldsymbol{y}(n)}\} = F(T_{\geq n})$.

Inductive step. Fix $k \geq 0$ and assume $G(T_{\geq m}) = F(T_{\geq m})$ for all nodes m of depth at most k. Let n have depth $k+1$.

If $n \in N^i(T)$ with unique successor n', then the backward recursion in the second case implies that $G(T_{\geq n}) = \boldsymbol{y}(n) \circ G(T_{\geq n'}) = \boldsymbol{y}(n) \circ F(T_{\geq n'}) = F(T_{\geq n})$, where the middle equality uses the induction hypothesis.

If $n \in N^d(T)$, backward recursion in the third case yields

$$G(T_{\geq n}) = \bigcup_{n' \in N^{+1}(n)} G(T_{\geq n'}) = \bigcup_{n' \in N^{+1}(n)} F(T_{\geq n'}) = F(T_{\geq n}).$$

If $n \in N^c(T)$, then the backward recursion in the fourth case implies that

$$G(T_{\geq n}) = \sum_{n' \in N^{+1}(n)} \pi(n' \mid n)\, G(T_{\geq n'}) = \sum_{n' \in N^{+1}(n)} \pi(n' \mid n)\, F(T_{\geq n'}) = F(T_{\geq n}).$$

Thus $G = F$ in all cases, and so at all nodes, proving uniqueness.

Step 3 (Root identification). At the root $n_0(T)$, one has $T_{\geq n_0(T)} = T$. Hence $F(T) = F(T_{\geq n_0(T)})$, by definition.

A.4 Proof of Theorem 3.1

Fix $(h, T) \in \mathcal{D}$ and a behaviour rule β. For each node $n \in N(T)$, write $h_n := (h; h(n))$. Define

$$\Phi^{\beta}_{h_n}(T_{\geq n}) := \left\{ \boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) : \mathbf{p} \in \mathbf{P}^{\beta}_{h_n}(T_{\geq n}) \right\} \subseteq F(T_{\geq n}).$$

We show that the family $\{\Phi^{\beta}_{h_n}(T_{\geq n}) : n \in N(T)\}$ satisfies the stated backward recurrence, and that it is the unique family of subsets of $\Delta(\mathbf{Y}(\mathbb{R} \times Y))$ which does so.

Step 1 (Existence). Fix $n \in N(T)$. The verification is identical to the proof of Theorem 2.2 except at decision nodes, where β restricts the admissible successors.

(a) Terminal node $n \in N^t(T)$. As in Step 1(a) of Theorem 2.2, the continuation subtree $T_{\geq n}$ has the terminal node n as its only node. Hence $\mathbf{P}^{\beta}_{h_n}(T_{\geq n})$ has the trivial path strategy ending at n as its only member. It follows that

$$\Phi^{\beta}_{h_n}(T_{\geq n}) = \{\delta_{\boldsymbol{y}(n)}\}.$$

(b) Intermediate timed node $n \in N^i(T)$ with $N^{+1}(n) = \{n'\}$. As in Step 1(b) of Theorem 2.2, every $\mathbf{p} \in \mathbf{P}^{\beta}_{h_n}(T_{\geq n})$ passes to n' and then continues as some $\mathbf{p}' \in \mathbf{P}^{\beta}_{h_{n'}}(T_{\geq n'})$. Using the recursion for $\boldsymbol{\eta}$ at an intermediate timed node and taking the range over $\mathbf{P}^{\beta}$, we obtain

$$\Phi^{\beta}_{h_n}(T_{\geq n}) = \boldsymbol{y}(n) \circ \Phi^{\beta}_{h_{n'}}(T_{\geq n'}).$$

(c) Decision node $n \in N^d(T)$. Let $\mathbf{p} \in \mathbf{P}^{\beta}_{h_n}(T_{\geq n})$. By β-admissibility, the unique successor $n' \in N^{+1}(n)$ selected by $\mathbf{p}$ at n satisfies $n' \in \beta_{h_n}(T_{\geq n})$, and thereafter $\mathbf{p}$

continues as some element of $\mathbf{P}^{\beta}_{h_{n'}}(T_{\geq n'})$. Hence

$$\Phi^{\beta}_{h_n}(T_{\geq n}) \subseteq \bigcup_{n' \in \beta_{h_n}(T_{\geq n})} \Phi^{\beta}_{h_{n'}}(T_{\geq n'}). \tag{16}$$

Conversely, fix $n' \in \beta_{h_n}(T_{\geq n})$ and $\boldsymbol{\lambda} \in \Phi^{\beta}_{h_{n'}}(T_{\geq n'})$. By definition of Φ^{β}, there exists $\mathbf{p}' \in \mathbf{P}^{\beta}_{h_{n'}}(T_{\geq n'})$ with $\boldsymbol{\eta}(T_{\geq n'}, \mathbf{p}') = \boldsymbol{\lambda}$. Prepending the move $n \to n'$ at n yields a β-admissible path strategy $\mathbf{p} \in \mathbf{P}^{\beta}_{h_n}(T_{\geq n})$ whose induced consequence lottery is $\boldsymbol{\lambda}$. Thus the reverse of inclusion (16) also holds, and therefore

$$\Phi^{\beta}_{h_n}(T_{\geq n}) = \bigcup_{n' \in \beta_{h_n}(T_{\geq n})} \Phi^{\beta}_{h_{n'}}(T_{\geq n'}).$$

(d) Chance node $n \in N^c(T)$. As in Step 1(d) of Theorem 2.2, conditional on each successor $n' \in N^{+1}(n)$ the continuation is determined by some $\mathbf{p}_{n'} \in \mathbf{P}^{\beta}_{h_{n'}}(T_{\geq n'})$, and the recursion for $\boldsymbol{\eta}$ at chance nodes yields a fixed-weight mixture. Taking the range over all such admissible continuations gives

$$\Phi^{\beta}_{h_n}(T_{\geq n}) = \sum_{n' \in N^{+1}(n)} \pi(n' \mid n)\, \Phi^{\beta}_{h_{n'}}(T_{\geq n'}).$$

This verifies the backward recurrence.

Step 2 (Uniqueness). Let $\{G_n : n \in N(T)\}$ be any family of subsets of $\Delta(\mathbf{Y}(\mathbb{R} \times Y))$ that satisfies the same backward recurrence as in Step 1. The induction argument in Step 2 of Theorem 2.2 applies verbatim, with the sole modification that at decision nodes the recurrence takes unions over $\beta_{h_n}(T_{\geq n})$ rather than over $N^{+1}(n)$. Since the induction hypothesis equates $G_{n'}$ with $\Phi^{\beta}_{h_{n'}}(T_{\geq n'})$ at successor nodes, each of the four node types yields $G_n = \Phi^{\beta}_{h_n}(T_{\geq n})$. Hence $G_n = \Phi^{\beta}_{h_n}(T_{\geq n})$ for every $n \in N(T)$.

Step 3 (Initial identification). At the initial node $n_0(T)$, one has $h(n_0(T)) = \varnothing$, so $h_{n_0(T)} = (h; \varnothing) = h$. Since $T_{\geq n_0(T)} = T$, the definition gives

$$\Phi^{\beta}_{h}(T) = \Phi^{\beta}_{h_{n_0(T)}}(T_{\geq n_0(T)}) = \Phi^{\beta}_{(h; h(n_0(T)))}(T_{\geq n_0(T)}).$$

A.5 Proof of Proposition 2

Recall that $F(T) = \{\boldsymbol{\eta}(T, \mathbf{p}) : \mathbf{p} \in \mathbf{P}(T)\}$.

(1) Equality $F(T^{\mathrm{R}}) = F(T)$. By construction of the reduced normal form tree T^{R}, it has a single decision node at the root $n_0^{\mathrm{R}}(T)$ and one terminal node $n^t_{\mathbf{p}}$ for each path strategy $\mathbf{p} \in \mathbf{P}(T)$. So the only decision strategies in T^{R} are the one-step strategies

$$\mathbf{q}_{\mathbf{p}} = \langle n_0^{\mathrm{R}}(T) \to n^t_{\mathbf{p}} \rangle,$$

and the terminal lottery at $n^t_{\mathbf{p}}$ is defined to be

$$\boldsymbol{\eta}(T^{\mathrm{R}}, n^t_{\mathbf{p}}) = \boldsymbol{\eta}(T, \mathbf{p}).$$

Therefore,

$$\boldsymbol{\eta}\big(T^{\mathrm{R}}, \mathbf{q}_{\mathbf{p}}\big) = \boldsymbol{\eta}(T, \mathbf{p}) \qquad \text{for every } \mathbf{p} \in \mathbf{P}(T).$$

Taking ranges gives

$$F\big(T^{\mathrm{R}}\big) = \{\boldsymbol{\eta}(T^{\mathrm{R}}, \mathbf{q}_{\mathbf{p}}) : \mathbf{p} \in \mathbf{P}(T)\} = \{\boldsymbol{\eta}(T, \mathbf{p}) : \mathbf{p} \in \mathbf{P}(T)\} = F(T). \tag{17}$$

(2) Equality $F(T^{\mathrm{C}}) = F(T)$. By construction of T^{C}, its only decision node is $n^{\mathrm{C}}_0(T)$. Also, it has just one terminal node $n^t_{\boldsymbol{\lambda}}$ corresponding to each $\boldsymbol{\lambda} \in F(T)$. So the only decision strategies in T^{C} take the form

$$\mathbf{r}_{\boldsymbol{\lambda}} = \langle n^{\mathrm{C}}_0(T) \to n^t_{\boldsymbol{\lambda}} \rangle,$$

Also, the terminal lottery at each terminal node $n^t_{\boldsymbol{\lambda}}$ is defined to be

$$\eta(T^{\mathrm{C}}, n^t_{\boldsymbol{\lambda}}) = \boldsymbol{\lambda}.$$

Hence

$$\boldsymbol{\eta}\big(T^{\mathrm{C}}, \mathbf{r}_{\boldsymbol{\lambda}}\big) = \boldsymbol{\lambda} \quad \text{for every } \boldsymbol{\lambda} \in F(T)$$

It follows that

$$F\big(T^{\mathrm{C}}\big) = \{\boldsymbol{\eta}(T^{\mathrm{C}}, \mathbf{r}_{\boldsymbol{\lambda}}) : \boldsymbol{\lambda} \in F(T)\} = F(T) \tag{18}$$

Combining the two equalities (17) and (18) yields $F(T^{\mathrm{R}}) = F(T^{\mathrm{C}}) = F(T)$, as claimed.

A.6 Proof of Theorem 4.1

We use two basic facts established above. First, for each finite tree T, the feasible set $F(T)$ is finite. Second, by construction, $\Phi^{\beta}_h(T) \subseteq F(T)$ for every behaviour rule β and every $T \in \mathcal{D}_h$.

Lemma 2. *Fix $h \in H$. For every non-empty finite set $F \subseteq \Delta(\mathbf{Y}_{>h}(\mathbb{R} \times Y))$, there exists a tree $T_F \in \mathcal{D}_h$ with $F(T_F) = F$.*

Proof. Let $F = \{\boldsymbol{\lambda}^1, \ldots, \boldsymbol{\lambda}^m\}$. For each ℓ, write $\boldsymbol{\lambda}^\ell$ as a finite lottery over chains in $\mathbf{Y}_{>h}(\mathbb{R} \times Y)$. By unrestricted domain, there is a history-compatible subtree implementing $\boldsymbol{\lambda}^\ell$ after h. Construct T_F by adding a decision root with one successor leading to each of these subtrees. Then choosing successor ℓ at the root induces $\boldsymbol{\lambda}^\ell$, and no other continuation lotteries are generated. Hence $T_F \in \mathcal{D}_h$ and $F(T_F) = F$. □

($\Rightarrow$) Suppose that β is consequentially normal-form invariant. Fix $h \in H$. Given any non-empty finite set $F \subseteq \Delta(\mathbf{Y}_{>h}(\mathbb{R} \times Y))$, Lemma 2 gives a tree $T_F \in \mathcal{D}_h$ with $F(T_F) = F$. Define

$$\mathcal{C}_h^\beta(F) := \Phi_h^\beta(T_F).$$

This is well defined. If $\widetilde{T}_F \in \mathcal{D}_h$ is another tree with $F(\widetilde{T}_F) = F$, then consequential normal-form invariance gives $\Phi_h^\beta(T_F) = \Phi_h^\beta(\widetilde{T}_F)$. Hence $\mathcal{C}_h^\beta(F)$ does not depend on the representing tree. Moreover, $\mathcal{C}_h^\beta(F) = \Phi_h^\beta(T_F) \subseteq F(T_F) = F$, and it is non-empty because β selects a non-empty set of admissible successors at every reached decision node. Thus $\mathcal{C}_h^\beta$ is a choice correspondence on finite subsets of $\Delta(\mathbf{Y}_{>h}(\mathbb{R} \times Y))$.

Finally, for any $T \in \mathcal{D}_h$, one can take $T_{F(T)} = T$ in the preceding construction. Therefore

$$\mathcal{C}_h^\beta(F(T)) = \Phi_h^\beta(T),$$

which proves (6).

($\Leftarrow$) Conversely, suppose that, for each $h \in H$, there exists a choice correspondence $\mathcal{C}_h^\beta$ such that $\Phi_h^\beta(T) = \mathcal{C}_h^\beta(F(T))$ for every $T \in \mathcal{D}_h$. Fix $h \in H$, and let $T, \widetilde{T} \in \mathcal{D}_h$ satisfy $F(T) = F(\widetilde{T})$. Then

$$\Phi_h^\beta(T) = \mathcal{C}_h^\beta(F(T)) = \mathcal{C}_h^\beta(F(\widetilde{T})) = \Phi_h^\beta(\widetilde{T}).$$

This is consequential normal-form invariance at history h. Since h was arbitrary, β is consequentially normal-form invariant on $\mathcal{D}$.

A.7 Proof of Theorem 5.1

Fix $h \in H$, and write $C_h := \mathbf{Y}_{>h}(\mathbb{R} \times Y)$.

($\Rightarrow$) Assume that $\succsim_h$ on $\Delta(C_h)$ is prerational and mixture continuous. By Definition 13, there exists a consequentially normal-form invariant behaviour rule β on $\mathcal{D}$ such that $\succsim_h = \succsim_h^\beta$. By Theorem 4.1, the h-section of β induces a choice correspondence $\mathcal{C}_h^\beta$ on finite subsets of $\Delta(C_h)$ satisfying $\Phi_h^\beta(T) = \mathcal{C}_h^\beta(F(T))$ for every $T \in \mathcal{D}_h$. Hence the dynamic structure of any tree $T \in \mathcal{D}_h$ affects behaviour after h only through the finite menu $F(T) \subseteq \Delta(C_h)$.

By the consequentialist normal-form reduction in Proposition 2, every such finite menu can be represented on the terminal domain with consequence space C_h. Therefore the static prerationality theorem of Hammond (1988a,b, 2022) applies to $\succsim_h$ on $\Delta(C_h)$. Since $\succsim_h$ is mixture continuous, there exists a Bernoulli index $u_h : C_h \to \mathbb{R}$ such that, for all $\boldsymbol{\lambda}, \boldsymbol{\lambda}' \in \Delta(C_h)$,

$$\boldsymbol{\lambda} \succsim_h \boldsymbol{\lambda}' \quad \iff \quad \sum_{\mathbf{y} \in \operatorname{supp}(\boldsymbol{\lambda})} \boldsymbol{\lambda}(\mathbf{y}) u_h(\mathbf{y}) \geq \sum_{\mathbf{y} \in \operatorname{supp}(\boldsymbol{\lambda}')} \boldsymbol{\lambda}'(\mathbf{y}) u_h(\mathbf{y}).$$

The same theorem gives cardinal uniqueness: any two Bernoulli indices representing $\succsim_h$ differ by a positive affine transformation.

($\Leftarrow$) Conversely, suppose that there exists a Bernoulli index $u_h : C_h \to \mathbb{R}$ such that the expected utility functional U_h defined in (**??**) represents $\succsim_h$.

First, $\succsim_h$ is mixture continuous. Indeed, for any $\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu} \in \Delta(C_h)$, the map

$$\alpha \mapsto U_h\big(\alpha \boldsymbol{\lambda} + (1-\alpha)\boldsymbol{\nu}\big)$$

is affine, hence continuous, on $[0,1]$. The upper and lower contour sets appearing in Definition 14 are therefore closed.

It remains to show prerationality. Define a choice correspondence on finite non-empty subsets $F \subseteq \Delta(C_h)$ by

$$\mathcal{C}_h^u(F) := \arg\max_{\boldsymbol{\lambda} \in F} U_h(\boldsymbol{\lambda}).$$

Since F is finite, $\mathcal{C}_h^u(F)$ is non-empty and contained in F. By the static prerationality theorem of Hammond (1988a,b, 2022), this choice correspondence is induced, on the terminal domain with consequence space C_h, by a consequentially normal-form invariant behaviour rule. Extending that rule to arbitrary trees in $\mathcal{D}_h$ by setting $\Phi_h^{\beta_u}(T) := \mathcal{C}_h^u(F(T))$, and choosing arbitrary consequentially normal-form invariant behaviour on histories other than h, gives a behaviour rule β_u on $\mathcal{D}$ whose h-section is consequentially normal-form invariant. Its induced base preference satisfies, for all $\boldsymbol{\lambda}, \boldsymbol{\lambda}' \in \Delta(C_h)$,

$$\boldsymbol{\lambda} \succsim_h^{\beta_u} \boldsymbol{\lambda}' \quad \iff \quad \boldsymbol{\lambda} \in \mathcal{C}_h^u\big(\{\boldsymbol{\lambda}, \boldsymbol{\lambda}'\}\big) \quad \iff \quad U_h(\boldsymbol{\lambda}) \geq U_h(\boldsymbol{\lambda}') \quad \iff \quad \boldsymbol{\lambda} \succsim_h \boldsymbol{\lambda}'.$$

Thus $\succsim_h = \succsim_h^{\beta_u}$, so $\succsim_h$ is prerational by Definition 13. This completes the proof.

A.8 Proof of Proposition 3

Fix $(h, T) \in \mathcal{D}$, and let $n \in N^d(T)$. Write $h_n = (h; h(n))$. By the closure property of the history-compatible domain, $(h_n, T_{\geq n}) \in \mathcal{D}$. Since β is a consequentially normal-form invariant behaviour rule on $\mathcal{D}$, its h_n-section induces the choice correspondence $\mathcal{C}_{h_n}^{\beta}$ and the associated base preference relation $\succsim_{h_n}^{\beta}$ on $\Delta(\mathbf{Y}_{>h_n}(\mathbb{R} \times Y))$.

By Definition 13, a relation on $\Delta(\mathbf{Y}_{>h_n}(\mathbb{R} \times Y))$ is prerational if it is induced by some consequentially normal-form invariant behaviour rule on $\mathcal{D}$. The relation $\succsim_{h_n}^{\beta}$ is induced by the same rule β. Hence $\succsim_{h_n}^{\beta}$ is prerational. Since $n \in N^d(T)$ was arbitrary, the conclusion holds for every decision node $n \in N^d(T)$.

A.9 Proof of Proposition 4

Fix $(h, T) \in \mathcal{D}$, and write $h_n = (h; h(n))$ for each $n \in N^d(T)$. Let $n, n' \in N^d(T)$ be such that n' is a proximate successor decision node of n, and let $\mathbf{y}$ be the decision-node-free timed consequence segment linking n to n'. Then $h_{n'} = (h_n; \mathbf{y})$.

Fix arbitrary $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \Delta(\mathbf{Y}_{>h_{n'}}(\mathbb{R} \times Y))$. By unrestricted domain, there exists a tree $B \in \mathcal{D}_{h_{n'}}$ such that $F(B) = \{\boldsymbol{\lambda}, \boldsymbol{\mu}\}$. By Theorem 4.1,

$$\Phi^{\beta}_{h_{n'}}(B) = \mathcal{C}^{\beta}_{h_{n'}}\big(\{\boldsymbol{\lambda}, \boldsymbol{\mu}\}\big).$$

Now construct a tree $B^{\mathbf{y}} \in \mathcal{D}_{h_n}$ whose initial decision node has a single successor, after which the tree follows the deterministic segment $\mathbf{y}$, with no further decision node before reaching the initial decision node of B. Hence

$$F(B^{\mathbf{y}}) = \{\mathbf{y} \circ \boldsymbol{\lambda}, \mathbf{y} \circ \boldsymbol{\mu}\}.$$

Since the segment after the initial singleton choice and before B contains no decision node, admissible choices in $B^{\mathbf{y}}$ are exactly the admissible choices in B, prefixed by $\mathbf{y}$. Thus, by the recursion for Φ^{β},

$$\Phi^{\beta}_{h_n}(B^{\mathbf{y}}) = \mathbf{y} \circ \Phi^{\beta}_{h_{n'}}(B).$$

Applying Theorem 4.1 again gives

$$\mathcal{C}^{\beta}_{h_n}\big(\{\mathbf{y} \circ \boldsymbol{\lambda}, \mathbf{y} \circ \boldsymbol{\mu}\}\big) = \mathbf{y} \circ \mathcal{C}^{\beta}_{h_{n'}}\big(\{\boldsymbol{\lambda}, \boldsymbol{\mu}\}\big).$$

By definition of the induced base preference relation,

$$\mathbf{y} \circ \boldsymbol{\lambda} \succsim^{\beta}_{h_n} \mathbf{y} \circ \boldsymbol{\mu} \quad \iff \quad \mathbf{y} \circ \boldsymbol{\lambda} \in \mathcal{C}^{\beta}_{h_n}\big(\{\mathbf{y} \circ \boldsymbol{\lambda}, \mathbf{y} \circ \boldsymbol{\mu}\}\big).$$

Using the equality just obtained and the injectivity of prefixing by the fixed segment $\mathbf{y}$, this is equivalent to

$$\boldsymbol{\lambda} \in \mathcal{C}^{\beta}_{h_{n'}}\big(\{\boldsymbol{\lambda}, \boldsymbol{\mu}\}\big),$$

which, again by definition of the induced base preference relation, is equivalent to

$$\boldsymbol{\lambda} \succsim^{\beta}_{h_{n'}} \boldsymbol{\mu}.$$

Therefore

$$\mathbf{y} \circ \boldsymbol{\lambda} \succsim^{\beta}_{h_n} \mathbf{y} \circ \boldsymbol{\mu} \quad \iff \quad \boldsymbol{\lambda} \succsim^{\beta}_{h_{n'}} \boldsymbol{\mu}.$$

Since n, n', $\boldsymbol{\lambda}$, and $\boldsymbol{\mu}$ were arbitrary, the induced family $\big(\succsim^{\beta}_{h_n}\big)_{n \in N^d(T)}$ is dynamically consistent.

A.10 Proof of Theorem 5.2

Fix $(h,T) \in \mathcal{D}$, and write $h_n = (h; h(n))$ for each $n \in N^d(T)$. By Proposition 3, each induced base preference relation $\succsim^{\beta}_{h_n}$ is prerational. By hypothesis, it is also mixture continuous and non-trivial on $\Delta\big(\mathbf{Y}_{>h_n}(\mathbb{R} \times Y)\big)$. Theorem 5.1 therefore gives, for each $n \in N^d(T)$, a non-constant Bernoulli utility index $u_{h_n} : \mathbf{Y}_{>h_n}(\mathbb{R} \times Y) \to \mathbb{R}$ such that

$$U_{h_n}(\boldsymbol{\lambda}) := \sum_{\mathbf{z} \in \operatorname{supp}(\boldsymbol{\lambda})} \boldsymbol{\lambda}(\mathbf{z}) u_{h_n}(\mathbf{z})$$

represents $\succsim^{\beta}_{h_n}$ on $\Delta\big(\mathbf{Y}_{>h_n}(\mathbb{R} \times Y)\big)$. We now prove the affine relation. Let $n, n' \in N^d(T)$ be such that n' is a proximate successor decision node of n. Let $\mathbf{y}$ be the deterministic chain of timed consequences realised between n and n'. Then $h_{n'} = (h_n; \mathbf{y})$. For all $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \Delta\big(\mathbf{Y}_{>h_{n'}}(\mathbb{R} \times Y)\big)$, define the preference relation $\succsim^{\mathbf{y}}_{h_n}$ on the common tail-lottery domain by

$$\boldsymbol{\lambda} \succsim^{\mathbf{y}}_{h_n} \boldsymbol{\mu} \quad \Longleftrightarrow \quad \mathbf{y} \circ \boldsymbol{\lambda} \succsim^{\beta}_{h_n} \mathbf{y} \circ \boldsymbol{\mu}. \tag{19}$$

Since U_{h_n} represents $\succsim^{\beta}_{h_n}$, the induced relation $\succsim^{\mathbf{y}}_{h_n}$ is represented by

$$U^{\mathbf{y}}_{h_n}(\boldsymbol{\lambda}) := U_{h_n}(\mathbf{y} \circ \boldsymbol{\lambda}) = \sum_{\mathbf{s} \in \operatorname{supp}(\boldsymbol{\lambda})} \boldsymbol{\lambda}(\mathbf{s}) u_{h_n}(\mathbf{y}, \mathbf{s}).$$

Thus the Bernoulli index representing $\succsim^{\mathbf{y}}_{h_n}$ on $\mathbf{Y}_{>h_{n'}}(\mathbb{R} \times Y)$ is $\mathbf{s} \mapsto u_{h_n}(\mathbf{y}, \mathbf{s})$. By Proposition 4, the family induced by the same global rule β is dynamically consistent. Hence, for all $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \Delta\big(\mathbf{Y}_{>h_{n'}}(\mathbb{R} \times Y)\big)$, one has

$$\boldsymbol{\lambda} \succsim^{\mathbf{y}}_{h_n} \boldsymbol{\mu} \quad \Longleftrightarrow \quad \boldsymbol{\lambda} \succsim^{\beta}_{h_{n'}} \boldsymbol{\mu}.$$

Therefore $\succsim^{\mathbf{y}}_{h_n}$ and $\succsim^{\beta}_{h_{n'}}$ are the same non-trivial expected-utility preference relation on the same mixture space $\Delta\big(\mathbf{Y}_{>h_{n'}}(\mathbb{R} \times Y)\big)$. The relation $\succsim^{\mathbf{y}}_{h_n}$ is represented by the Bernoulli index $\mathbf{s} \mapsto u_{h_n}(\mathbf{y}, \mathbf{s})$, while $\succsim^{\beta}_{h_{n'}}$ is represented by $u_{h_{n'}}$. By the cardinal uniqueness part of Theorem 5.1, there exist unique numbers $v_{h_n}(\mathbf{y}) \in \mathbb{R}$, and $\xi_{h_n}(\mathbf{y}) > 0$, such that

$$u_{h_n}(\mathbf{y}, \mathbf{s}) = v_{h_n}(\mathbf{y}) + \xi_{h_n}(\mathbf{y}) u_{h_{n'}}(\mathbf{s})$$

for every $\mathbf{s} \in \mathbf{Y}_{>h_{n'}}(\mathbb{R} \times Y)$. This proves the local affine bridge for every realised proximate segment in T.

It remains to justify the extension of the coefficients to arbitrary continuation chains after h_n. Fix $n \in N^d(T)$ and let $\mathbf{y} \in \mathbf{Y}_{>h_n}(\mathbb{R} \times Y)$ be arbitrary. By the unrestricted-domain assumption, there exists an auxiliary tree $\widehat{T} \in \mathcal{D}_{h_n}$ with decision nodes $\widehat{n}, \widehat{n}' \in N^d(\widehat{T})$ such that $\widehat{n}'$ is a proximate successor decision node of $\widehat{n}$, the initial history at $\widehat{n}$ is h_n, and the deterministic segment between $\widehat{n}$ and $\widehat{n}'$ is precisely

$\mathbf{y}$. Applying the local argument just proved to this auxiliary tree yields numbers $\widehat{v}_{h_n}(\mathbf{y}) \in \mathbb{R}$, and $\widehat{\xi}_{h_n}(\mathbf{y}) > 0$, such that

$$u_{h_n}(\mathbf{y}, \mathbf{s}) = \widehat{v}_{h_n}(\mathbf{y}) + \widehat{\xi}_{h_n}(\mathbf{y}) u_{(h_n;\mathbf{y})}(\mathbf{s})$$

for every $\mathbf{s} \in \mathbf{Y}_{>(h_n;\mathbf{y})}(\mathbb{R} \times Y)$. These coefficients are independent of the auxiliary tree used to realise $\mathbf{y}$. Indeed, once h_n and $\mathbf{y}$ are fixed, the induced preference relation on $\Delta\big(\mathbf{Y}_{>(h_n;\mathbf{y})}(\mathbb{R} \times Y)\big)$ is given by

$$\boldsymbol{\lambda} \succsim_{h_n}^{\mathbf{y}} \boldsymbol{\mu} \quad \Longleftrightarrow \quad \mathbf{y} \circ \boldsymbol{\lambda} \succsim_{h_n}^{\beta} \mathbf{y} \circ \boldsymbol{\mu}.$$

This relation depends only on $\succsim_{h_n}^{\beta}$ and the fixed segment $\mathbf{y}$. The continuation preference $\succsim_{(h_n;\mathbf{y})}^{\beta}$ also depends only on the successor history $(h_n; \mathbf{y})$. Thus any two auxiliary trees realising the same pair $h_n, (h_n; \mathbf{y})$ generate the same two non-trivial expected-utility preferences on the same mixture space. Cardinal uniqueness therefore gives the same affine coefficients. Define $v_{h_n}(\mathbf{y}) := \widehat{v}_{h_n}(\mathbf{y})$, and $\xi_{h_n}(\mathbf{y}) := \widehat{\xi}_{h_n}(\mathbf{y})$. Since $\mathbf{y} \in \mathbf{Y}_{>h_n}(\mathbb{R} \times Y)$ was arbitrary, this defines functions

$$v_{h_n} : \mathbf{Y}_{>h_n}(\mathbb{R} \times Y) \to \mathbb{R}, \qquad \xi_{h_n} : \mathbf{Y}_{>h_n}(\mathbb{R} \times Y) \to \mathbb{R}_{>0}.$$

On realised proximate segments in the original tree T, these extended coefficients agree with the coefficients already obtained from the local argument, again by uniqueness. This completes the proof.

A.11 Proof of Theorem 5.3

Fix $(h, T) \in \mathcal{D}$, and write $h_n = (h; h(n))$ for each $n \in N^d(T)$. Since T is finite, the decision nodes can be ordered by reverse decision depth. We define $V(n)$ by backward induction. The induction also establishes the following auxiliary claim: $V(n)$ is the maximal u_{h_n}-value attainable in $F(T_{\geq n})$, and, for every $a \in N^{+1}(n)$, $Q_{h_n}(n, a)$ is the maximal u_{h_n}-value attainable in $F_a(T_{\geq n})$.

Consider first a decision node n such that no later decision node follows n. For each $a \in N^{+1}(n)$ and each $\ell \in \mathcal{L}(n, a)$, the arrival node $m(\ell)$ is terminal. Hence $R_{h_n}(\ell) = u_{h_n}(\mathbf{y}(\ell))$. The quantity $Q_{h_n}(n, a)$ is therefore well defined by (11). Since $N^{+1}(n)$ is finite and non-empty, $V(n) := \max_{a \in N^{+1}(n)} Q_{h_n}(n, a)$ is well defined. In this terminal case, once a has been chosen, the transition links in $\mathcal{L}(n, a)$ exhaust the induced chance evolution until termination. Thus $Q_{h_n}(n, a)$ is the maximal u_{h_n}-value attainable in $F_a(T_{\geq n})$, and $V(n)$ is the maximal u_{h_n}-value attainable in $F(T_{\geq n})$.

Now let $n \in N^d(T)$, and suppose that $V(m)$ has already been defined for every decision node m that can be reached from n without passing through another decision node first. Suppose also that, for each such m, $V(m)$ is the maximal u_{h_m}-value attainable in $F(T_{\geq m})$. Fix $a \in N^{+1}(n)$. For each link $\ell \in \mathcal{L}(n,a)$, there are two cases. If $m(\ell) \in N^t(T)$, then $\mathbf{y}(\ell)$ is a complete continuation chain after h_n, so its value is $u_{h_n}(\mathbf{y}(\ell))$. If $m(\ell) \in N^d(T)$, then $h_{m(\ell)} = (h_n; \mathbf{y}(\ell))$. For any continuation lottery $\boldsymbol{\lambda} \in F(T_{\geq m(\ell)})$, Theorem 5.2 gives

$$U_{h_n}(\mathbf{y}(\ell) \circ \boldsymbol{\lambda}) = v_{h_n}(\mathbf{y}(\ell)) + \xi_{h_n}(\mathbf{y}(\ell)) U_{h_{m(\ell)}}(\boldsymbol{\lambda}).$$

Since $\xi_{h_n}(\mathbf{y}(\ell)) > 0$, maximising the value of $\mathbf{y}(\ell) \circ \boldsymbol{\lambda}$ at h_n is equivalent to maximising $U_{h_{m(\ell)}}(\boldsymbol{\lambda})$ at $h_{m(\ell)}$. By the induction hypothesis, the maximal continuation value at $m(\ell)$ is $V(m(\ell))$. Hence the maximal contribution of link ℓ is $R_{h_n}(\ell)$, as defined in (12).

A path strategy specifies choices independently at all decision nodes reached later in the tree. Thus, after the control a has been chosen at n, the maximal value attainable from the induced transition links is the probability-weighted sum in (11). Hence $Q_{h_n}(n,a)$ is the maximal u_{h_n}-value attainable in $F_a(T_{\geq n})$. Taking the maximum over $a \in N^{+1}(n)$ gives (13) and establishes that $V(n)$ is the maximal u_{h_n}-value attainable in $F(T_{\geq n})$. This completes the backward-induction construction. Uniqueness follows from the same induction, since each $V(n)$ is determined by successor-decision values that have already been uniquely determined.

It remains to prove (14). By Proposition 3 and the mixture-continuity hypothesis, Theorem 5.1 applies to $\succsim^{\beta}_{h_n}$. Hence the choice correspondence $\mathcal{C}^{\beta}_{h_n}$ is rationalised by $\succsim^{\beta}_{h_n}$ on finite feasible sets. Lemma 1 then gives

$$\Phi^{\beta}_{h_n}(T_{\geq n}) = \max\left(F(T_{\geq n}), \succsim^{\beta}_{h_n}\right).$$

Fix $a \in \beta_{h_n}(T_{\geq n})$. Since β is non-empty at every decision node in its domain, the immediate choice a can be extended to a β-admissible path strategy $\mathbf{p} \in \mathbf{P}^{\beta}_{h_n}(T_{\geq n})$. Hence

$$\boldsymbol{\eta}(T_{\geq n}, \mathbf{p}) \in \Phi^{\beta}_{h_n}(T_{\geq n}).$$

By the preceding display, $\boldsymbol{\eta}(T_{\geq n}, \mathbf{p})$ is $\succsim^{\beta}_{h_n}$-maximal in $F(T_{\geq n})$. Therefore its u_{h_n}-value is $V(n)$. Since $\mathbf{p}$ chooses a at n, the lottery $\boldsymbol{\eta}(T_{\geq n}, \mathbf{p})$ belongs to $F_a(T_{\geq n})$. Since $Q_{h_n}(n,a)$ is the maximal u_{h_n}-value attainable in $F_a(T_{\geq n})$, we have

$$Q_{h_n}(n,a) \geq V(n).$$

By (13),

$$V(n) = \max_{b \in N^{+1}(n)} Q_{h_n}(n,b),$$

so $Q_{h_n}(n,a) = V(n)$. Hence

$$a \in \arg\max_{b \in N^{+1}(n)} Q_{h_n}(n,b).$$

Since $a \in \beta_{h_n}(T_{\geq n})$ was arbitrary,

$$\beta_{h_n}(T_{\geq n}) \subseteq \arg\max_{a \in N^{+1}(n)} Q_{h_n}(n,a),$$

as required.

A.12 Proof of Corollary 2

Fix $n \in N^d(T)$. By the auxiliary claim established in the backward-induction construction in the proof of Theorem 5.3, $Q_{h_n}(n,a)$ is the maximal u_{h_n}-value attainable in $F_a(T_{\geq n})$, and $V(n)$ is the maximal u_{h_n}-value attainable in $F(T_{\geq n})$. Hence

$$Q_{h_n}(n,[a]_n) = V(n) \iff F_a(T_{\geq n}) \cap \max\left(F(T_{\geq n}), \succsim^{\beta}_{h_n}\right) \neq \varnothing.$$

By Lemma 1,

$$\Phi^{\beta}_{h_n}(T_{\geq n}) = \max\left(F(T_{\geq n}), \succsim^{\beta}_{h_n}\right).$$

Therefore

$$Q_{h_n}(n,[a]_n) = V(n) \iff F_a(T_{\geq n}) \cap \Phi^{\beta}_{h_n}(T_{\geq n}) \neq \varnothing.$$

By definition of $\widehat{\beta}_{h_n}(T_{\geq n})$, this is equivalent to

$$[a]_n \in \widehat{\beta}_{h_n}(T_{\geq n}).$$

Since

$$V(n) = \max_{[b]_n \in \bar{N}^{+1}(n)} Q_{h_n}(n,[b]_n),$$

the claim follows.

References

Anscombe, F. J. and R. J. Aumann (1963): "A definition of subjective probability," *Annals of mathematical statistics*, 34, 199–205.

Arrow, K. J. (1959): "Rational choice functions and orderings," *Economica*, 26, 121–127.

Bandyopadhyay, T. (1988): "Revealed preference theory, ordering and the axiom of sequential path independence," *The Review of Economic Studies*, 55, 343–351.

Bratman, M. (1987): "Intention, plans, and practical reason," *Harvard University Press.*

Campbell, D. E. (1978): "Realization of choice functions," *Econometrica*, 171–180.

Cubitt, R. P. and R. Sugden (2001): "On money pumps," *Games and Economic Behavior*, 37, 121–160.

Davidson, D., J. C. C. McKinsey, and P. Suppes (1955): "Outlines of a formal theory of value, I," *Philosophy of science*, 22, 140–160.

Debreu, G. (1959): *Theory of value: An axiomatic analysis of economic equilibrium*, Yale University Press.

Gustafsson, J. E. (2022): *Money-pump arguments*, Cambridge University Press.

Hammond, P. J. (1977): "Dynamic restrictions on metastatic choice," *Economica*, 44, 337–350.

——— (1988a): "Consequentialism and the Independence Axiom," in *B.R. Munier (ed.) Proceedings of the 3rd International Conference on the Foundations and Applications of Utility, Risk and Decision Theories (Dordrecht: D. Reidel*, Springer, 503–516.

——— (1988b): "Consequentialist foundations for expected utility," *Theory and decision*, 25, 25–78.

——— (1998a): "Objective expected utility: A consequentialist perspective," *Handbook of utility theory*, 1, 143–211.

——— (1998b): "Subjective expected utility," *Handbook of utility theory*, 1, 213–271.

——— (2022): "Prerationality as avoiding predictably regrettable consequences," *Revue économique*, 73, 943–976.

Hammond, P. J. and A. Troccoli-Moretti (2026): "Bayesian Rationality in Finite Decision Trees with Menu Consequences," *Working Paper*.

Jensen, N. E. (1967): "An introduction to Bernoullian utility theory: I. Utility functions," *The Swedish Journal of Economics*, 163–183.

Karni, E. and D. Schmeidler (1991): "Utility theory with uncertainty," *Handbook of mathematical economics*, 4, 1763–1831.

Koopmans, T. C. (1964): "On the Flexibility of Future Preferences," in *Human Judgments and Optimality, ed. by M. W. Shelly and G. L. Bryan.*, Wiley, New York.

Kreps, D. (1988): *Notes on the Theory of Choice*, Westview Press.

Kreps, D. M. (1979): "A representation theorem for "preference for flexibility"," *Econometrica*, 565–577.

——— (2013): *Microeconomic Foundations I: Choice and Competitive Markets*, Princeton university press.

Kreps, D. M. and E. L. Porteus (1978): "Temporal resolution of uncertainty and dynamic choice theory," *Econometrica*, 185–200.

Malinvaud, E. (1952): "Note on von Neumann-Morgenstern's strong independence axiom," *Econometrica*, 20, 679.

Marschak, J. (1950): "Rational behavior, uncertain prospects, and measurable utility," *Econometrica*, 111–141.

Phelps, E. S. and R. A. Pollak (1968): "On second-best national saving and game-equilibrium growth," *The Review of Economic Studies*, 35, 185–199.

Plott, C. R. (1973): "Path independence, rationality, and social choice," *Econometrica*, 1075–1091.

Raiffa, H. (1968): *Decision analysis: Introductory lectures on choices under uncertainty.*, Addison-Wesley.

Samuelson, P. A. (1952): "Probability, utility, and the independence axiom," *Econometrica*, 670–678.

Savage, L. J. (1954): *The foundations of statistics*, John Wiley.

Strotz, R. H. (1956): "Myopia and inconsistency in dynamic utility maximization," *Review of Economic Studies*, 23, 165–180.

Vickrey, W. S. (1964): "Metastatics and macroeconomics," *(No Title)*.

von Neumann, J. (1928): "Zur Theorie der Gesellschaftsspiele," *Mathematische Annalen*, 100, 295–320.

von Neumann, J. and O. Morgenstern (1944): "Theory of games and economic behavior," *Princeton: Princeton University Press*.