Many important choices require decision makers (DMs) to choose between alternatives that they do not fully understand. We model a DM making a choice under uncertainty who may imperfectly understand acts—mappings from states to outcomes. We model coarse understanding of acts using partitions of the state space: for each cell of the partition, the DM knows the set of outcomes that she could receive if the true state lies in that cell, but within each cell she is unable to match states with outcomes. A key feature of the model is that the DM may understand different acts using different partitions, depending on the acts’ specific outcomes. We argue that this allows us to differentiate limited understanding of acts from coarse contingencies and ambiguity aversion, both related phenomena, using only static choice of acts. Our main results axiomatically characterize this model and uniquely identify the partitions that the DM uses to understand acts.

Keywords: limited understanding; complexity; coarse contingencies; ambiguity

JEL Classification: D81, D83
1. **Introduction**

1.1. **Motivation**

The world around us is extremely complex. Most choices require decision makers (DMs) to choose between actions that are uncertain, in that a given action may lead to different outcomes depending on the realized state of nature. In subjective expected utility theory, to assess an action, DMs must determine likelihoods for the states of nature and determine which outcome the action leads to in each state. While much attention has been paid to difficulties in the assessing likelihoods, economists typically assume that the latter determination is costless for DMs. However, many actions lead to a large number of consequences, so realistically DMs may struggle to identify which outcome will come about in each state of nature.

In this paper, we study a decision maker who is constrained in her ability to match actions’ possible outcomes with the states of nature, which she perfectly understands. From choice, we identify the extent to which the DM is constrained and model how she simplifies actions she subjectively views as complex. To fix ideas, consider a DM choosing a new health insurance policy. A state of nature is a collection of treatments the DM may receive while she is covered by the policy, which we will refer to as treatment plans. Each policy is an action, which associates every possible treatment plan with a cost borne by the DM. The DM we model can perfectly describe every treatment plan and can perfectly identify the set of possible costs that each policy may leave her responsible for. However, for a given policy, the DM may be constrained in her ability to match treatment plans with their corresponding payments. This may be the case because the DM simply does not have time to closely read the policy, or because she is cognitively constrained in her ability to understand the finer details of the policy.

How might decision makers simplify actions (acts) that are too complex for them understand? Since complexity here refers to the DM being unable to match states and outcomes for a given act \( f \), the DM may attempt to reduce this complexity by matching a set of states \( E \) with the set of outcomes \( f(E) \) that are possible given the act \( f \). This simplification procedure naturally leads to a partition \( P_f \) of the set of states \( S \) where for each \( E \in P_f \), the DM views outcomes in \( f(E) \) as possible. Finer partitions correspond to understanding more details, or reading more closely. Indeed, the “standard” case—where the DM fully understands acts—corresponds to the special case in which each \( E \in P_f \) is a single state. This simplification procedure may depend on the structure of \( f \), so the DM may simplify different acts in different ways.

To illustrate, we can revisit the health insurance example. Suppose the DM has injured her

---

For simplicity, suppose that the DM has no choice in which treatments she receives. This may be the case if they are prescribed by a benevolent doctor and all the relevant treatments are necessary for the DM’s survival.
knee playing basketball. To repair it, she needs one of two possible surgeries. Following the surgery, she will be on one of two different rehabilitation plans that are common to both possible surgeries. Suppose that prior to her surgery, the DM needs to choose a new health insurance policy. Additionally, suppose for simplicity that the DM is certain she will not need any other medical care for the duration of the plan. As such, the states of the world are given by

\[ S = \{s_1, s_2\} \times \{r_1, r_2\} = \{(s_1, r_1), (s_1, r_2), (s_2, r_1), (s_2, r_2)\}, \]

where \(s_1\) refers to the first possible surgery, \(r_2\) refers to the second possible rehab plan, and so on. In this case, an insurance policy is an act that describes the DM’s total required payment in each of the four states above. The first insurance policy available to the DM fully covers \(s_2\) and \(r_2\), but \(s_1\) and \(r_1\) cost the DM $50 each. This policy is the act \(f = (100, 50, 50, 0)\), where the first entry corresponds to the DM’s total payment in state \((s_1, r_1)\), the second corresponds to state \((s_1, r_2)\), etc. However, the DM needs to read the policy closely to know that \(f((s_1, r_1)) = 100\), and so on. If she is unable to do so, she may simplify \(f\) using the partition \(P = \{(s_1, r_1), (s_2, r_1)\};\{(s_1, r_2), (s_2, r_2)\}\), i.e. she views the plan \(f\) as costing $50 or $100 in states \((s_1, r_1)\) and \((s_2, r_1)\), and $0 or $50 in states \((s_1, r_2)\) and \((s_2, r_2)\). This might correspond to the DM “focusing” on the rehab, coarsely understanding the policy as “rehab one will cost me at most $100, otherwise I’ll pay at most $50”. Keep in mind that the DM’s understanding is subjective, and other DM’s may simplify \(f\) in a different way. Moreover, we interpret \(P\) as how the DM is understanding \(f\) at the time of choice. She does not reason further about \(f\)—the partition \(P\) is the outcome of any reasoning she has done about \(f\).

The second insurance policy available to the DM fully covers both rehab plans and surgery two, but requires a $100 payment for surgery one. This policy is the act \(g = (100, 100, 0, 0)\). In this case, the DM may simplify \(g\) using the partition \(Q = \{(s_1, r_1), (s_1, r_2)\};\{(s_2, r_1), (s_2, r_2)\}\), i.e. she views \(g\) as costing $100 in states \((s_1, r_1)\) and \((s_1, r_2)\), and $0 in states \((s_2, r_1)\) and \((s_2, r_2)\). Notice that in this case, even though the DM may have limited understanding of acts, her ability to simplify \(g\) using the partition \(Q\) is sufficient to perfectly understand \(g\) since it is \(Q\)-measurable. In this sense, the DM views \(g\) as a “simple” policy because her constraints do not prevent her from understanding it. We refer to the partitions that the DM may use to understand acts, like \(P\) and \(Q\) above, as understanding strategies.
In the sequel, we provide an axiomatic model of a DM who has a limited capacity to understand acts, like the DM above. When the DM views an act as simple, she is an expected utility agent. When an act is complex, she simplifies it using an understanding strategy and then computes the act’s expected utility. One of our model’s key hypotheses is that if $P$ is an understanding strategy, then the DM views $P$-measurable acts as simple (as with the policy $g$ and partition $Q$ above). As such, we identify the understanding strategies from choice by looking at the set of partitions $P$ such that the DM’s preference satisfies the Independence axiom when choosing between $P$-measurable acts. In other words, if the DM violates Independence for some triple $f, g, h$ of $P$-measurable acts, then the DM must not be able to understand acts using the strategy $P$.\footnote{Notice that this hypothesis rules out the possibility that the DM does not “notice” that an act is $P$-measurable when $P$ is an understanding strategy.}

Our axiomatic analysis identifies a novel behavior that reveals the DM’s uneasiness with, or aversion to, her limited understanding. Given an act $f$, we show in Section 3 how to identify the understanding strategy $P$ the DM is using when assessing $f$. As above, given an event $E \in P$, the DM can only see the set of outcomes $f(E)$. Suppose that for some constant outcome $x$, we observe the DM’s preference to be

$$f > xE f,$$

where $xE f$ is the act $f$ outside of $E$ but yields $x$ for every $s \in E$. Since the DM’s limited understanding lies in her inability to match states and outcomes, $xE f$ is as simple as possible on $E$, and the same as $f$ elsewhere. As such, if the DM is indeed averse to her limited understanding, we expect that she would not reverse her preference when $xE f$ is made more complicated, i.e. she also exhibits the preference

$$f > xA f$$

for any $A \subseteq E$.

A slightly weaker version of the above behavior displaying aversion to limited understanding, along with other axioms that are implied by the subjective expected utility model (with a finite state space and full support prior), characterizes the representation

$$V(f) = \max_{P \in \mathcal{P}} \sum_{E \in P} \pi(E) \min_{s \in E} u(f(s)),$$

where $\mathcal{P}$ is the set of understanding strategies identified using the procedure described above, $u$ is a vNM utility function and $\pi$ is a suitably defined set function representing the DM’s beliefs. The representation suggests that the DM first “chooses” a partition $P \in \mathcal{P}$ to use to simplify the act $f$. Then for each $E \in P$, she knows that the possible outcomes of $f$ are given by $f(E)$. Since the DM is aware that her reading of the act is incomplete, she proceeds cautiously by considering only the
worst outcome in $f(E)$ when computing her expected utility of $f$ given $P$. In other words, she displays an aversion to her limited understanding. We discuss the representation further following its formal presentation in Definition 2.

Our main contribution is twofold. The standard subjective expected utility model (Savage (1954) and Anscombe and Aumann (1963)) suggests that DMs understand the choice environment well enough to fully conceptualize the state space and form a single prior over it. There is an enormous literature on DMs’ inability to form a single prior (ambiguity), a smaller literature on DMs who may be unable to conceptualize the state space specified by the analyst (coarse contingencies), and some work discussing how to distinguish the two (see Section 1.2 for more details on the relevant papers). Our first contribution is to observe a third way that DMs may not understand their environment. Namely, even when a DM may be able to fully conceptualize the state space and form a prior, acts themselves may be difficult to understand, as in the examples presented above. Our second contribution is to provide a parsimonious model of this form of limited understanding. Since the phenomenon we are modeling is uncertainty about the very structure of acts themselves, it would be natural to encode this uncertainty by expanding the state space to include every possible “true act” that an act could be, and supposing the DM has a prior over each of these enlarged state spaces. However, this would require a different state space for each act, resulting in a very cumbersome model. Moreover, this proposed model provides no guidance on how the DM’s beliefs should vary with the act. On the other hand, the model presented here identifies a belief over the traditional state space, and identifies how the agent understands each act using only the usual primitives (a preference over acts defined on a common state space).

1.2. Relationship with Coarse Contingencies

The DM we are modeling perfectly understands the state space $S$ and the set of outcomes that could arise following any act $f$. However, the DM may coarsely understand the mapping $f$, and thus identify each cell of the understanding strategy used to understand $f$ with a set of outcomes. As such, the DM may appear to coarsely understand the state space. The key difference between coarse understanding of the state space as it is usually modeled and apparent coarseness resulting from limited understanding of acts is that in the former case the DM’s coarseness is usually described by a single partition, while in the latter the DM may use different partitions to understand different acts, as we described above. In this subsection, we briefly discuss some related work on coarse understanding of the state space and how it differs from our model.

A number of papers have focused on the relationship between coarse understanding of $S$
(or coarse contingencies) and ambiguity, beginning with Mukerji (1997) and Ghirardato (2001).³ Mukerji (1997) takes both the analyst’s and the DM’s state spaces as observable and models the mapping between them. He considers the DM’s preference over acts defined on the analyst’s state space and shows that under some conditions it may be represented by a Choquet integral (a model of ambiguity sensitive preferences developed in Schmeidler (1989)). Ghirardato (2001) takes as primitive a preference over correspondences whose domain is the analyst’s state space $S$. The fact that choice objects are correspondences instead of functions captures the DM’s coarse understanding of the state space (it could also be interpreted as a model of limited understanding of acts). He shows that under natural adaptations of the Savage (1954) axioms, the DM’s preference may also be represented by a Choquet integral. A key difference between this paper, Mukerji (1997), and Ghirardato (2001) is that the latter two papers take the DM’s coarse understanding to be observable, while one of our main contributions is identifying the how the DM understands acts from choice. Indeed, Ghirardato writes that identifying the DM’s perception of acts is “arguably...the most interesting part of the exercise” (p. 251).

In the menu choice literature following Kreps (1979) and Dekel, Lipman, and Rustichini (2001, henceforth DLR), Epstein, Marinacci, and Seo (2007) argue that coarse contingencies would induce a DM to strictly prefer randomizing between two menus, as in the ambiguity aversion literature. Relying on this intuition, Epstein et al. (2007) provide two models of coarse contingencies within the DLR framework that can be interpreted as subjective state space analogs of the maxmin expected utility model of Gilboa and Schmeidler (1989). Aside from the primitive, the major difference between Epstein et al. (2007) and this paper is that in the former paper the DM’s state space is fixed and coarseness manifests via multiple priors. In this paper, the “coarse state space” (partition) varies with the action and the DM has a single belief.


The fact that coarseness arises in our framework due to limited understanding of acts allows us to distinguish ambiguity aversion from coarse contingencies using only static choice between acts. We are able to do so because limited understanding of acts naturally leads to multiple “coarse state spaces” (i.e. partitions), which relaxes the common assumption that the DM’s coarse state space

³Each of these papers build on Gilboa and Schmeidler (1994), which offers an interpretation of Choquet expected utility as modeling a DM completing a misspecified model.
space is independent of the act she is considering.\textsuperscript{4} To see how we differentiate coarseness from ambiguity aversion, recall that ambiguity aversion is typically identified with a preference for hedging, or smoothing out payoffs across states.\textsuperscript{5} Now, return to the health insurance example above, and consider a third policy $h = (100, 0, 100, 0)$; i.e. $h$ charges $100$ for $r_1$ and fully covers $s_1$, $s_2$, and $r_2$. Notice that $h$ is $P$-measurable, so the DM fully understands $h$ given our hypothesis. Suppose the DM is indifferent between $h$ and $g$, which is also perfectly understood since it is $Q$-measurable. Suppose further that $P$ and $Q$ are the DM’s only understanding strategies. In this case, notice that policy $f$, which smooths out the DM’s payments across states, is not understood. As such, if the DM is averse to her limited understanding, she may strictly prefer $g$ or $h$ to $f$, which violates the weak preference for hedging usually identified with ambiguity aversion.

The key intuition in the previous example is that mixing two acts that are measurable with respect to two different understanding strategies creates a more complicated act in the sense that the mixture is measurable with respect to a finer partition. Since this finer partition may not be a feasible understanding strategy, the DM may be unable to understand the mixture and thus may disprefer it. This suggests that the distinction between limited understanding and ambiguity aversion relies crucially on the multiplicity of the DM’s partitions, which is a key feature of our model. If the DM only has a single understanding strategy, every act is understood using this partition, and mixing two indifferent acts weakly improves them, as it makes them closer to constant on each cell of the DM’s understanding strategy. In this sense, limited understanding of acts manifests itself as ambiguity when the DM has a single understanding strategy, even though the DM has a single belief. We explore this idea formally in Section 4.3.

Two final notes on related literature: First, while Ahn and Ergin (2010) are concerned with framing and not limited understanding, there are some connections between this paper and their paper on framing, which we discuss in Section 4.6. Second, Gul, Pesendorfer, and Strzalecki (2017) study a general equilibrium model where consumers are restricted to choose consumption plans that are measurable with respect to a partition in some exogenously specified set. Our paper can be understood as providing foundations for a model of such consumers and showing how to identify which partitions are actually feasible for consumers (i.e. endogenizing the set of partitions).

The paper proceeds as follows: In Section 2, we state our primitives and formally define the representation. Section 3 presents the axiomatic model. Our main results are in Section 4, including further discussion about the relationship between limited understanding and ambiguity, as well as more formal comparisons to the literature.

\textsuperscript{4}This is the case in Mukerji (1997), Ghirardato (2001), Epstein et al. (2007), Kochov (2018), and Minardi and Savochkin (2017).

\textsuperscript{5}See Section 4.3 for a formal definition of uncertainty aversion.
2. Setup & Model

In this section we describe the choice environment and formally present the utility representation we intend to characterize.

2.1. Primitives

We use a version of the classic Anscombe and Aumann (1963) choice setting. Let $S$ denote a nonempty finite set of states of the world, and let $\Sigma := 2^S$ be the power set of all events—subsets of $S$. Let $X$ be the set of outcomes, which we assume to be a convex subset of a metrizable vector space. For example, this is the case if $X$ is the set of simple probability distributions over some prize set $Z$, as in the textbook presentation of the Anscombe and Aumann (1963) model (Kreps, 1988).

An act is function from $S$ to $X$. Let $\mathcal{F} := X^S$ be the set of all acts, endowed with the product topology. Since $X$ is convex, we define mixtures in $\mathcal{F}$ pointwise: for any $f, g \in \mathcal{F}$ and $\alpha \in (0, 1)$, define $\alpha f + (1 - \alpha)g \in \mathcal{F}$ as $(\alpha f + (1 - \alpha)g)(s) := \alpha f(s) + (1 - \alpha)g(s)$ for every $s \in S$. For any event $E \in \Sigma$ and acts $f, g \in \mathcal{F}$, we use $fE$ to denote the act that yields $f(s)$ if $s \in E$ and $g(s)$ otherwise. Additionally, for any event $E \in \Sigma$ and act $f \in \mathcal{F}$, let $f(E) := \{f(s) | s \in E\}$.

The primitive is the DM’s preference on $\mathcal{F}$, a binary relation denoted $\succ$. We say that a function $V : \mathcal{F} \rightarrow \mathbb{R}$ represents $\succ$ if for every $f, g \in \mathcal{F}$, $f \succ g \iff V(f) \geq V(g)$.

Two remarks are in order regarding the interpretation of the primitive. First, notice that we are using the language of states, outcomes, and acts that is common and well understood among economists. If the DM’s understanding of acts is limited, then there are acts $f$ that the DM does not view as the same function that the analyst does. However, every act $f$ corresponds to real world alternative that the DM may choose, even if she understands it differently than the analyst. In our introductory example, the insurance plan $f$ is not understood by the DM in the same way that the analyst understands it, but the DM can choose that plan nonetheless. Indeed, a main objective of this paper is to elicit how the DM understands actions using choice data.

Second, we assume throughout that the DM is standard aside from her limited ability to understand acts. Since constant acts are trivial in the sense that they map every state to the same outcome, we assume that the DM is standard when evaluating constant acts in that her limited understanding does not affect her preferences over outcomes. This is natural because no matter how the DM chooses to understand a constant act, for each event she looks at there will only be one outcome and thus she will conclude that the act is constant. Equivalently, since the DM

---

6Since $X$ is metrizable and $S$ is finite, the product topology is equivalent to the topology induced by the supnorm metric on $\mathcal{F}$ defined by $d(f, g) = \max_{s \in S} \hat{d}(f(s), g(s))$ where $\hat{d}$ is a metric on $X$. 
can see the set of outcomes associated with each act, she knows that constant acts are constant because they only lead to one outcome.

2.2. Model

Before defining and discussing the model, we will introduce a bit more notation. A partition of $S$ is a collection of nonempty, pairwise disjoint events in $\Sigma$ whose union is equal to $S$. Given partitions $P$ and $Q$, we write $P \gg Q$ to denote that $P$ is (strictly) finer than $Q$. Since partitions can be understood as functions from $S$ to $\Sigma$, we write $P(s)$ to denote the unique cell in $P$ that contains $s$. Given a partition $P$, let $\sigma(P)$ denote the algebra generated by $P$. In other words, $\sigma(P)$ is the collection of events that can be formed by taking unions of events in $P$ (with the empty set added in). Given a set of partitions $\mathcal{P}$, let

$$\mathcal{E}(\mathcal{P}) := \bigcup_{P \in \mathcal{P}} \sigma(P)$$

denote the set of events that are in the algebra generated by some partition in $\mathcal{P}$. Note that in the absence of any further assumptions on $\mathcal{P}$, $\mathcal{E}(\mathcal{P})$ is in general not an algebra itself, as it need not be closed under unions or intersections. The following definition will be used in the formal definition of the representation below.

**Definition 1.** A set of partitions $\mathcal{P}$ is rich if for any $P, Q \in \mathcal{P}$, $A \in \sigma(P)$, and $B \in \sigma(Q)$ such that $A \subset B$, there exists $R \in \mathcal{P}$ such that $B \setminus A, A \in \sigma(R)$.

Loosely, a set of partitions $\mathcal{P}$ is rich if it is closed under a type of combination operation. To gain intuition, fix $P, Q \in \mathcal{P}$. The join, or coarsest common refinement, of $P$ and $Q$, denoted $P \vee Q$, is the set of all events $E$ such that $E = A \cap B$ for some $A \in P$ and $B \in Q$. If $\mathcal{P}$ were closed under the join operation, then for any $P, Q \in \mathcal{P}$, $A \in \sigma(P)$, and $B \in \sigma(Q)$ such that $A \subset B$, it follows that $A, B \setminus A \in \sigma(P \vee Q)$; so $P \vee Q$ plays the role of $R$ in the definition above. However, $P \vee Q$ also consists of intersections of events in $P$ and $Q$ that may not be nested. Requiring that $\mathcal{P}$ be rich instead of closed under the join operation relaxes this second requirement, by allowing the partition $R$ to be arbitrary on $B^c$. Following the representation definition, we discuss richness in the context of the model.

Lastly, given a set of events $\mathcal{A} \subseteq \Sigma$, say that a function $\pi : \mathcal{A} \to [0, 1]$ is normalized if $\pi(S) = 1$ and $\pi(\emptyset) = 0$ whenever $\emptyset, S \in \mathcal{A}$, additive if $\pi(A \cup B) = \pi(A) + \pi(B)$ for every disjoint $A, B \in \mathcal{A}$ such that $A \cup B \in \mathcal{A}$, and strictly monotone if $A \subset B$ implies $\pi(A) < \pi(B)$ for every $A, B \in \mathcal{A}$.

**Definition 2.** A preference $\succeq$ has a **revealed understanding** representation if there exists

---

7. $P \gg Q$ if for every $E \in P$, there exists $E' \in Q$ such that $E \subseteq E'$, and at least one of the set inclusions is strict.

8. Throughout, we use $\subset$ to denote a proper subset. When equality is permitted, $\subseteq$ is used.
• a continuous, mixture linear function \( u : X \rightarrow \mathbb{R} \),

• a nonempty, rich set of partitions \( \mathcal{P} \),

• and a normalized, additive, strictly monotone set function \( \pi : \mathcal{E}(\mathcal{P}) \rightarrow [0, 1] \),

such that

\[
V(f) = \max_{P \in \mathcal{P}} \sum_{E \in P} \pi(E) \min_{s \in E} u(f(s))
\]  

(1)

represents \( \gtrapprox \).

We refer to the revealed understanding (RU) representation as a tuple \( \langle u, \pi, \mathcal{P} \rangle \). The DM’s tastes and beliefs are represented by \( u \) and \( \pi \) respectively, as in the standard model. The set of partitions \( \mathcal{P} \) captures the different ways in which the DM can understand actions.

The representation suggests that the DM coarsely understands actions, and simplifies them using a partition \( P \in \mathcal{P} \). More specifically, given an act \( f \) and a partition \( P \) that solves (1), for each \( E \in P \), the DM knows that \( f \) could lead to any outcome in \( f(E) \). The DM is aware that she is missing details regarding the mapping within each \( E \), so she proceeds cautiously by assigning the utility value \( \min_{s \in E} u(f(s)) \) to each \( E \in P \), i.e. she is averse to her limited understanding. She then combines these utility values with her beliefs \( \{\pi(E)\}_{E \in \mathcal{P}} \) using an expected utility calculation. We offer two possible justifications for the DM’s extremely cautious attitude regarding her limited understanding. First, our motivation primarily stems from situations in which the state space (and thus events in these partitions) would be large relative to the sets of outcomes. Thus, we view the “min” component as a parsimonious approximation to the DM minimizing over all the possible functions mapping a large event \( E \) to a (relatively smaller) set of outcomes \( f(E) \). Second, one could interpret it through an evolutionary lens. If the DM used alternative procedure, self interested agents who are setting the DM’s menu (e.g. insurance companies) would over time be able to extract large rents. Thus, over time one would expect the DM to “learn” that the worst case outcome in each event is the most important consideration when evaluating acts she coarsely understands.

The set \( \mathcal{P} \) being possibly nonsingleton is the key feature of the model that is particular to our interpretation of the DM as one who does not understand acts. As we argued in Section 1, this is the distinguishing feature between a DM who does not understand acts and a DM who does not understand the state space. In the latter case, one would expect that if the DM could understand the contingencies in \( P \) and the contingencies in \( Q \), then the contingencies in \( P \lor Q \) should also be understood. Iterating this reasoning would result in a singleton \( \mathcal{P} \).\footnote{Without loss of generality, see Theorem 2.} However, the DM we
are modeling understands the state space, and the partitions are “tools” she uses to simplify or better understand acts. As such, it is natural that \( P \) is not closed under the join operation, since understanding an act’s possible outcomes for each event in \( P \lor Q \) is more demanding than doing so for \( P \) or \( Q \) individually. In our introductory example, the DM understood policy \( f \) by thinking only about which rehab she would need, and understood policy \( g \) by thinking only about which surgery she would need. However, she may nonetheless be unable to understand policies for which each surgery and rehab plan require distinct payments, since that would require a closer examination of the policy.

While one could imagine more exotic ways for the DM to “choose” a partition for each act, the “max” operator is parsimonious and has two nice features. First, it interacts with the cautious attitude within cells nicely; any solution to (1) will be as fine as possible. More formally, if \( P, Q \in P \) and \( P \gg Q \), if \( Q \) solves (1), then so must \( P \). This fact also suggests an awareness regarding the DM’s limited understanding, as the DM tries to understand the act to the best of her ability. A corollary of this fact is that the DM’s limited understanding only causes her to depart from expected utility when acts cannot be expressed in a way the agent can understand. More specifically, if \( f \) is \( P \)-measurable for some \( P \in P \), then \( P \) is a solution to (1) and the utility of \( f \) is simply the expected utility of \( f \) computed with \( \pi \).\(^{10}\) Thus, the DM’s limited understanding only affects her evaluation of acts that are not \( P \)-measurable for all \( P \in P \). As such, if \( P \) contains the partition of all singletons \( \{\{s\}_{s \in S}\} \), then the DM is an expected utility agent.

We argued above that the DM may not be able to “combine” partitions, in the sense that \( P \) may not be closed under the join operation. However, the representation requires that \( P \) be rich, which we described above as a weaker requirement than \( P \) being closed under the join operation. To see why this is a natural condition given our story, first recall that richness requires that if there exists \( P, Q \in P \), \( A \in \sigma(P) \), and \( B \in \sigma(Q) \) such that \( A \subset B \), then there exists \( R \in P \) such that \( A \) and \( B \setminus A \) are in \( \sigma(R) \). Richness is natural in this setting because we interpret the partitions in \( P \) as different ways that the DM can understand acts. In particular, since \( B \in \sigma(Q) \) and \( Q \in P \), the DM can understand an act \( f \) by considering the sets of outcomes \( f(B) \) and \( f(B^c) \) (and similarly for \( A \)). Since \( A \subset B \), if the DM is understanding \( f \) by looking at \( f(B) \), the DM should be able to “focus” on \( A \) within \( B \), i.e. splitting \( f(B) \) into \( f(A) \) and \( f(B \setminus A) \). This corresponds to the partition \( R \) in the definition of richness. It is important to note that \( R \) may be very coarse on \( B^c \), which permits this “focusing” to be costly for the DM. This would not be possible if \( P \) were closed under the join operation, which is why we impose only richness. In particular, since \( A \subset B \), it follows that \( B^c \subset A^c \). So richness also requires that there exist \( R' \in P \) such that \( B^c, A^c \setminus B^c \in \sigma(R') \), but does not imply that \( R = R' \). In other words, the DM may face a tradeoff in choosing whether to focus on \( A \) within \( B \) or \( B^c \) within \( A^c \). Moreover, requiring that \( P \) be closed under the join

\(^{10}\)See Lemma 7 for a proof of this fact.
operation would require that if \( A \) and \( B \) are as above but not nested, then the DM is able to split them into \( A \setminus B, B \setminus A, \) and \( A \cap B, \) i.e. there is some \( R'' \in \mathcal{P} \) such that all three events are in \( \sigma(R'') \).

However, to perform this “split”, the DM would need to look at \( A \) and \( B \) simultaneously, which she is only able to do if there exists \( P' \in \mathcal{P} \) such that \( A \cup B \in \sigma(P') \). If this is true, then this type of focusing is implied by richness. If it is not, then requiring that \( \mathcal{P} \) be closed under the join operation supposes the DM has more reasoning ability than she has revealed. Aside from richness, the set \( \mathcal{P} \) is arbitrary, but in applications there may be natural parametrizations, e.g. the set of all partitions with \( k \) or fewer cells, where \( k \in \mathbb{N} \).

Lastly, we note that while “inattention” typically refers to other things in the decision theory literature, the RU representation could be interpreted as modeling a DM who is attentive to acts’ outcomes on each state. In this interpretation, the partitions in \( \mathcal{P} \) describe which parts of the state space the DM can pay closer attention to (and thus know what the outcomes are on those events), and which tradeoffs the DM faces in allocating her attention to different parts of acts’ outcomes. The “min” part of the representation then corresponds to the DM’s caution toward her inattention (e.g. she wants to avoid low payoff surprises), while the “max” part corresponds to her wanting to distort the act as little as possible. As we mentioned, this differs from the two types of inattention typically studied: inattention to information (e.g. de Oliveira, Denti, Mihm, and Ozbek (2017); Ellis (2018)) and inattention to items in the choice set (e.g. Masatlioglu, Nakajima, and Ozbay (2012); Lleras, Masatlioglu, Nakajima, and Ozbay (2017)).

3. Axioms

In this section we present our behavioral model. The first axiom collects some standard behaviors common to many models of subjective uncertainty.

**Axiom 1.**

(i) (Weak Order) \( \preceq \) is complete and transitive.

(ii) (Continuity) For every \( f \in \mathcal{F}, \) the sets \( \{ g \mid g \succeq f \} \) and \( \{ g \mid f \succeq g \} \) are closed.

(iii) (Certainty Independence) For every \( f, g \in \mathcal{F}, \) \( x \in X, \) and \( \alpha \in (0, 1), \)

\[
f \succeq g \iff \alpha f + (1 - \alpha)x \succeq \alpha g + (1 - \alpha)x.
\]

\(^{11}\)This parametrization is compatible with richness. For example, if \( |S| = 4, \) the set of all partitions with \( k = 2 \) cells is \( \mathcal{P} = \{ \{s_1, s_2, s_3\}, \{s_1, s_4\} \} \cup \{ \{s_2, s_3\}, \{s_2, s_4\} \} \cup \{ \{s_1, s_4\}, \{s_2, s_3\} \}, \) which is rich.

\(^{12}\)Similarly, one could view acts as products with multiple attributes, and \( \pi \) as the DM’s subjective weights on the importance of each set of attributes. In this setting, the inattention interpretation of the model would describe a DM who cannot pay attention to each attribute of a product, so instead maps the product into a simpler one.
Parts (i) and (ii) are standard. Part (i) requires that the preference is a weak order, so the DM’s choices from menus of acts would satisfy the Weak Axiom of Revealed Preference (WARP). We have hypothesized that the DM views each act in isolation when trying to understand them because the process of understanding may depend on the acts’ possible outcomes. As such, the model naturally rules out any menu effects. This is another feature of our model that distinguishes it from coarse understanding as it is traditionally understood, as some models of coarse contingencies imply that the DM may violate WARP. Indeed, the ex post (stochastic) choice from menus implied the model in Epstein et al. (2007) may violate (the stochastic choice analog of) WARP because the ambiguity implied by coarse contingencies in their model causes the DM’s beliefs to change with the menu she is considering.\textsuperscript{13} We view part (ii) as a technical condition requiring that the DM’s preference be continuous, so we will not discuss it at any length.

Part (iii) requires that the DM satisfies \textit{Certainty Independence}, due to Gilboa and Schmeidler (1989). This axiom requires that the DM’s preference between two acts is unchanged if they are each mixed with a common outcome using the same mixing coefficient. The reason we find this axiom appealing in this setting is that a DM who may not understand acts will nonetheless be indifferent to the timing of the resolution of uncertainty when that uncertainty is regarding mixing with some common outcome. To see this more formally, let $\alpha f \oplus (1 - \alpha)x$ denote the mixture of $f$ and $x$ where the uncertainty is resolved \textit{before} the DM makes a choice, and similarly for $ag \oplus (1 - \alpha)x$. If $f \succeq g$, then it is natural to assume that $\alpha f \oplus (1 - \alpha)x \succeq ag \oplus (1 - \alpha)x$, in line with the usual intuition regarding the Independence axiom (keeping in mind that this preference is hypothetical, as these objects are not in our domain of preference). The key to the argument underlying \textit{Certainty Independence} is that the DM would be indifferent between $\alpha f \oplus (1 - \alpha)x$ and $\alpha f + (1 - \alpha)x$, where the randomization occurs after the choice is made. To see why, first observe that in the former case, the DM will only concern herself with understanding $f$ because she perfectly understands the outcome $x$ in the event that the randomization results in $x$. Moreover, our ongoing interpretation of the model is that the DM perfectly understands the set of outcomes that may arise following an action. As such, in the latter case, the DM can see that every outcome of $\alpha f + (1 - \alpha)x$ is a mixture between some outcome in $f(S)$ and $x$, all with the same probability of receiving $x$. Thus, she will ignore the probability $1 - \alpha$ event in which she receives $x$ regardless of the state, and focus on the rest of the act. In this case, she again will only concern herself with understanding $f$. As such, we expect that she will have the same sets of outcomes in mind when considering $\alpha f \oplus (1 - \alpha)x$ and $\alpha f + (1 - \alpha)x$, and thus be indifferent between the two.

Notice that the argument above hinges crucially on $x$ being state independent. If $x$ were replaced with some arbitrary act $h \in \mathcal{F}$, then we would expect the DM to prefer early resolution.

\textsuperscript{13}WARP violations are not common to every model of coarse contingencies, e.g. Kochov (2018) and Minardi and Savochkin (2017) both model a DM whose preference is a weak order.
of uncertainty, since in this case the DM knows what action she is evaluating when determining what events to focus on. For this reason we do not assume that the DM’s preference satisfies the usual Independence axiom. Lastly, notice that *Certainty Independence* implies that the DM satisfies Independence on \( X \), which is consistent with our interpretation of the model in which the DM understands the pure outcomes part of her environment perfectly.

We now move on to identifying the DM’s understanding strategies from choice. For any partition \( P \), let \( \mathcal{F}_P \subseteq \mathcal{F} \) denote the set of acts that are \( P \)-measurable, i.e. \( f \in \mathcal{F}_P \) if and only if \( f(s) = f(s') \) for all \( s, s' \in E \) and \( E \in P \). As above, the DM may violate the Independence axiom when the acts being mixed are such that the DM is using different understanding strategies when considering the acts in isolation. However, we would not expect such violations if the DM is using the same understanding strategy. Additionally, if a partition \( P \) is an understanding strategy, we have hypothesized that the DM will use this strategy when evaluating \( P \)-measurable acts. Combining these two observations yields our strategy for identifying understanding strategies: A partition \( P \) is an understanding strategy if the DM’s preference satisfies the Independence axiom when it is restricted to \( P \)-measurable acts. In other words, a partition \( P \) is an understanding strategy if the DM does not depart from expected utility when choosing between acts that only require the DM to understand \( P \) to evaluate. This is borne out in the following definition.\(^{14}\)

**Definition 3.** A partition \( P \) is an understanding strategy if for all \( f, g, h \in \mathcal{F}_P \) and \( \alpha \in (0, 1) \),

\[
    f \succeq g \iff \alpha f + (1 - \alpha)h \succeq \alpha g + (1 - \alpha)h.
\]

We use \( \mathcal{P}_\succsim \) to denote the set of all understanding strategies identified from the preference in this fashion. If \( P \succ Q \) and \( P \) is an understanding strategy, then \( Q \) must also be an understanding strategy since \( \mathcal{F}_Q \subseteq \mathcal{F}_P \). Notice that \( \mathcal{P}_\succsim \) is nonempty because \( \succeq \) satisfies Independence on \( X \) by *Certainty Independence*, so the trivial partition \( \{S\} \in \mathcal{P}_\succsim \).

Applying the above intuition a second time, we can identify from choice which understanding strategy the DM is using when evaluating a given act. More formally, fix any \( f \in \mathcal{F} \). If the DM is using the understanding strategy \( P \) to evaluate \( f \), then the DM’s preference between \( f \) and acts in \( \mathcal{F}_P \) should not depart from expected utility. In particular, if \( g, h \in \mathcal{F}_P \) and \( f \succ g \), then mixing \( f \) and \( g \) with \( h \) should not result in a preference reversal, since the DM was using the understanding strategy \( P \) when evaluating each of the three acts in isolation. To see why, first notice that \( \alpha g + (1 - \alpha)h \) is perfectly understood by the DM for any \( \alpha \), since \( P \in \mathcal{P}_\succsim \) and the mixture is \( P \)-measurable. If the DM is using \( P \) when evaluating \( f \), then the DM sees the set of outcomes \( \{f(E)\}_{E \in P} \). When \( f \) is mixed with \( h \), the DM would then see the set of outcomes are all of the

\(^{14}\)See Ghirardato, Maccheroni, and Marinacci (2004) for a related idea in an ambiguity setting.
form \( \{ \alpha f(E) + (1 - \alpha)h(E) \}_{E \in \mathcal{F}} \), where each \( h(E) \) is a single outcome since \( h \) is \( P \)-measurable.\(^{15}\) As such, the usual intuition suggests that the DM knows that the choice between the two mixtures is again a choice between \( f \) and \( g \). The following definition formalizes this intuition.

**Definition 4.** An act \( f \in \mathcal{F} \) is understood using \( P \in \mathcal{P}_\Sigma \) if for all \( g, h \in \mathcal{F}_P \) and \( \alpha \in (0, 1) \),

\[
f \succ g \implies \alpha f + (1 - \alpha)h > \alpha g + (1 - \alpha)h,
\]

and there is no \( Q \in \mathcal{P}_\Sigma \) with \( Q \gg P \) satisfying (2).

We let \( \mathcal{U}(P) \subseteq \mathcal{F} \) denote the set of all acts that are understood using \( P \). The requirement that there is no \( Q \gg P \) that satisfies the definition is what makes the definition sharp, for if \( P \) satisfies (2), then so does any \( R \ll P \) since \( \mathcal{F}_R \subseteq \mathcal{F}_P \). Interpreting \( \mathcal{U} \) as a correspondence, let \( \mathcal{U}^{-1}(f) \) denote the set of \( P \in \mathcal{P}_\Sigma \) such that \( f \in \mathcal{U}(P) \). We later show (Lemma 1) that every act is understood using some \( P \in \mathcal{P}_\Sigma \).

Noticeably absent from our analysis to this point is a statewise monotonicity condition, which would require that if \( f(s) \succ g(s) \) for all \( s \in S \), then \( f \succ g \). Since the DM may not understand acts finely enough to notice statewise dominance, we refrain from imposing this axiom on the DM’s preference. In this setting, the natural monotonicity condition involves the sets of outcomes that the DM sees when evaluating a given act. More specifically, if \( g \) is understood using \( P \) and for every \( E \in P \), some act \( f \) dominates \( g \) in the sense that the worst outcome in \( f(E) \) is better than the best outcome in \( g(E) \), then the DM should notice this dominance and weakly prefer \( f \). To formalize this, given any \( E \in \Sigma \) and \( f \in \mathcal{F} \), let \( b^E_f \in f(E) \) be the outcome such that \( b^E_f \succeq f(s) \) for all \( s \in E \). Define \( w^E_f \in f(E) \) analogously with the reverse preference.

**Axiom 2 (Understanding Monotonicity).** For every \( P \in \mathcal{P}_\Sigma \), \( g \in \mathcal{U}(P) \), and \( f \in \mathcal{F} \), if \( w^E_f \succeq b^E_g \) for every \( E \in P \), then \( f \succeq g \).

As described above, the axiom requires that if \( g \) is understood using \( P \), and \( f \) dominates \( g \) in the strong sense that \( w^E_f \succ b^E_g \) for every \( E \in P \), then \( f \) is preferred to \( g \). This generalizes the standard statewise monotonicity condition, for in the standard model every act is understood using the singletons partition, so setwise dominance in the axiom above becomes statewise dominance.

The above axiom imposes a weak form of monotonicity on the DM’s preference given the way the DM is understanding acts. The following axiom describes when the DM should satisfy a stronger form of monotonicity. We discuss it following the statement. Recall that \( \mathcal{E}(\mathcal{P}_\Sigma) = \bigcup_{P \in \mathcal{P}_\Sigma} \sigma(P) \) is the set of events that the DM is able to use when understanding acts, though not necessarily at the same time.

\(^{15}\)We use the notation \( \alpha f(E) + (1 - \alpha)x \) to mean the set of outcomes \( \{ ay + (1 - \alpha)x \mid y \in f(E) \} \).
Axiom 3 (Contingent Monotonicity). For any $A, B \in \mathcal{E}(\mathcal{P}_\xi)$ such that $A \subseteq B$, any $f \in \mathcal{F}_{\{A,B\}_{\neq}}$, $s \in B$, and $x \in X$,

$$f(s) > x \implies f > x \{s\} f.$$ 

To interpret the axiom, fix any $A, B \in \mathcal{E}(\mathcal{P}_\xi)$ such that $A \subseteq B$. Recall that this means that for any act $f$, the DM is able to think of $f$ in terms of its outcomes on $B$ and $B^c$. Further, suppose that $f \in \mathcal{F}_{\{A,B\}_{\neq}}$, as in the axiom. This means that $f$ yields three (or fewer) outcomes—one on $A$, one on $B \setminus A$, and one on $B^c$. Since the DM is able to split $f$ into $f(B)$ and $f(B^c)$ and the latter is a single outcome, the DM’s limited understanding only affects her understanding of $f(B)$, where she sees the two outcomes $f(A)$ and $f(B \setminus A)$. The axiom requires that if a strictly worse outcome is added to either $f(A)$ or $f(B \setminus A)$, the DM strictly disfavors this change. We view this as a natural requirement of the DM we are modeling because $A \in \mathcal{E}(\mathcal{P}_\xi)$. In other words, the DM has already revealed her ability to focus on $A$ when understanding acts, so she is able to see that $f(A)$ is distinct from $f(B \setminus A)$. Thus, she “notices” when a strictly worse outcome is added to either set, resulting in a strictly worse act. Notice that while the axiom alludes to the individual state $s \in B$, the reasoning just given does not rely on the DM focusing on $s$ itself. Since $f$ is constant on $A$ and $B \setminus A$, the DM notices the set of outcomes $f(A)$ or $f(B \setminus A)$ is worse without knowing exactly which state is worse since previously the set was a singleton.

The final axiom concerns the DM’s attitude toward her limited understanding. We will focus on the case where the DM is averse to her limited understanding, since that seems to be the more plausible attitude.\(^{16}\) To see what behavior reveals the DM’s attitude toward her limited understanding, recall that if the DM is using the understanding strategy $P$ when evaluating the act $f$, then for each $E \in P$ the DM sees the set of outcomes $f(E)$ and nothing else. If $f(E)$ is a singleton, then the DM’s limited understanding of that part of $f$ is irrelevant, as she perfectly understands that she will receive $f(E)$ if any $s \in E$ is realized. As such, we would expect a DM who dislikes her limited understanding to prefer the situation when $f(E)$ is a singleton. Now, suppose that we observe the preference

$$f > x Ef.$$ 

This reveals that although $x Ef$ is simpler than $f$, the outcome $x$ is not attractive enough to compensate the DM for her limited understanding of $f(E)$. If the DM is averse to her limited understanding, we would then expect that making $x Ef$ more complex does not result in a preference.

\(^{16}\)Additionally, the literature on ambiguity, especially in applications, has focused on aversion, and similarly in the coarse contingencies literature. For example, Dekel, Lipman, and Rustichini (1998) discuss notions of “unforeseen contingency aversion”. 

reversal; i.e.

\[ f > xAf \]

for any \( A \subseteq E \). The following axiom is slightly weaker than what was just alluded to; we revisit this point following Theorem 1 below.

**Axiom 4 (Aversion to Limited Understanding).** For all \( P \in \mathcal{P}_\succeq \), \( f \in \mathcal{U}(P) \), \( x \in X \), and \( E \in P \), if \( f > xEf \) and \( f > xE'f \) for every nonsingleton \( E' \in Q \in \mathcal{U}^{-1}(f) \) such that \( E \cap E' \neq \emptyset \), then

\[ f > xAf \]

for any \( A \subseteq E \).

The axiom formalizes the intuition given above. It is weaker in that it requires not only that \( f > xEf \) for \( E \in P \in \mathcal{U}^{-1}(f) \), but also that \( f > xE'f \) for any \( E' \) that intersects with \( E \). In other words, it requires not only that \( x \) not be enough to compensate the DM for her limited understanding of \( f(E) \), it must also not be sufficient to compensate the DM for her limited understanding of \( f(E') \), where \( E' \) overlaps with \( E \). In this case, the axiom requires that the DM’s preference does not reverse when \( xAf \) is made more complicated by “mixing” \( x \) with \( f(E) \). When \( \mathcal{U}^{-1}(f) \) is a singleton, this additional requirement has no bite and the axiom is precisely the condition described prior to its statement.

4. Results

4.1. Representation

This section contains our main results. The first subsection presents results that characterize the RU representation using Axioms 1-4, as well as uniqueness results. All proofs are in Appendix A.

**Theorem 1.** The preference \( \succeq \) satisfies Axioms 1-4 if and only if it admits a RU representation.

The proof of this result follows three main steps. First, we observe that Understanding Monotonicity implies that for any \( P \in \mathcal{P}_\succeq \) and \( f, g \in \mathcal{F}_P \) such that \( f(E) \succeq g(E) \) for all \( E \in P \), \( f \succeq g \). Since the preference satisfies Independence on \( \mathcal{F}_P \) by assumption, we can invoke the Anscombe and Aumann (1963) SEU theorem to show that the restriction of \( \succeq \) to each \( \mathcal{F}_P \) has a SEU representation whenever \( P \in \mathcal{P}_\succeq \). Thus, for each \( P \in \mathcal{P}_\succeq \), there is a probability \( \pi_P \) defined on \((S, \sigma(P))\) and vNM utility function \( u \) such that \( V_P(f) = \sum_{E \in P} \pi_P(E)u(f(E)) \) represents the preference on each \( \mathcal{F}_P \). The utility function \( u \) is independent of \( P \) because the preference satisfies Independence on \( \mathcal{F}_{\{S\}} = X \), and \( P \succ \{S\} \) for every \( P \in \mathcal{P}_\succeq \), so the utility over outcomes must be independent of the partition chosen.
Second, we show that Aversion to Limited Understanding implies that for every $f \in \mathcal{F}$, there exists $P \in \mathcal{P}_\pi$ such that $f \sim w_f^P$, where $w_f^P \coloneqq \left\{w_f^E, E \in P\right\}$. Thus, it follows that for every $f \in \mathcal{F}$, $V(f) = \sum_{E \in P} \pi_P(E) \min_{s \in E} u(f(s))$ represents $\pi$, where $P$ is such that $f \in \mathcal{U}(P)$. Moreover, for any $Q \in \mathcal{P}_\pi$, Understanding Monotonicity implies that $f \geq w_f^Q$, so we can rewrite $V$ as $V(f) = \max_{P \in \mathcal{P}_\pi} \sum_{E \in P} \pi_P(E) \min_{s \in E} u(f(s))$.

Notice that only Axioms 1, 2, and 4 were used to this point. The final step is showing that the collection of probabilities $\{\pi_P\}_{P \in \mathcal{P}_\pi}$ can be patched together to form $\pi$. This is where Contingent Monotonicity comes in. This step is slightly more subtle because it may not be the case that there exists a well defined $\pi$. More specifically, the difficulty in this is that there may be disjoint events $E_1, \ldots, E_n$ such that $E_i \in P_i$ for some $P_i \in \mathcal{P}_\pi$, while $E = \bigcup_{i=1}^n E_i \in P$ for some $P \in \mathcal{P}_\pi$. In general, it need not be the case that $\pi_P(E) = \sum_{i=1}^n \pi_{P_i}(E_i)$, which leaves indeterminacy in defining $\pi$. To overcome this, we show that Contingent Monotonicity and Aversion to Limited Understanding imply that $\mathcal{P}_\pi$ is rich. To see how richness allows us to overcome the problem, suppose that $i = 2$. In this case, notice that since $E_1 \in P_1 \in \mathcal{P}_\pi$ and $E = E_1 \cup E_2 \in P \in \mathcal{P}_\pi$ by assumption, richness implies that there must exist $Q \in \mathcal{P}_\pi$ such that $E_1, E_2 \in \sigma(Q)$. Moreover, since $E_1 \in \sigma(P_1) \cap \sigma(Q)$ and $V_{P_1}$ and $V_Q$ represent restrictions of the same preference, we must have $\pi_{P_1}(E_1) = \pi_Q(E_1)$, and similarly $\pi_{P_1}(E_2) = \pi_Q(E_2)$. Thus, since $\pi_Q$ is additive, it follows that $\pi_{P_1}(E_1) + \pi_{P_1}(E_2) = \pi_Q(E)$, so there is no longer indeterminacy in defining $\pi$. For $i > 2$, we can use induction on the size of $i$ (it must be finite since $S$ is finite).

Two remarks before we proceed. First, absent other assumptions, it can be demonstrated by example that there may not exist a probability on $S$ that extends $\pi$. To the extent that this calls into question the separation between ambiguity and coarse understanding of acts, one way of ensuring such an extension exists is to assume that the partition $\{\{s\}, \{S \setminus s\}\} \in \mathcal{P}$ for every $s \in S$. This is true if and only if $x > y$ implies $x \{s\} y > y$ for every $x, y \in X$. Adding these additional assumptions ensure not only that an extension exists, but that it is unique. All the results that follow also hold in this case.

Second, a remark about necessity. Consider the stronger version of Aversion to Limited Understanding that we described prior to the statement of the axiom. We show in Appendix A.2 that $f \in \mathcal{U}(P)$ if and only if $P$ solves (1) and there is no $R \gg P$ that also does so. If $P$ is the unique (minimal) solution and $f > xEf$, then $w_f^E > x$, so replacing $f(A)$ with $x$ yields a strictly worse act if the new act is understood using $P$. In this case, the stronger version of the axiom is satisfied. However, if both $P$ and $Q$ solve (1), and $s \in E \in P$ is such that $f(s) > x \geq w_f^{Q(s)}$ for all $s \in A$, then replacing $f(A)$ with $x$ does not change the utility of $f$ given the understanding strategy $Q$, so $f \sim xAf$. The weaker version that we use rules out this possibility by ensuring that $w_f^{Q(s)} > x$. The stronger version of the axiom is generically necessary under a joint nonredundancy condition on $\pi$ and $\mathcal{P}$.
We now move on to discussing the uniqueness properties of the representation. We say that a RU representation \( \langle u, \pi, P \rangle \) is nontrivial if \( u \) is nonconstant. Given two functions \( u, u' : X \rightarrow \mathbb{R} \), we write \( u \approx u' \) if there exist \( \alpha > 0 \) and \( \beta \in \mathbb{R} \) such that \( u = \alpha u' + \beta \). The following defines a similar equivalence on the space of sets of partitions.

**Definition 5.** If \( P \) and \( Q \) are two collections of partitions, we write \( P \approx Q \) if

(i) \( P = Q \),

(ii) or for every \( P \in P \setminus Q \) and \( Q \in Q \setminus P \), there exists \( P', Q' \in P \cap Q \) such that \( P' \gg P \) and \( Q' \gg Q \).

In other words, \( P \approx Q \) if \( P = Q \) or if \( P \) and \( Q \) differ only by adding coarsenings of partitions common to both, i.e. their “minimal” sets are equal. Recall that \( E(P) = \bigcup_{P \in P} \sigma(P) \), so \( P \approx Q \) implies that \( E(P) = E(Q) \). The following result shows that the parameters of the RU representation are essentially unique.

**Theorem 2.** Suppose \( \langle u, \pi, P \rangle \) and \( \langle u', \pi', P' \rangle \) are two nontrivial RU representations of the same preference. Then \( P \approx P' \), \( u \approx u' \), and \( \pi = \pi' \).

Let \( V \) and \( V' \) denote the functionals corresponding to the two RU representations in the statement of the theorem. Uniqueness of \( u \) is standard. The key idea for the rest of the uniqueness result is that \( P \) is unique regardless of what \( \pi \) is. More specifically, notice that \( \succeq \) satisfies Independence on \( F_P \) for every \( P \in P \), regardless of \( \pi \). If there exists \( Q \in P \setminus P' \) and \( Q \) is not a coarsening of a partition common to both, then we can show that \( V' \) must exhibit an Independence violation on \( F_Q \), regardless of what \( \pi' \) is. This contradicts the fact that \( \succeq \) satisfies Independence on \( F_P \) for all \( P \in P \). Once we have established that \( P \approx P' \) independently of \( \pi \) and \( \pi' \), uniqueness of \( \pi \) on \( E(P) \) follows from the fact that each \( \pi_P \) is unique, and the extension to \( E(P) \) leaves no further indeterminacy.

For the remainder of this section, we assume that the utility function \( u \) in any RU representation is nonconstant so we can work with essentially unique RU representations.

**4.2. Comparing Understanding**

Our framework allows us to behaviorally compare different DMs who have a limited understanding of acts. Thus, let \( \succeq_1 \) and \( \succeq_2 \) denote two preference relations over \( F \) corresponding to two different DMs. The following is a choice based way of identifying which DM has a finer understanding of acts.
**Definition 6.** $\succeq_1$ understands more than $\succeq_2$ if for every $P \in \mathcal{P}_{\succeq_2}$, $f \in \mathcal{F}$, and $g \in \mathcal{F}_P$,

$$f \succeq_2 g \implies f \succeq_1 g.$$ 

The definition parallels comparative definitions of risk and ambiguity aversion.\footnote{See Epstein (1999), Ghirardato and Marinacci (2002), and the references therein.} DM2’s understanding strategies can be identified from $\succeq_2$ following Definition 3. Recall that if $P \in \mathcal{P}_{\succeq_2}$, then our hypothesis is that $P$-measurable acts are perfectly understood in the same way that constant acts are “ambiguity free” when defining comparative ambiguity aversion (e.g. in Ghirardato and Marinacci (2002)). As such, the definition says that DM1 understands more than DM2 if whenever DM2 prefers an act $f$ to a perfectly understood act $g$, so does DM1. The following result characterizes DM1 understanding more than DM2 in terms of the parameters of the RU representation.

**Proposition 1.** Suppose $\succeq_1$ and $\succeq_2$ admit RU representations $\langle u_1, \pi_1, P_1 \rangle$ and $\langle u_2, \pi_2, P_2 \rangle$. Then the following are equivalent:

(i) $\succeq_1$ understands more than $\succeq_2$.

(ii) $P_2 \subseteq P_1$, $\pi_1(E) = \pi_2(E)$ for all $E \in \mathcal{E}(P_2)$, and $u_1 \approx u_2$.

The result shows that if DM1 understands more than DM2, then intuitively her set of understanding strategies $\mathcal{P}_1$ is a superset of $\mathcal{P}_2$. The definition also implies that the two DMs have the same risk preferences and beliefs on $\mathcal{E}(P_2)$. Importantly, the part of the above result concerning the DMs’ understanding does not rely on the DMs’ attitude toward their coarseness. More specifically, for any pair of preferences $\succeq_1, \succeq_2$ that satisfy Axiom 1, if we elicit $\mathcal{P}_{\succeq_1}$ and $\mathcal{P}_{\succeq_2}$ as in Definition 3, then $\succeq_1$ understands more than $\succeq_2$ if and only if $u_1 \approx u_2$ (where each $u_i$ represents the restriction of $\succeq_i$ to $X$ and exists by Axiom 1) and $\mathcal{P}_{\succeq_2} \subseteq \mathcal{P}_{\succeq_1}$.

### 4.3. Ambiguity

In the Introduction, we argued that the distinguishing feature of a DM who may not understand acts is $\mathcal{P}$ being nonsingleton. Otherwise, the DM is indistinguishable from an ambiguity averse DM. In this section, we argue this point more formally by showing that given Axioms 1-4, ambiguity aversion is equivalent to $\mathcal{P}$ being a singleton (without loss of generality). Moreover, we show that in this case the model lies in the intersection of two of the most commonly used models of ambiguity aversion, maxmin expected utility (Gilboa and Schmeidler, 1989) and Choquet expected utility (Schmeidler, 1989).
We begin with some additional notation and definitions. Fix an RU representation \( \langle u, \pi, P \rangle \). Unless \( P = \{S\} \), then \( P \) is not a singleton since if \( P \in P \), then every partition \( Q \) such that \( P \gg Q \) is also in \( P \). However, given the uniqueness result in Theorem 2, \( P \) may be taken to be a singleton if there exists a partition \( P \in P \) such that \( P \gg Q \) for all \( Q \in P \). If such a partition \( P \) exists, it is easy to see that \( P \) must be the coarsest common refinement of every \( Q \in P \). Thus, given a set of partitions \( P \), define \( \Omega(P) \) to be the coarsest common refinement of all partitions in \( P \). Notice that if \( \Omega(P) \in P \), then \( \Omega(P) \approx P \), so \( \langle u, \pi, P \rangle \) and \( \langle u, \pi, \Omega(P) \rangle \) represent the same preference. Therefore every act is understood using the same partition \( \Omega(P) \).

Next, recall the following behavioral definition of ambiguity aversion from Schmeidler (1989) and Gilboa and Schmeidler (1989).

**Definition 7.** A preference \( \succcurlyeq \) is uncertainty averse if for every \( f, g \in F \) and \( \alpha \in (0, 1) \),

\[
 f \sim g \implies \alpha f + (1 - \alpha)g \succcurlyeq g.
\]

The following result shows that within the RU model, uncertainty aversion is equivalent to the DM understanding every act using the same strategy.

**Proposition 2.** Suppose \( \succcurlyeq \) has an RU representation \( \langle u, \pi, P \rangle \). Then \( \Omega(P) \in P \) if and only if \( \succcurlyeq \) is uncertainty averse.

The result above shows that when the DM uses the same understanding strategy when evaluating every act, it is as if the DM views her limited understanding as generating ambiguity. Indeed, the following is a trivial corollary of the above result. Say that a preference is MMEU (maxmin expected utility, Gilboa and Schmeidler (1989)) if there exists \( \langle u, M \rangle \), where \( u : X \to \mathbb{R} \) is a utility function and \( M \subseteq \Delta(S) \) is a closed, convex set of probabilities on \( S \) such that

\[
 V_{\text{MMEU}}(f) = \min_{\mu \in M} \int u(f(s)) d\mu(s)
\]

represents \( \succcurlyeq \).\(^\text{18}\) Let \( \overline{\text{co}} (\cdot) \) denote the closed convex hull operator.

**Corollary 1.** Suppose \( \succcurlyeq \) has an RU representation \( \langle u, \pi, P \rangle \) and \( \Omega(P) \in P \). Then \( \succcurlyeq \) admits a MMEU representation \( \langle u, M \rangle \) where

\[
 M = \overline{\text{co}} \left( \{ \mu \in \Delta(S) \mid \forall E \in \Omega(P), \mu(E) = \pi(E) = \mu(s) \text{ for some } s \in E \} \right).
\]

The additional structure from the RU representation allows us to characterize the set \( M \) beyond the closedness and convexity that comes from the Gilboa and Schmeidler (1989) representation.

\(^{18}\)Since \( S \) is finite, endow \( \Delta(S) \) with Euclidean distance.
In this special case, the set $M$ is the set of probability measures $\mu$ that agree with $\pi$ on each $E \in \Omega(\mathcal{P})$ and whose conditionals $\mu_E$ are degenerate on some $s \in E$.

Interestingly, if $\Omega(\mathcal{P}) \in \mathcal{P}$, the RU model falls into another class of ambiguity models, the Choquet expected utility model (CEU, Schmeidler (1989)). A capacity is a normalized, monotone function $\nu : \Sigma \rightarrow [0, 1]$. A preference $\succeq$ is CEU if there exists a pair $\langle u, \nu \rangle$, where $u$ is a utility function as above and $\nu$ is a capacity, such that

$$V_{CEU}(f) = \int u(f) \, d\nu$$

represents $\succeq$ (the integration here is the Choquet integral; see Schmeidler (1989) for details).

**Proposition 3.** Suppose $\succeq$ has an RU representation $\langle u, \pi, \mathcal{P} \rangle$ and $\Omega(\mathcal{P}) \in \mathcal{P}$. Then there exists a unique capacity $\nu$ such that the CEU representation $\langle u, \nu \rangle$ represents $\succeq$, where

$$\nu(E) = \max_{A \in \sigma(\Omega(\mathcal{P})): A \subseteq E} \pi(A)$$

for every $E \in \Sigma$.

Again, the additional structure brought by the RU model allows us to characterize the capacity $\nu$ in terms of the key parameters of the RU model $\pi$ and $\mathcal{P}$. The capacity $\nu$ agrees with $\pi$ on its domain, which in this case is $\sigma(\Omega(\mathcal{P}))$. For $E \in \Sigma \setminus \sigma(\Omega(\mathcal{P}))$, $\nu(E)$ is an “inner approximation” using the DM’s belief $\pi$.

A natural question that follows given the proposed connection between ambiguity and limited understanding of acts is exactly what events the DM perceives as ambiguous. If the DM indeed reduces her limited understanding to ambiguity when there is a single understanding strategy, then the events in this partition $\Omega(\mathcal{P})$ should be unambiguous to the DM. To address this question, recall the following definition of unambiguous events from Epstein and Zhang (2001).

**Definition 8.** An event $E \in \Sigma$ is unambiguous if for every $f \in \mathcal{F}$, $x, y, z, z' \in X$, and disjoint events $A, B \subseteq E^c$,

$$\begin{bmatrix} z & E \\ x & A \\ y & B \\ f & E^c \setminus [A \cup B] \end{bmatrix} \succ \begin{bmatrix} z' & E \\ x & A \\ y & B \\ f & E^c \setminus [A \cup B] \end{bmatrix} \Rightarrow \begin{bmatrix} z' & E \\ x & A \\ y & B \\ f & E^c \setminus [A \cup B] \end{bmatrix},$$

and the above holds when $E$ is replaced with $E^c$.

The following proposition shows that unambiguous events (as defined above) are exactly those that are in the algebra generated by $\Omega(\mathcal{P})$. 
**Proposition 4.** Suppose ⪰ has an RU representation \( \langle u, \pi, \mathcal{P} \rangle \) and \( \Omega(\mathcal{P}) \in \mathcal{P} \). Then \( E \in \sigma(\Omega(\mathcal{P})) \) if and only if \( E \) is unambiguous.

At an intuitive level, when the DM’s preference admits an RU representation \( \langle u, \pi, \mathcal{P} \rangle \) and \( \Omega(\mathcal{P}) \in \mathcal{P} \), the DM’s preference satisfies a type of separability with respect to the events in \( \Omega(\mathcal{P}) \), akin to the Sure Thing Principle (Savage, 1954, P2). This type of separability is exactly what Epstein and Zhang (2001) were focused on in their definition of unambiguous events.

In the special case of the model where \( \Omega(\mathcal{P}) \in \mathcal{P} \), we can revisit our comparative definition of coarse understanding. A corollary of Proposition 1 is the following.

**Corollary 2.** Suppose \( \succeq_1 \) and \( \succeq_2 \) admit RU representations \( \langle u_1, \pi_1, \mathcal{P}_1 \rangle \) and \( \langle u_2, \pi_2, \mathcal{P}_2 \rangle \) with \( \Omega(\mathcal{P}_i) \in \mathcal{P}_i \) for \( i = 1, 2 \). Then the following are equivalent:

1. \( \succeq_1 \) understands more than \( \succeq_2 \).
2. \( \Omega(\mathcal{P}_1) \gg \Omega(\mathcal{P}_2) \), \( \pi_1(E) = \pi_2(E) \) for all \( E \in \Omega(\mathcal{P}_2) \), and \( u_1 \approx u_2 \).

The result says that when two DMs’ preferences have RU representations and each have a single understanding strategy, if DM1 understands more than DM2 as in Definition 6, then DM1’s strategy is finer than DM2’s. In other words, DM1 can “read” more closely than DM2. In this case, another implication is that DM1 is less ambiguity averse in the sense the set of measures \( \mathcal{M}_1 \) in the MMEU representation of DM1’s preference is a subset of \( \mathcal{M}_2 \). As such, our behavioral definition of “understands more” becomes a definition for “more ambiguity averse” when the DMs each only have one understanding strategy.

### 4.4. Dynamic Consistency

A natural question that follows our analysis is how a DM who does not understand acts plans for the future and reacts to learning new information. While characterizing a dynamic model is beyond our scope, in this section we present a natural extension of the RU model to accommodate ex ante and ex post choice. Throughout this section, we consider a pair of preferences \( (\succeq, \succeq_A) \) over \( \mathcal{F} \), where \( A \subset S \). The interpretation is that \( \succeq_A \) represents the DM’s choices after being told that the event \( A \) occurred. We assume that the DM is directly told that the event \( A \) occurred, so that hearing this information would enable her to focus on \( A \) if she was not able to before.\(^{19}\)

In other words, even if \( A \) is too fine for the DM to focus on according to her ex ante preference \( \succeq \), we assume that each partition ex post is a partition of \( A \).

Before presenting the model, we need some additional notation. First, if \( Q \) is a set of partitions of \( S \) and \( Q \) is any other partition of \( S \), let \( Q \vee Q := \{ Q' \vee Q \mid Q' \in Q \} \) denote the set of partitions

\(^{19}\)This echoes the intuition in Ahn and Ergin (2010), where explicit descriptions of the event \( A \) ensure the DM is aware of \( A \).
obtained by taking the coarsest common refinement of $Q$ with each $Q' \in Q$. Second, if $E \in \Sigma$ and $Q$ is a set of partitions such that $E \in \sigma(Q)$ for all $Q \in Q$, let $Q|_E$ denote the set of partitions obtained by truncating each $Q \in Q$ at $E$; i.e. $Q|_E$ is the set of partitions of $E$ such that for every $Q \in Q|_E$, there exists $Q' \in Q$ such that $Q \subseteq Q'$ and if $B \in Q' \setminus Q$, then $B \subseteq E^c$. We can now state the model.

**Definition 9.** A pair of preferences $(\succ, \succ_A)$ has a **dynamic revealed understanding** representation if $\succ$ admits a revealed understanding representation $\langle u, \pi, \mathcal{P} \rangle$ and $\succ_A$ admits a revealed understanding representation $\langle u, \pi_A, \mathcal{P}_A \rangle$, where $\mathcal{P}_A = [\mathcal{P} \cup \{A, A^c\}]|_A$ and $\pi_A : \mathcal{E}(\mathcal{P}_A) \rightarrow [0, 1]$ is a normalized, additive, strictly monotone function satisfying

$$
\pi_A(E) = \frac{\pi(A \cap E)}{\max_{B \in \mathcal{E}(\mathcal{P}_A) : B \subseteq A} \pi(B)}
$$

for any $E \in \mathcal{E}(\mathcal{P}_A)$ such that $A \cap E \in \mathcal{E}(\mathcal{P})$.

When $A$ is fixed, we can compactly summarize the dynamic revealed understanding (DRU) model using the tuple $\langle u, \pi, \pi_A, \mathcal{P} \rangle$. One can see the key modeling assumptions brought forth in the DRU definition. The set of partitions is “updated” in the sense that each partition ex post is the coarsest common refinement of a partition ex ante with the partition $\{A, A^c\}$. Moreover, the set of partitions $\mathcal{P}_A$ are all partitions of $A$, so the DM “believes” and understands her information that $A$ occurred. Lastly, the posterior belief $\pi_A$ is updated from the prior using a rule that reduces to Bayes rule whenever it is well defined. If $A \notin \mathcal{E}(\mathcal{P})$, then there are other natural choices for the denominator in the definition of $\pi_A(E)$; however it will become clear that this will not matter for our purposes in this subsection.

Recall the following definition that connects ex ante and ex post preferences in the SEU model (see Ghirardato (2002)).

**Definition 10.** A pair of preference relations $(\succ, \succ_A)$ satisfies **Dynamic Consistency** if for every $f, g \in \mathcal{F}$,

$$
f Ah \succ g Ah \iff f \succ_A g
$$

for some $h \in \mathcal{F}$.

Loosely, dynamic consistency reveals that the DM’s ex post choices are guided by what she planned to choose ex ante if confronted with the information that $A$ occurred. Intuitively, a DM who does not understand acts may violate Dynamic Consistency. The main idea is that if the DM does not focus on the event $A$ ex ante when evaluating $fAh$ or $gAh$, then it is intuitive that
learning $A$ could lead to changes in her preferences, violating dynamic consistency. The following result confirms this intuition—imposing dynamic consistency in the DRU model requires the DM to always focus on the event $A$ when making choices ex ante.

**Proposition 5.** Suppose the pair of preferences $(\succ, \succ_A)$ admit a DRU representation given by $(u, \pi, P)$ and $(u, \pi_A, P_A)$. If the pair of preferences also satisfy Dynamic Consistency, then $A \in \sigma(P)$ for all $P \in \mathcal{P}$.

The result says that if the DM’s preferences have a DRU representation and also satisfy dynamic consistency, then the DM must be able to focus of the event $A$ regardless of what other parts of the acts she is focusing on. This result has far reaching implications for how limited understanding interacts with dynamic consistency. To see this, suppose that we instead took as primitive a collection of preferences $\{\succ_A\}_{A \in \Sigma}$ and assumed each pair $(\succ, \succ_A)$ admitted a DRU representation. Then if each pair $(\succ, \succ_A)$ satisfies dynamic consistency, Proposition 5 implies that for every $A \in \Sigma$, $A \in \sigma(P)$ for every $P \in \mathcal{P}$. In other words, we must have $\sigma(P) = \Sigma$ for every $P \in \mathcal{P}$, so the DM must be SEU (and hence will also be SEU ex post, where $\pi_A$ is the Bayesian update of $\pi$). Thus, if the DM is dynamically consistent following any arrival of information, this rules out limited understanding.\(^{20}\) However, some limited understanding can coexist with dynamic consistency if the DM is only dynamically consistent with respect to the events she always focuses on.

One may wonder if the converse of this statement holds. It is easy to see that if $\mathcal{P}$ is a singleton, then the converse of the result holds. However, the converse may fail in general. The reason is that we may have acts $f, g, h \in \mathcal{F}$ such that $fAh \succeq gAh$, but the understanding strategies being used when evaluating $fAh$ and $gAh$ differ on $A^c$, so the ex ante preference may still fail to be separable on $A$. The converse would hold if every $P, Q \in \mathcal{P}$ agreed on $A$ or $A^c$, or some combination of both.

### 4.5. Foresight

Intuitively, foresight relates to which events the DM plans for when making choices. At an informal level, to plan for an event when choosing between acts requires focusing on acts’ payoffs on that event. In this subsection we make this connection formal by considering a different dynamic extension of the model, which facilitates comparison with the behavioral definition of foreseen events in Kochov (2018). In the main text, he considers a preference over two period streams of state dependent outcomes, or equivalently pairs of acts. The interpretation is that the DM makes a choice at time 0, the state is realized at time 1, and the DM receives a state dependent outcome in time 1 and time 2. Pairs of acts are elements of $\mathcal{F}^2$. To differentiate generic elements of $\mathcal{F}^2$

\(^{20}\) A similar message appears in Minardi and Savochkin (2017) as well, where they argue that if the DM satisfies dynamic consistency following any arrival of information, then she must be SEU.
from elements of $\mathcal{F}$, we use $f, g, h$, etc. to denote elements of $\mathcal{F}^2$, i.e. $f = (f_1, f_2)$, where $f_1, f_2 \in \mathcal{F}$.

Consider the following dynamic extension of the RU model.

**Definition 11.** A preference $\succsim \subset \mathcal{F}^2 \times \mathcal{F}^2$ has an intertemporal revealed understanding representation if there exists

- a continuous, mixture linear function $u : X \rightarrow \mathbb{R}$,
- a rich set of partitions $\mathcal{P}$,
- a discount factor $\delta \in (0, 1)$,
- and a normalized, additive, strictly monotone function $\pi : \mathcal{E}(\mathcal{P}) \rightarrow [0, 1]$,

such that

$$W(f) = \max_{P \in \mathcal{P}} \sum_{E \in P} \pi(E) \min_{s \in E} u(f_1(s)) + \delta \max_{P \in \mathcal{P}} \sum_{E \in P} \pi(E) \min_{s \in E} u(f_2(s))$$

represents $\succsim$.

We can summarize the IRU model using the tuple of parameters $\langle u, \pi, \delta, \mathcal{P} \rangle$. The model extends the RU model to intertemporal choice by hypothesizing that the DM discounts the future and that the DM chooses an understanding strategy for each act in the stream. Notice that the IRU model is not a special case of the representation in Kochov (2018), since it may violate his Coarse Recursivity axiom.\(^{21}\)

To introduce his definition of foreseen events, we need to introduce some preliminary definitions. For any $f \in \mathcal{F}^2$, let $f(s) := (f_1(s), f_2(s)) \in X^2$. If $f, g \in \mathcal{F}^2$ and $E \in \Sigma$, let $f_E = (f_1E, f_2E)$. An act $f$ is effectively certain if $f(s) \sim f(s')$ for all $s, s' \in S$. An effectively certain act $f$ is subjectively certain if $f \sim f(s)$ for some $s \in S$. An event $A \in \Sigma$ is foreseen if for every $d, d' \in X^2$ such that $dA \subseteq dA'$ is effectively certain, $dA'$ is subjectively certain. The following result shows that within the IRU model, foreseen events are exactly those in some $P \in \mathcal{P}$; i.e. an event that the DM is able to focus on when evaluating acts.

**Proposition 6.** Suppose the preference $\succsim \subset \mathcal{F}^2 \times \mathcal{F}^2$ admits an IRU representation $\langle u, \pi, \delta, \mathcal{P} \rangle$. Then an event $A \in \Sigma$ is foreseen if and only if there exists $P \in \mathcal{P}$ such that $A \in P$.

This result also holds if we consider the alternative model where the partition varies with the act but not with time, i.e. there is only a single “max” in the representation. Notice that even in the case where $\mathcal{P}$ may be taken to be a singleton, the IRU model (and even its counterpart with a single “max” term) allow unforeseen events given the proposed definition.

\(^{21}\)Coarse Recursivity requires that for any collection of minimal foreseen events (in terms of set inclusion) $A_1, \ldots, A_n$ and acts $f, g \in \mathcal{F}^2$, if $fA_i \succsim g$ for all $i = 1, \ldots, n$, then $f \left[ \bigcup_{i=1}^n A_i \right] \succsim g$. 


Proposition 6 suggests that the type of foresight captured by Kochov’s definition is similar to the type of focus we consider. In particular, in his model, if $A$ is foreseen, then the DM’s preference between (streams of) acts that are bets on $A$ conforms to discounted subjective expected utility. Similarly, in our model, for any $A \in P \in \mathcal{P}$, the DM’s preference between acts that are bets on $A$ conforms to subjective expected utility. However, the set of events $\mathcal{E}(P)$ and Kochov’s set of foreseen events differ in a significant way. First, notice that Kochov’s definition identifies each foreseen event individually, and the axioms he imposes on the preference imply that the DM can foresee any collection of foreseen events simultaneously. For example, if $A$ and $B$ are disjoint foreseen events, Kochov’s axioms imply that the DM’s preference between $\{A, B, [A \cup B]^c\}$-measurable acts conforms to discounted SEU. On the contrary, since we identify entire partitions (understanding strategies) and not individual events, our model does not require this type of “combination” of coarse events. Returning to the previous example where $A$ and $B$ are disjoint foreseen events (and hence each in some $P, Q \in \mathcal{P}$ by Proposition 6), our model permits the DM violate to Independence when choosing between $\{A, B, [A \cup B]^c\}$-measurable acts, unlike in Kochov (2018). This makes sense in our setting because this is precisely the type of bounded rationality our model intends to capture; the DM’s cognitive limitations may preclude her from focusing on $A$ and $B$ simultaneously even though she can focus on them separately.


We conclude with a comparison with Ahn and Ergin (2010), which is facilitated by the results from Section 4.3. They study a DM whose preferences over acts depend on the way the act is described, or framed; these descriptions take the form of partitions of $S$. For example, if $P$ and $Q$ are partitions of $S$ and $P \gg Q$, they allow the DM’s preferences over $Q$-measurable acts to depend on whether the state space was described using $P$ or $Q$. Their main motivation is explaining how a DM’s beliefs may depend on how acts are described, and they provide a behavioral notion of events that are “transparent”, or independent of how acts are described. The authors argue that one could interpret transparent events as those that are perfectly “foreseen” by the DM.

As discussed in the previous subsection, foresight of an event is related to which events the DM has in mind when evaluating an act. In this subsection, we provide a formal comparison of the “foreseen” events in Ahn and Ergin (2010) with the events that the DM focuses on. Transparent events are not directly comparable with the events in the partitions that we identify due to the different primitives considered. However, in the special case of our model where $\Omega(P) \in \mathcal{P}$, one could compare the algebra generated by this partition with the collection of transparent events, which is also an algebra. This is natural since their foreseen events are independent of the description and the act the DM is evaluating.
Ahn and Ergin (2010) present a model with two parameters: a utility function $u$ and a support function $\lambda : \Sigma \rightarrow \mathbb{R}_+$ such that $\sum_{E \in P} \lambda(E) > 0$ for every partition $P$ in some exogenously specified set. Ahn and Ergin present a behavioral definition of “transparent” events; let $A$ denote the set of transparent events. In their Proposition 2, they show that $A$ has the following properties:

- $A$ is an algebra,
- $\lambda$ is additive on $A$,
- $A = \Sigma$ if and only if $\lambda$ is additive on $\Sigma$,
- and $A \in A$ if and only if
  $$\lambda(E) = \lambda(E \cap A) + \lambda(E \cap A^c)$$
  for all $E \in \Sigma$.

In the special case of the RU representation where $\Omega(P) \in \mathcal{P}$, the natural object to compare $A$ with is $\sigma(\Omega(P))$, the set of events that are generated by the DM’s fixed understanding strategy (recall these events “are unambiguous” in the ambiguity interpretation). By definition, $\sigma(\Omega(P))$ is an algebra so it is similar to $A$ in that respect. Similarly, $\pi$ is additive on $\sigma(\Omega(P))$ by definition (and hence so is $\nu$ given Proposition 3). The following result shows that $\sigma(\Omega(P))$ also shares the other two properties that $A$ has.

**Proposition 7.** Suppose $\succeq$ has an RU representation $\langle u, \pi, \mathcal{P} \rangle$ and $\Omega(P) \in \mathcal{P}$. Let $\langle u, \nu \rangle$ be the CEU representation obtained in Proposition 3. Then the following hold:

(i) $A \in \sigma(\Omega(P))$ if and only if $\nu(E) = \nu(E \cap A) + \nu(E \cap A^c)$ for all $E \in \Sigma$.

(ii) $\sigma(\Omega(P)) = \Sigma$ if and only if $\nu$ is additive.

In words, this means that the capacity $\nu$ obtained in Proposition 3 has many of the same properties regarding $\sigma(\Omega(P))$ as the support function in Ahn and Ergin (2010) and moreover, their transparent events share properties with the events that the DM studied here focuses on. This strengthens the connection between foresight and focus, suggesting that the events for which their DM is able to form beliefs that are invariant to the description of the state space are closely related to the events that our DM focuses on when making a choice. Both ideas capture a sophistication on the part of the DM that the models permit to fail.
A. Proofs

A.1. Proof of Theorem 1—Sufficiency

Throughout, we assume that there exist \( f, g \in \mathcal{F} \) with \( f > g \), otherwise the result follows trivially. We also assume that for every \( P \in \mathcal{P}_\succ \), there exists \( E \in P \) with \( |E| \geq 2 \). Otherwise, \( \{s \}_{s \in S} \in \mathcal{P}_\succ \), so \( \succ \) satisfies Independence everywhere and is SEU, so the result follows trivially.\(^{22}\)

Now, the vNM theorem implies that there exists a mixture linear function \( u : X \rightarrow \mathbb{R} \) such that \( u \) represents the restriction of \( \succ \) to \( X \). It is routine to show that Continuity implies that \( u \) is continuous (in the given metric on \( X \), which is metrizable by assumption). The first lemma shows that \( \mathcal{F} = \bigcup_{P \in \mathcal{P}_{\succ}} \mathcal{U}(P) \).

**Lemma 1.** For every \( f \in \mathcal{F} \), there exists \( P \in \mathcal{P}_{\succ} \) such that \( f \in \mathcal{U}(P) \).

**Proof.** Notice that by Certainty Independence, for every \( x, y, z \in X \) and \( \alpha \in [0, 1] \), we have that

\[
x \succ y \iff \alpha x + (1 - \alpha) z \succ \alpha y + (1 - \alpha) z.
\]

Thus, the trivial partition \( \{S\} \in \mathcal{P}_{\succ} \). Applying Certainty Independence again, for every \( f \in \mathcal{F} \), \( x, y \in X \), and \( \alpha \in (0, 1) \), we have

\[
f \succ x \implies \alpha f + (1 - \alpha) y > \alpha x + (1 - \alpha) y.
\]

Since \( X = \mathcal{F}_{\{S\}} \), it follows that either \( f \in \mathcal{U}(\{S\}) \), or \( f \in \mathcal{U}(P) \) for some \( P \gg \{S\} \).\( \square \)

Similarly to the text, for any partition \( P \), let \( w^P_f = (w^E_f, E \in P) \). We next show that each \( f \in \mathcal{F} \) is indifferent to \( w^P_f \) for some \( P \in \mathcal{P}_{\succ} \).

**Lemma 2.** For every \( f \in \mathcal{F} \), there exists \( P \in \mathcal{P}_{\succ} \) such that \( f \sim w^P_f \).

**Proof.** We prove the result by strong induction on the size of \( f(S) \). We begin with the \( |f(S)| = 2 \) case (the singleton case is trivial). Thus, take any \( f \in \mathcal{F} \) with \( |f(S)| = 2 \). We begin with the following claim.

**Claim 1.** For every \( f \in \mathcal{F} \), there exists \( P \in \mathcal{U}^{-1}(f) \) and \( E \in P \) such that \( f \preceq w^E_f \).

**Proof of Claim 1.** Suppose toward a contradiction that for every \( P \in \mathcal{U}^{-1}(f) \) and \( E \in P \), \( f > w^E_f \). This implies that each \( E \) contains both outcomes of \( f \), for otherwise \( w^E_f = f \), contradicting the supposed strict preference. Moreover, notice that for any such \( E \) and \( P \), if \( s \in E \) is such that

\(^{22}\)It is easy to see that Understanding Monotonicity implies the Anscombe and Aumann (1963) monotonicity axiom if \( \{s \}_{s \in S} \in \mathcal{P}_{\succ} \).
We now show that Claim 1 applies to

We now move to the inductive step. So suppose that for every

Aversion to Limited Understanding implies that either \(w_f^E f \succeq f\) or that there exists \(Q \in \mathcal{U}^{-1}(f)\) and \(s \in E\) such that \(w_f^E Q(s) f \succeq f\). Since the former possibility is ruled out by assumption, we proceed under the latter assumption. However, since \(f\) only has two outcomes and \(w_f^E\) is the dispreferred of the two, \(w_f^E = w_f^{Q(s)}\), so the desired result follows.

From here, the proof proceeds iteratively. Take \(E \in P \in \mathcal{U}^{-1}(f)\) as established in the claim above, and consider \(w_f^E f\). If \(|w_f^E f(S)| = 1\), then \(w_f^E f = w_f^S\) (recall \(|f(S)| = 2\), so \(f(S \setminus E) = w_f^E\). Thus, apply Understanding Monotonicity to conclude that \(P\) is the desired partition. So suppose \(|w_f^E f(S)| = 2\). Then we can apply Claim 1 to \(w_f^E f\), keeping in mind it need not be understood using \(P\). Continue on in this fashion until one of two things occurs:

1. \(w_f^S \succeq f\), in which case Understanding Monotonicity implies that \(f \succeq w_f^P \succeq w_f^S \succeq f\), so any \(P \in \mathcal{P}_\succeq\) is satisfactory.

2. There exists \(P \in \mathcal{P}_\succeq\) such that \(w_f^P \succeq f\). By Definition 4, it follows that there exists \(Q \in \mathcal{P}_\succeq\) such that \(Q \gg P\) and \(w_f^P \in \mathcal{U}(Q)\). Since \(Q \gg P\), for every \(E \in P\), there exists \(A^1_E, \ldots, A^n_E \in Q\) such that \(E = \bigcup_{i=1}^n A^i_E\). By definition, for every \(A \in Q\), \(b^A_w = w_f^E\) for any \(E \in P\) such that \(A \subseteq E\). Since \(A \subseteq E\), \(w_f^A \succeq w_f^E\). Thus, Understanding Monotonicity implies that \(f \sim w_f^P\), as desired.

We now move to the inductive step. So suppose that for every \(f \in \mathcal{F}\) with \(|f(S)| \leq n\), there exists \(P \in \mathcal{P}_\succeq\) such that \(f \sim w_f^P\). Consider any \(f \in \mathcal{F}\) with \(|f(S)| = n + 1\). Recalling that \(\mathcal{U}^{-1}(f)\) is nonempty by Lemma 1, take any \(P \in \mathcal{U}^{-1}(f)\) and any nonsingleton \(E \in P\) such that \(w_f^S \in f(E)\).

We now show that Claim 1 applies to \(E\). Again, suppose toward a contradiction that \(f > w_f^E f\). Let \(s \in E\) be any state such that \(f(s) = w_f^E\). Then since \(\{s\} \subseteq E\) and \(w_f^E\) is the dispreferred of \(f\), the contrapositive of Aversion to Limited Understanding implies that either \(w_f^E f \succeq f\) or that there exists \(Q \in \mathcal{U}^{-1}(f)\) and \(s \in E\) such that \(w_f^E Q(s) f \succeq f\), where \(Q(s)\) is nonsingleton. Recall that by assumption, \(w_f^E\) is the worst outcome in \(f(S)\) such that there exists \(R \in \mathcal{U}^{-1}(f)\) with \(R \left(f^{-1}(w_f^E)\right)\) nonsingleton. Thus it follows that \(w_f^E \sim w_f^{Q(s)}\), so \(Q\) is the desired partition. As above, repeat this procedure for \(w_f^{Q(s)}\) and so on until one of the \(Q(s)\) events contains more than one outcome. If this does not occur, then the conclusion of Lemma 2 follows. At this point, we have constructed an act \(g \succeq f\) such that \(g(s) \succeq f(s)\) for every \(s \in S\) and \(|g(S)| \leq n\). As such, the inductive assumption implies that there exists \(R \in \mathcal{P}_\succeq\) such that \(w_g^R \sim g \succeq f\). Since \(f(s) \succeq g(s)\) for every \(s \in S\), \(w_f^R \succeq w_g^R \succeq f\). An argument identical to case 2 above then shows that \(w_f^R \sim f\), as desired.

\footnote{If no such \(P\) exists, move on to the next worst outcome in \(f(S)\) and so on. One such \(P\) must exist, otherwise the preference is SEU and the proof of sufficiency is complete.}
If \( f \in \mathcal{F}_P \) and \( E \in P \), let \( u(f(E)) := u(f(s)) \) for any \( s \in E \). This is well defined because \( u(f(s)) = u(f(s')) \) for all \( s, s' \in E \in P \) whenever \( f \in \mathcal{F}_P \). Moreover, notice that if \( P \in \mathcal{P}_\succneq \), Understanding Monotonicity implies that for any \( f, g \in \mathcal{F}_P \), if \( f(E) \succneq g(E) \) for all \( E \in P \), then \( f \succneq g \) (since \( g \in \mathcal{F}_P \), either \( g \in \mathcal{U}(P) \) or \( g \in \mathcal{U}(Q) \) for some \( Q \succ P \), the statement holds in either case). By definition, \( \succneq \) satisfies Independence on \( \mathcal{F}_P \) for any \( P \in \mathcal{P}_\succneq \). Hence it follows from standard results that there exists a unique probability measure \( \pi_P : \sigma(P) \rightarrow [0, 1] \) such that

\[
V(f) = \sum_{E \in P} \pi_P(E)u(f(E))
\]

represents the restriction of \( \succneq \) to \( \mathcal{F}_P \). The utility function \( u \) is independent of \( P \) above because \( X \subseteq \mathcal{F}_P \) for every \( P \in \mathcal{P}_\succneq \).

By Lemma 2, for every \( f \in \mathcal{F} \), there exists \( P \in \mathcal{P}_\succneq \) such that \( f \sim w^P_f \). Moreover, for any \( Q \in \mathcal{P}_\succneq \), it follows from Understanding Monotonicity that \( f \succneq w^Q_f \). Thus, we can extend the function \( V \) defined above from \( \bigcup_{P \in \mathcal{P}_\succneq} \mathcal{F}_P \) to all of \( \mathcal{F} \) by setting

\[
V(f) = \max_{P \in \mathcal{P}_\succneq} \sum_{E \in P} \pi_P(E) \min_{s \in E} u(f(s)).
\]

The above lemmas ensure \( V \) is well defined and represents \( \succneq \). Notice that \( V \) satisfies monotonicity, i.e. if \( f(s) \succneq g(s) \) for all \( s \in S \), then \( V(f) \geq V(g) \). As such, \( \succneq \) satisfies monotonicity, which we will use in the sequel.

The remainder of the proof is to show that we can take each \( \pi_P \) above to be independent of \( P \), show that this new set function is defined on the appropriate domain, and satisfies the conditions set out in Definition 2. This relies on \( \mathcal{P}_\succneq \) being rich, which we next show is indeed true.

**Lemma 3.** For any \( P, Q \in \mathcal{P}_\succneq \), \( A \in \sigma(P) \), and \( B \in \sigma(Q) \) such that \( A \subset B \), there exists \( Q' \in \mathcal{P}_\succneq \) such that \( A, B \setminus A \in \sigma(Q') \).

**Proof.** Take \( P, Q, A, B \) as in the statement of the lemma. Assume \( A, B \) are nonempty, for if either is empty the result follows trivially. Let \( E = B \setminus A \). Suppose toward a contradiction that there is no \( Q' \in \mathcal{P}_\succneq \) such that \( A, E \in \sigma(Q') \). Fix any \( x, y, z \in X \) such that \( x > y > z \), and let \( f = xAyEz \). Let \( R \in \mathcal{P}_\succneq \) be such that \( f \in \mathcal{U}(R) \), which exists by Lemma 1. By assumption, either \( A \notin \sigma(R) \), \( E \notin \sigma(R) \), or both. If \( A \notin \sigma(R) \) and \( E \in \sigma(R) \), then there exists \( E' \in R \) such that \( E' \cap A \neq \emptyset \) and \( E' \cap B^c \neq \emptyset \) (in words, \( E' \) contains states in \( A \) and in \( B^c \)). By definition, \( x, z \in f(E') \) where \( x > z \) by assumption. Therefore, we can apply Claim 1 to \( E' \) and conclude that

\[
f \succeq f[(A \setminus E') \cup E] z,
\]

a contradiction to Contingent Monotonicity since \( f[(A \setminus E') \cup E] z \succeq z(s) f \) for any \( s \in E' \cap A \).
by monotonicity (and $E' \cap A$ is nonempty).

The case where $A \in \sigma(R)$ and $E \not\in \sigma(R)$ is symmetric. Lastly, suppose that $A \not\in \sigma(R)$ and $E \not\in \sigma(R)$. Then neither $A$ nor $E$ is in $R$. If $B \in \sigma(R)$, then there exists $B' \subseteq B$ such that $B' \in R$, $B' \cap A \neq \emptyset$, and $B' \cap E \neq \emptyset$. In this case, applying Claim 1 to $B'$ yields $f \preceq yB'f$, a contradiction to Contingent Monotonicity. If $B \not\in \sigma(R)$, then there exists $B' \in R$ such that $B' \cap B^C \neq \emptyset$ and either $B' \cap A \neq \emptyset$, $B' \cap E \neq \emptyset$, or both. In this case, similar arguments to above imply that $f \preceq zB'f$, a contradiction to Contingent Monotonicity.

Recall that $\mathcal{E}(P_\omega) = \bigcup_{P \in \mathcal{P}_\omega} \sigma(P)$ denotes the set of all events that are in the algebra generated by some partition in $\mathcal{P}_\omega$. We now define a set function $\pi : \mathcal{E}(P_\omega) \rightarrow [0, 1]$ by $\pi(E) = \pi_P(E)$ for any $P \in \mathcal{P}_\omega$ such that $E \in \sigma(P)$. Since $S \in \sigma(P)$ for all $P \in \mathcal{P}_\omega$ and each $\pi_P$ is a probability, $\pi$ is normalized. Moreover, the monotonicity of $\preceq$ implies that $\pi$ is weakly monotone. Notice that $\pi$ is well defined, for if there exists $E \in \mathcal{E}(P_\omega)$ such that $E \in \sigma(P) \cap \sigma(P')$ for some $P, P' \in \mathcal{P}_\omega$, we can choose $x, y \in X$ such that $u(x) = 1$ and $u(y) = 0$, so we must have

$$V(xEy) = \pi_P(E) = \pi_{P'}(E).$$

We will now show that $\pi$ is additive on $\mathcal{E}(P_\omega)$ (recall that this set is finite), whenever $\mathcal{E}(P_\omega)$ is closed under disjoint unions.

**Lemma 4.** Take any $E \in \mathcal{E}(P_\omega)$ such that $E = \bigcup_{i=1}^n E_i$, where $E_i \in \mathcal{E}(P_\omega)$ for all $i = 1, \ldots, n$ and $E_i \cap E_j = \emptyset$ for all $i, j$. Then $\pi(E) = \sum_{i=1}^n \pi(E_i)$.

**Proof.** We will prove the result by induction on $n$. Notice that $n$ must always be finite because $\mathcal{E}(P_\omega)$ is finite. We assume throughout that all events are nonempty, as the empty cases are trivial. We begin with the $n = 2$ case. So suppose $E, E_1, E_2, \in \mathcal{E}(P_\omega)$, where $E = E_1 \cup E_2$ and $E_1 \cap E_2 = \emptyset$. Let $P, P_1, P_2 \in \mathcal{P}_\omega$ denote their corresponding partitions in $\mathcal{P}_\omega$, so $E \in \sigma(P)$, $E_1 \in \sigma(P_1)$, and $E_2 \in \sigma(P_2)$. By Lemma 3, there exists $Q \in \mathcal{P}_\omega$ such that $E_1, E_2 \in \sigma(Q)$. Since $\pi_Q$ is additive and $\pi$ is well defined, we must have

$$\pi(E) = \pi_P(E) = \pi_Q(E) = \pi_Q(E_1) + \pi_Q(E_2),$$

as desired.

Now, take any $m \in \mathbb{N}$, and assume that the lemma is true for all $m' < m$. Take any $E \in \mathcal{E}(P_\omega)$ such that $E = \bigcup_{i=1}^m E_i$, where $E_i \in \mathcal{E}(P_\omega)$ for all $i = 1, \ldots, m$ and $E_i \cap E_j = \emptyset$ for all $i, j$. Let

---

24A set function $\mu$ is weakly monotone if for every $A, B$ such that $A \subseteq B$, $\mu(B) \geq \mu(A)$. 

A = \bigcup_{i=1}^{m-1} E_i. Then by the inductive step,

$$\pi(A) = \sum_{i=1}^{m-1} \pi(E_i).$$

By Lemma 3, there exists Q \in \mathcal{P}_z such that A, E_m \in \sigma(Q). Since \pi_Q is additive and \pi is well defined, we must have

$$\pi(E) = \pi_Q(E) = \pi_Q(A) + \pi_Q(E_m) = \sum_{i=1}^{m-1} \pi(E_i) + \pi(E_m) = \sum_{i=1}^{m} \pi(E_i),$$

as desired. \qed

The following lemma shows that \pi is strictly monotone on its domain, completing the proof of Theorem 1.

**Lemma 5.** \pi is strictly monotone on \mathcal{E}(\mathcal{P}_z).

**Proof.** Take any A, B \in \mathcal{E}(\mathcal{P}_z) such that A \subset B. By definition, there exists P, P' \in \mathcal{P}_z such that A \in \sigma(P) and B \in \sigma(P'). Therefore, by Lemma 3, there exists Q \in \mathcal{P}_z such that A, B \setminus A \in \sigma(Q).

Toward a contradiction, suppose that \pi(A) = \pi(B) (recall \pi is weakly monotone). Since \pi is additive, it follows that \pi(B \setminus A) = 0. Therefore, for any x, y, z \in X such that x \succ y \succ z, we have

$$V(xA \{B \setminus A\} z) = \pi(A)u(x) + \pi(B^c)u(z) \leq \pi(A)u(x) + \pi(B_x^c)u(z) \leq V(xA z),$$

so xA z \succeq xAy \{B \setminus A\} z (the first equality above follows from Lemma 7, whose proof does not rely on the strict monotonicity of \pi). However, for any s \in B \setminus A, Contingent Monotonicity implies that

$$xAy \{B \setminus A\} z \succ xAy \{(B \setminus A) \setminus \{s\}\} z,$$

while monotonicity implies that

$$xAy \{(B \setminus A) \setminus \{s\}\} z \succeq xA z,$$

a contradiction. \qed

A.2. Proof of Theorem 1—Necessity

Throughout, fix a RU representation \langle u, \pi, \mathcal{P} \rangle that represents a preference \succeq. For any f \in \mathcal{F}, define

$$\varphi(f) := \text{arg max}_{P \in \mathcal{P}} \sum_{E \in \mathcal{P}} \pi(E) \min_{s \in E} u(f(s))$$

to be the argmax correspondence. Define

$$\varphi^*(f) := \{P \in \varphi(f) \mid \exists Q \in \varphi(f) \text{ s.t. } Q \succeq P\}$$

to be the set of “minimal” solutions in the argmax.
Necessity of Order and Certainty Independence is routine and omitted. Continuity follows
given the following lemma, which we will use throughout the rest of this subsection as well.

**Lemma 6.** If \( f_n \to f \), there exists \( N \in \mathbb{N} \) such that \( \varphi(f_n) \subseteq \varphi(f) \) for all \( n \geq N \).

*Proof.* Take any sequence \( \{f_n\} \) such that \( f_n \to f \). Take any \( P \notin \varphi(f) \). Then there exists \( \delta > 0 \) such that \( V(f) - V(f | P) > \delta \). Since \( f_n \to f \) and \( u \) is continuous, there exists \( N_P \in \mathbb{N} \) such that \( |V(f) - V(f_n)| < \frac{\delta}{2} \) and \( |V(f | P) - V(f_n | P)| < \frac{\delta}{2} \) for all \( n \geq N_P \). This implies that \( P \notin \varphi(f_n) \) for all \( n \geq N_P \). Taking \( N = \max_{P \in \mathcal{P} \setminus \varphi(f)} N_P \) (which exists since \( \mathcal{P} \) is finite) yields the desired conclusion. \( \blacksquare \)

Notice that the lemma above also holds for \( \varphi^* \). The following two results are used in proving Lemma 8 and the necessity of Understanding Monotonicity and Contingent Monotonicity.

**Lemma 7.** For every \( P \in \mathcal{P} \) and \( f \in \mathcal{F}_P \), \( P \in \varphi(f) \).

*Proof.* Take any \( P \in \mathcal{P} \) and \( f \in \mathcal{F}_P \). Enumerate the events in \( P \) such that \( P = \{E_i\}_{i=1}^n \) and \( x_1 \succeq x_2 \succeq \cdots \succeq x_n \), where \( f = (x_i, E_i)_{i=1}^n \). Now, take any \( Q \in \mathcal{P} \). We will show that \( V(f | Q) \leq V(f | P) \). For each \( i \), let

\[
B_i = \bigcup \left\{ E \in Q \mid u(x_i) = \min_{s \in E} u(f(s)) \right\}.
\]

Thus, we can write

\[
V(f | Q) = \sum_{E \in Q} \pi(E) \min_{s \in E} u(f(s)) = \sum_{i=1}^n \pi(B_i) u(x_i) \leq V((x_i, B_i)_{i=1}^n)
\]

(since each \( B_i \) is a union of cells in \( Q \), each \( \pi(B_i) \) is well defined). Notice that by the definition of \( V(f | Q) \), we have

\[
B_1 \subseteq E_1 \\
B_2 \subseteq [E_1 \setminus B_1] \cup E_2 \\
B_3 \subseteq [E_2 \setminus B_2] \cup E_3 \\
\vdots \\
B_n \subseteq [E_{n-1} \setminus B_{n-1}] \cup E_n.
\]

Thus, it follows from monotonicity of \( V \) and the arguments above that

\[
V(f | Q) \leq V((x_i, B_i)_{i=1}^n) \leq V((x_i, E_i)_{i=1}^n) = V(f | P),
\]

as desired. \( \blacksquare \)
Notice that in the above proof, in the displayed set inclusions, at least one of inclusion must be strict whenever \( Q \succ P \). Thus, if we took \( f = (x_i, E_i)_{i=1}^n \) such that \( x_1 \succ x_2 \succ \cdots \succ x_n \), then it would follow from the strong monotonicity of \( \alpha \) that \( V(f \mid P) > V(f \mid Q) \) for any \( Q \succ P \). We will rely on this version of Lemma 7 in the proof of Lemma 8 below.

**Lemma 8.** For every \( P \in \mathcal{P} \) and \( f \in \mathcal{F} \), \( f \in \mathcal{U}(P) \) if and only if \( P \in \varphi^*(f) \).

*Proof.* Take any \( P \in \mathcal{P} \) and \( f \in \mathcal{U}(P) \). Toward a contradiction, suppose that \( P \notin \varphi^*(f) \), and hence \( P \notin \varphi(f) \).\(^{25}\) Notice also that if \( P' \gg P \), \( P' \notin \varphi(f) \), for this would contradict the fact that \( f \in \mathcal{U}(P) \) (\( P \) would no longer be minimal). This implies that \( V(f) > V(f \mid P') \), where \( P' = P \) or \( P' \gg P \). Now, fix some \( g \in \mathcal{F}_P \) such that \( f \sim g \). Then by the definition of \( \mathcal{U}(P) \), we must have \( \alpha f + (1 - \alpha)h \geq \alpha g + (1 - \alpha)h \) for all \( h \in \mathcal{F}_P \) and \( \alpha \in [0, 1] \).\(^{26}\) Take any \( h \in \mathcal{F}_P \) such that \( \varphi^*(h) = P' \), where \( P' = P \) or \( P' \gg P \).\(^{27}\) Then Lemma 6 implies that there exists \( \alpha \) sufficiently close to 0 yet positive such that \( P' \in \varphi(\alpha f + (1 - \alpha)h) \). Then we have

\[
V(\alpha f + (1 - \alpha)h) = V(\alpha f + (1 - \alpha)h \mid P')
\]

\[
= \alpha V(f \mid P') + (1 - \alpha)V(h)
\]

\[
< \alpha V(f) + (1 - \alpha)V(h)
\]

\[
= \alpha V(g) + (1 - \alpha)V(h)
\]

\[
= V(\alpha g + (1 - \alpha)h),
\]

so \( \alpha g + (1 - \alpha)h \succ \alpha f + (1 - \alpha)h \), a contradiction.

Conversely, if \( P \in \varphi^*(f) \), then \( f \sim w_f^P \). Therefore, Definition 4 implies that \( f \in \mathcal{U}(P') \) for some \( P' \gg P \). However, if \( P' \neq P \), then the first part of this proof would imply that \( P \) is not minimal, a contradiction. \( \blacksquare \)

Given this result, **Understanding Monotonicity** follows trivially.

**Lemma 9.** The preference \( \succeq \) satisfies **Contingent Monotonicity**.

*Proof.* Take any \( A, B \in \mathcal{E}(\mathcal{P}) \) such that \( A \subseteq B \). We may assume \( B \) is nonempty, for otherwise the axiom is not well defined. Let \( E = B \setminus A \). Take any \( f \in \mathcal{F}_{(A, E, B^c)} \). By the richness of the RU representation, there exists \( Q \in \mathcal{P} \) such that \( A, E \in Q \). Lemma 7 implies that \( Q \in \varphi(f) \).

Take any \( x \in X \) and \( s \in B \) such that \( f(s) \succ x \). Let \( g = x \{s\} f \). Without loss of generality, assume that \( s \in E \).\(^{28}\) Moreover, it is without loss of generality to assume that \( x > f(B^c), f(A) \)

\(^{25}\) If \( P \in \varphi(f) \setminus \varphi^*(f) \), then one can immediately see that \( P \) is not minimal and so \( f \notin \mathcal{U}(P) \).

\(^{26}\) This is not the exact definition of \( \mathcal{U}(P) \) stated in Definition 4, but can be obtained by considering a sequence \( g_n \to g \) with \( f > g_n \) for all \( n \in \mathbb{N} \).

\(^{27}\) This can be done because of the discussion following the proof of Lemma 7 and the fact that \( \mathcal{P} \) is finite.

\(^{28}\) If \( A = B \), so \( E \) is empty, then replace \( E \) with \( A \) and the argument goes through unchanged.
since the representation is monotonic. Again since the representation is monotonic, \( f \succeq g \), so suppose toward a contradiction that \( f \sim g \).

Take any \( P \in \varphi(g) \), so \( V(g) = V(g \mid P) \). Then there exists \( A_i, i = 1, 2, 3, 4 \) such that each \( A_i \in \sigma(P) \) and

\[
V(g \mid P) = \pi(A_1)u(f(E)) + \pi(A_2)u(x) + \pi(A_3)u(f(B^c)) + \pi(A_4)u(f(A)),
\]

where \( A_1, A_2 \subseteq E \). We consider two cases.

First, suppose \( A_2 \) is nonempty. Then since \( A_2 \subseteq E \) and \( \pi(A_2) > 0 \), \( V(g) = V(g \mid P) < V(f \mid P) \leq V(f) \), which is the desired conclusion.

Second, suppose \( A_2 \) is empty. Then it follows that \( V(f \mid Q) = V(g \mid P) = V(f \mid P) \). There are then two possibilities. Either \( P = Q \), which contradicts \( A_2 \) being empty, or \( P \neq Q \) and \( P \in \varphi(f) \). Given the discussion following Lemma 7, in the second case it must also hold that \( f(B^c) \sim f(A) \). In this case, it must be that \( A_1 \subseteq E \), and

\[
V(f \mid Q) = \pi(E)u(f(E)) + \pi(A)u(f(A)) + \pi(B^c)u(f(B^c))
\]

\[
= \pi(A_1)u(f(E)) + \pi(A \cup B^c \cup [E \setminus A_1])u(f(B^c))
\]

\[
= V(f \mid P),
\]
a contradiction since \( f(E) \succeq f(B^c) \) and \( \pi \) is strictly monotone.

The final part of the proof is to show that Aversion to Limited Understanding is satisfied.

\begin{lemma}
The preference satisfies Aversion to Limited Understanding.
\end{lemma}

\begin{proof}
Take any \( f \in F \) and \( x \in X \) such that \( f > xEf \) for some \( E \in P \in \mathcal{U}^{-1}(f) \) and \( f > xE'f \) for every nonsingleton \( E' \) such that \( E' \cap E \neq \emptyset \) and \( E' \subseteq Q \subseteq \mathcal{U}^{-1}(f) \). In other words, \( f > xQ(s)f \) for every \( s \in E \) and \( Q \in \mathcal{U}^{-1}(f) \) such that \( Q(s) \) is nonsingleton. Notice that this implies that \( w^E_f > x \) and \( w^{Q(s)}_f > x \) for all such \( Q(s) \), otherwise we would have \( xQ(s)f \succeq f \) and similarly for \( E \).

Now, let \( g = xAf \) for any \( A \subseteq E \). Take any \( R \in \varphi^*(g) \). There are two cases. First, if \( R \notin \varphi^*(f) \), then \( V(f) > V(f \mid R) \geq V(g \mid R) = V(g) \), as desired. So suppose \( R \in \varphi^*(f) \). Then by assumption, \( w^R_f > x \) for every \( s \in E \) such that \( R(s) \) is nonsingleton. Moreover, if \( R(s) \) is a singleton for every \( s \in A \), then since \( f(s) \succeq w^R_f \geq g(s) \) for every \( s \in E \), it follows that \( V(f) = V(f \mid R) > V(g \mid R) = V(g) \), as desired. So there must exist \( s \in A \) such that \( R(s) \) is nonsingleton. In this case, by assumption, \( w^{R(s)}_f > x \), so again \( V(f) = V(f \mid R) > V(g \mid R) = V(g) \), as desired.

This completes the proof of necessity.
\end{proof}
A.3. Proofs of Other Results in the Text

We first present a simple lemma that will be used in the proof of Theorem 2 below. Recall from Schmeidler (1989) that two acts \( f, g \in \mathcal{F} \) are comonotonic if there are no \( s, s' \in S \) such that \( f(s) > f(s') \) and \( g(s) < g(s') \).

**Lemma 11.** For any \( f, g \in \mathcal{F} \), there exists \( \alpha \in (0, 1) \) such that \( f \) and \( \alpha f + (1 - \alpha)g \) are comonotonic.

**Proof.** Take any \( s, s' \in S \), and without loss of generality suppose that \( f(s) > f(s') \) and \( g(s') > g(s) \), otherwise the result follows trivially. Per contra, suppose that for all \( \alpha \in (0, 1) \),

\[
\alpha f(s') + (1 - \alpha)g(s') > \alpha f(s) + (1 - \alpha)g(s).
\]

Then taking a sequence \( \{\alpha_n\} \subset (0, 1) \) such that \( \alpha_n \to 1 \), we have that \( \alpha_n f(s') + (1 - \alpha_n)g(s') > \alpha_n f(s) + (1 - \alpha_n)g(s) \) for all \( n \in \mathbb{N} \), so by Continuity we must have \( f(s') \geq f(s) \), a contradiction. ■

**Proof of Theorem 2.** Uniqueness of \( u \) is routine and omitted. Henceforth without loss of generality assume \( u = u' \). Let \( V \) and \( V' \) be the functionals corresponding to each RU representation in the statement of the theorem. We now show that \( \mathcal{P} \approx \mathcal{P}' \). In pursuit of a contradiction, suppose \( \mathcal{P} \not\approx \mathcal{P}' \), so (without loss of generality) there exists \( P \in \mathcal{P} \setminus \mathcal{P}' \) and there is no \( P' \in \mathcal{P} \cap \mathcal{P}' \) such that \( P' \gg P \). Notice that by Lemma 7, \( V \) satisfies Independence on \( \mathcal{F}_P \), so \( \succeq \) satisfies Independence on \( \mathcal{F}_P \). Thus, since \( V' \) also represents \( \succeq \), \( V' \) must satisfy Independence. We now proceed to show that \( V' \) must in fact violate Independence, a contradiction.

We consider two cases. First, suppose there exist comonotonic acts \( f, g \in \mathcal{F}_P \) such that \( \varphi'(f) \cap \varphi'(g) = \emptyset \). Without loss of generality, assume \( f \sim g \). Let \( \{\alpha_n\} \subset (0, 1) \) be such that \( \alpha_n \to 1 \). By Lemma 6, there exists \( N \in \mathbb{N} \) such that \( \varphi'(\alpha_n f + (1 - \alpha_n)g) \subseteq \varphi'(f) \) for all \( n \geq N \). Fix some \( n \geq N \) and let \( \alpha = \alpha_n \) for simplicity. Take any \( Q \in \varphi'(\alpha f + (1 - \alpha)g) \) and \( R \in \varphi'(g) \). Since \( f \) and \( g \) are comonotonic, it follows that

\[
V'(\alpha f + (1 - \alpha)g) = V'(\alpha f + (1 - \alpha)g | Q)
\]

\[
= \alpha V'(f | Q) + (1 - \alpha) V'(g | Q)
\]

\[
< \alpha V'(f) + (1 - \alpha) V'(g)
\]

\[
= V'(f),
\]

a contradiction since \( V' \) must satisfy Independence.

Now, let \( \{\mathcal{F}_P^c\}_{c \in C} \) denote the collection of subsets of \( \mathcal{F}_P \) such that for every \( f, g \in \mathcal{F}_P \), \( f \) and \( g \) are comonotonic. Clearly, \( \bigcup_{c \in C} \mathcal{F}_P^c = \mathcal{F}_P \), and if \( f, g \in \mathcal{F}_P^c \), then \( \alpha f + (1 - \alpha)g \in \mathcal{F}_P^c \) for every \( \alpha \in [0, 1] \). Moving to the second case, suppose that for every \( c \in C \) and \( f, g \in \mathcal{F}_P^c \), \( \varphi'(f) \cap \varphi'(g) \neq \emptyset \).
We claim that this implies that for every $c \in C$, there exists $Q_c \in P'$ such that $Q_c \in \varphi'(f)$ for every $f \in F_c^c$.

**Claim.** Under the hypotheses of Theorem 2, if for every $c \in C$ and $f, g \in F_c^c$, $\varphi'(f) \cap \varphi'(g) \neq \emptyset$, then there exists $Q_c \in P'$ such that $Q_c \in \varphi'(f)$ for every $f \in F_c^c$.

**Proof of Claim.** Fix any $f, g \in F_c^c$, and suppose toward a contradiction that there exists $h \in F_c^c$ such that $\varphi'(f) \cap \varphi'(g) \cap \varphi'(h) = \emptyset$. By assumption, there exists $Q \in \varphi'(f) \cap \varphi'(g)$. Moreover, since $f$ and $g$ are comonotonic, for any $R \in P'$ and $\alpha \in [0, 1]$, we have $V'(\alpha f + (1 - \alpha)g \mid R) = \alpha V'(f \mid R) + (1 - \alpha) V'(g \mid R)$, so it follows that for any $\alpha \in (0, 1)$, $\varphi'(f) \cap \varphi'(g) \subseteq \varphi'(\alpha f + (1 - \alpha)g)$.

We now show that the reverse inclusion holds. Thus, for some $\alpha \in (0, 1)$, suppose there exists $Q \in \varphi'(\alpha f + (1 - \alpha)g)$ such that $Q \notin \varphi'(f) \cap \varphi'(g)$. Without loss of generality, suppose $Q \notin \varphi'(f)$ and that $f \sim g$. Then we have

$$V'(\alpha f + (1 - \alpha)g) = V'(\alpha f + (1 - \alpha)g \mid Q)$$

$$= \alpha V'(f \mid Q) + (1 - \alpha) V'(g \mid Q)$$

$$< \alpha V'(f) + (1 - \alpha) V'(g)$$

$$= V(f),$$

a contradiction to Independence since $f, g \in F_p$. Thus, for every $\alpha \in (0, 1)$, $\varphi'(f) \cap \varphi'(g) = \varphi'(\alpha f + (1 - \alpha)g)$. Since $\varphi'(f) \cap \varphi'(g) \cap h = \emptyset$, this implies that $\varphi'(\alpha f + (1 - \alpha)g) \cap \varphi'(h) = \emptyset$, which contradicts the assumption that $\varphi'(f') \cap \varphi'(g') \neq \emptyset$ for all $f', g' \in F_c^c$, since $\alpha f + (1 - \alpha)g, h \in F_c^c$. ■

Given the claim above, suppose that for every $c \in C$, there exists $Q_c \in P'$ such that $Q_c \in \varphi'(f)$ for all $f \in F_c^c$. There are two subcases. First, suppose that there exists $c, c' \in C$ such that $Q_c \neq Q_c'$. Then take $f \in F_c^c$ and $g \in F_c'^c$ such that $f \sim g$ and $Q_c \notin \varphi'(g)$, which can be done by assumption (if $Q_c \in \varphi'(g)$ for all $g \in F_c'^c$, then we may without loss of generality assume $Q_c = Q_c'$). By Lemma 11, there exists $\alpha \in (0, 1)$ such that $f$ and $\alpha f + (1 - \alpha)g$ are comonotonic, so $Q_c \in \varphi'(\alpha f + (1 - \alpha)g)$. Thus, it follows that

$$V'(\alpha f + (1 - \alpha)g) = V'((\alpha f + (1 - \alpha)g \mid Q_c)$$

$$= \alpha V'(f \mid Q_c) + (1 - \alpha) V'(g \mid Q_c)$$

$$< \alpha V'(f) + (1 - \alpha) V'(g)$$

$$= V'(f),$$

a contradiction since $V'$ must satisfy Independence.

Lastly, suppose that $Q_c = Q_{c'} = Q$ for all $c, c' \in C$. In this case, $V'(f) = V'(f \mid Q)$ for all $f \in F_p$. Take any $E \in Q$. Since $Q$ is not finer than $P$, there exists disjoint $A_1, A_2 \in \sigma(P)$ such
that \( A_i \cap E \neq \emptyset \) for \( i = 1, 2 \) and \( E \subseteq A_1 \cup A_2 \). Take any \( x, x', y, y' \in X \) with \( x > y, x' > y' \), and \( xA_1y \sim x'A_2y' \). Without loss of generality, we may assume \( x > x' > y' > y \). Then we have

\[
V'(xA_1y) = \sum_{B \in \mathcal{Q}; \mathcal{B} \subseteq A_1 \setminus E} \pi(B)u(x) + \sum_{B \in \mathcal{Q}; \mathcal{B} \not\subseteq A_1 \setminus E} \pi(B)u(y)
\]

\[
V'(x'A_2y') = \sum_{B \in \mathcal{Q}; \mathcal{B} \subseteq A_2 \setminus E} \pi(B)u(x') + \sum_{B \in \mathcal{Q}; \mathcal{B} \not\subseteq A_2 \setminus E} \pi(B)u(y').
\]

Since \( A_1 \) and \( A_2 \) are disjoint, for any \( \alpha \in [0, 1] \) we have

\[
V'(\alpha xA_1y + (1 - \alpha)x'A_2y') = \sum_{B \in \mathcal{Q}; \mathcal{B} \subseteq A_1 \setminus E} \pi(B)u(\alpha x + (1 - \alpha)y') + \sum_{B \in \mathcal{Q}; \mathcal{B} \not\subseteq A_2 \setminus E} \pi(B)u(\alpha y + (1 - \alpha)x')
\]

\[
+ \pi(E) \min \{u(\alpha x + (1 - \alpha)y'), u(\alpha y + (1 - \alpha)x')\}
\]

\[
+ \sum_{B \in \mathcal{Q}; \mathcal{B} \not\subseteq A_1 \setminus E, \mathcal{B} \subseteq A_2 \setminus E} \pi(E)u(\alpha y + (1 - \alpha)y').
\]

Define a function \( V^* : [0, 1] \rightarrow \mathbb{R} \) by \( V^*(\alpha) = V'(\alpha xA_1y + (1 - \alpha)x'A_2y') \) for each \( \alpha \in [0, 1] \). It is easy to see that \( V^*(0) = V^*(1) \) since \( xA_1y \sim x'A_2y' \), \( V^* \) is continuous, and \( V^* \) is weakly concave. Moreover, since \( \pi(E) > 0 \) and \( x > x' > y' > y \), there exists \( \beta, \lambda, \alpha \in (0, 1) \) such that \( V^*(\alpha \lambda + (1 - \alpha)\beta) > \alpha V^*(\lambda) + (1 - \alpha)V^*(\beta) \). Thus, \( V^* \) is strictly concave on \([0, 1]\), so letting \( \alpha^* = \arg\max_{\alpha \in [0, 1]} V^*(\alpha) \) (which exists by the Weierstrass Theorem), we have \( V^*(\alpha^*) > V^*(0) \), and hence \( \alpha^* xA_1y + (1 - \alpha^*)x'A_2y' > xA_1y \sim x'A_2y' \), a contradiction to Independence.

Thus, we have shown \( \mathcal{P} \approx \mathcal{P}' \). Notice that if \( \mathcal{P} \neq \mathcal{P}' \), they only differ by adding a coarsening of some \( P \in \mathcal{P} \cap \mathcal{P}' \). Removing any such coarsenings does not affect the underlying preference, so assume without loss of generality that \( \mathcal{P} = \mathcal{P}' \). In this case, \( \mathcal{E}(\mathcal{P}) = \mathcal{E}(\mathcal{P}') \), and \( \pi \) and \( \pi' \) are extended from the same collection of unique probabilities \( \{\pi_P\}_{P \in \mathcal{P}} \), so it follows that \( \pi = \pi' \).

**Proof of Proposition 1.** The (ii) implies (i) case is trivial. So suppose that for every \( f \in \mathcal{F}, P \in \mathcal{P}_2, \) and \( g \in \mathcal{F}_P \),

\[
f \succ_2 g \implies f \succ_1 g.
\]

Notice then that for every \( P \in \mathcal{P}_2, \succ_1 \) and \( \succ_2 \) agree on \( \mathcal{F}_P \), and hence on \( X \), so \( u_1 = u_2 = u \) without loss of generality. We next show that \( \mathcal{P}_2 \subseteq \mathcal{P}_1 \). Toward a contradiction, suppose that \( \mathcal{P}_2 \not\subseteq \mathcal{P}_1 \), so there exists \( P \in \mathcal{P}_2 \) such that \( P \notin \mathcal{P}_1 \). By **Lemma 7**, \( \succ_2 \) satisfies Independence on \( \mathcal{F}_P \), and so must \( \succ_1 \). Thus, we can mimic the arguments in the proof of **Theorem 2** to obtain a contradiction since \( P \notin \mathcal{P}_1 \). So we have \( \mathcal{P}_2 \subseteq \mathcal{P}_1 \). Since \( V_1(f) = V_2(f) \) for all \( f \in \mathcal{F}_P \) and \( P \in \mathcal{P}_2 \) (by passing to certainty equivalents), it follows that we must have \( \pi_1(E) = \pi_2(E) \) for every \( E \in \mathcal{E}(\mathcal{P}_2) \).

**Proof of Proposition 2.** Throughout, we shorten \( \Omega(\mathcal{P}) \) to just \( \Omega \). First, suppose that \( \Omega \notin \mathcal{P} \). Then
$\succeq$ is represented by

$$V(f) = \sum_{E \in \Omega} \pi(E) \min_{s \in E} u(f(s)).$$

Let

$$\mathcal{M}_\Omega \coloneqq \{ \mu \in \Delta(S) \mid \forall E \in \Omega, \exists s \in E \text{ s.t. } \mu(s) = \pi(E) \}$$

denote the set of measures $\mu \in \Delta(S)$ such that for every $E \in \Omega$, $\mu(E) = \pi(E)$ and each conditional measure $\mu_E$ is a degenerate measure on some $s \in E$. Then it follows that we can write $V$ as

$$V(f) = \min_{\mu \in \mathcal{M}_\Omega} \sum_{s \in S} \mu(s) u(f(s)).$$

Taking the closed convex hull of $\mathcal{M}_\Omega$ if necessary (which does not affect the underlying preference), $V$ is a multiple priors representation (Gilboa and Schmeidler, 1989), so $\succeq$ is uncertainty averse.

Second, suppose $\succeq$ is uncertainty averse, but $\Omega \not\in \mathcal{P}$. Then $|\mathcal{P}| \geq 2$, for if it were a singleton that single partition would be $\Omega$. Now, fix some $P \in \mathcal{P}$. Note that there must exist $Q \in \mathcal{P}$ and $g \in \mathcal{F}_Q$ such that $V(g) = V(g \mid Q) > V(g \mid P)$, otherwise every $Q \in \mathcal{P}$ such that $Q \neq P$ is redundant, and the preference could be represented by $V(\cdot \mid P)$, in which case $P = \Omega$ and $\Omega \in \mathcal{P}$. Thus, take any $g$ as described, and any $f \in \mathcal{F}_P$ such that $f \sim g$. Then by Lemma 7, $P \in \varphi(f)$. Moreover, for $\alpha \in (0, 1)$ sufficiently large, we must have $P \in \varphi(\alpha f + (1 - \alpha)g)$ (see Lemma 6). This implies that

$$V(\alpha f + (1 - \alpha)g) = V(\alpha f + (1 - \alpha)g \mid P) = \alpha V(f) + (1 - \alpha)V(g \mid P) < V(f) + (1 - \alpha)V(g) = V(f),$$

contrary to uncertainty aversion. ■

Proof of Proposition 3. First, notice that if $\Omega \in \mathcal{P}$, the functional

$$V(f) = \sum_{E \in \Omega} \pi(E) \min_{s \in E} u(f(s))$$

represents $\succeq$. It is easy to see that $V$ satisfies Comonotonic Independence (and so must $\succeq$), so by Schmeidler (1989), there exists a unique capacity $\nu$ such that the Choquet representation $\langle u, \nu \rangle$ represents $\succeq$. Moreover, $\succeq$ is biseparable (Ghirardato and Marinacci, 2001). Without loss of generality, fix $x, y \in X$ with $u(x) = 1$ and $u(y) = 0$. Hence, by Ghirardato and Marinacci (2001), for every $E \in \Sigma$,

$$\nu(E) = V(xEy) = \sum_{A \in \Omega} \pi(A) \mathbb{1}(A \subseteq E) = \max_{A \in \sigma(\Omega) : A \subseteq E} \pi(A),$$
as desired (the final equality follows from the additivity of $\pi$).

\textbf{Proof of Proposition 4.} Suppose that $\Omega \in \mathcal{P}$. Then by Proposition 3, $\succeq$ has a CEU representation $\langle u, \nu \rangle$ where

\begin{equation}
\nu(E) = \max_{A \in \sigma(\Omega), A \subseteq E} \pi(A)
\end{equation}

for every $E \in \Sigma$. Recall also that the preference can be represented by $V(\cdot \mid \Omega)$. Notice that for every $x, y \in X$, $f, g \in \mathcal{F}$, and $E \in \sigma(\Omega)$, $xEf \succeq xEg$ if and only if $yEf \succeq yEg$. This implies any $E \in \sigma(\Omega)$ is unambiguous.

Suppose $E$ is unambiguous but $E \notin \sigma(\Omega)$. Let $\{E_i\}_{i=1}^n$ enumerate the events such that $E_i \subseteq E$ and $E_i \in \Omega$ for all $i$. Since $E \notin \sigma(\Omega)$, there exists $A_1, A_2 \in \Sigma$ such that $A_1 \subseteq E$, $A_2 \subseteq E^c$, and $A_1 \cup A_2 \in \Omega$. Notice that since $A_2$ is a proper subset of a cell in $\Omega$, $\nu(A_2) = 0$. If $A_2$ is nonsingleton, fix any $s \in A_2$, so $\nu(\{s\}) = \nu(A_2) = 0$. However, we have

$$
\nu(E \cup A_2) = \pi \left( [A_1 \cup A_2] \bigcup [\cup_i E_i] \right) > \nu(E \cup \{s\}) = \pi(B^*)
$$

for some $B^* \subset [A_1 \cup A_2] \bigcup [\cup_i E_i]$ (recall $\pi$ is strictly monotone and the definition of $\nu$ from Proposition 3). Thus, by Lemma 7.2 of Epstein and Zhang (2001), $E$ cannot be unambiguous, a contradiction.

If $A_2$ is a singleton, then $\nu(E^c \setminus A_2) = \nu(E^c)$, for if $B \in \sigma(\Omega)$ and $B \subseteq E^c$, then $A_2 \notin B$, since $A_1 \notin B$ and $A_1 \cup A_2 \in \Omega$ by assumption. Therefore, $B$ must also be a subset of $E^c \setminus A_2$. However,

$$
\nu(E \cup E^c) = 1 > \nu(E \cup E^c \setminus A_2) = 1 - \pi(A_1 \cup A_2),
$$

a contradiction to Lemma 7.2 in Epstein and Zhang (2001) given that $\pi$ is strictly monotone.

\textbf{Proof of Proposition 5.} Throughout, assume that $\mathcal{P}$ is minimal without loss of generality. Suppose that Dynamic Consistency holds. Toward a contradiction, suppose there exists $P \in \mathcal{P}$ such that $A \notin \sigma(P)$. Then there exists $E \in P$ such that $E \cap A \neq \emptyset$ and $E \cap A^c \neq \emptyset$. Fix any $x, y, z \in X$ such that $x > y > z$. Find $h \in \mathcal{F}_P$ such that $\varphi(h) = \{P\}$, which must exist by the uniqueness of $\mathcal{P}$ (Theorem 2). By Lemma 6, there exists $\alpha \in (0, 1)$ such that $P \in \varphi(\alpha x [A \cap E] z + (1 - \alpha)h) \cap \varphi(\alpha y [A \cap E] z + (1 - \alpha)h)$. Thus, it follows that

$$
V(\alpha x [A \cap E] z + (1 - \alpha)h) = \alpha V(x [A \cap E] z \mid P) + (1 - \alpha) V(h) = \alpha u(z) + (1 - \alpha) V(h)
$$
\[ V(\alpha x [A \cap E] z + (1 - \alpha)h) = \alpha V(y [A \cap E] z \mid P) + (1 - \alpha)V(h) = \alpha u(z) + (1 - \alpha)V(h), \]

so \( \alpha x [A \cap E] z + (1 - \alpha)h \sim ay [A \cap E] z + (1 - \alpha)h \). Thus, since \( \alpha x [A \cap E] z + (1 - \alpha)h(s) = ay [A \cap E] z + (1 - \alpha)h(s) \) for all \( s \in A^c \), Dynamic Consistency would require that \( \alpha x [A \cap E] z + (1 - \alpha)h \sim_A ay [A \cap E] z + (1 - \alpha)h \). However, \( P \cap \{A, A^c\} \in \mathcal{P}_A \), so \( A \cap E \in P_A \). Since \( \alpha x [A \cap E] z + (1 - \alpha)h \) and \( ay [A \cap E] z + (1 - \alpha)h \) are \( P_A \)-measurable, it follows that

\[
V_A(\alpha x [A \cap E] z + (1 - \alpha)h) = \alpha [\pi_A(A \cap E)u(x) + (1 - \pi_A(A \cap E))u(z)] + (1 - \alpha)V_A(h) \\
> \alpha [\pi_A(A \cap E)u(y) + (1 - \pi_A(A \cap E))u(z)] + (1 - \alpha)V_A(h) \\
= V_A(\alpha x [A \cap E] z + (1 - \alpha)h),
\]

a contradiction to dynamic consistency. \[\blacksquare\]

**Proof of Proposition 6.** First, take any \( P \in \mathcal{P} \) and \( A \in P \). We will show \( A \) is foreseen. As such, fix any \( x, x', y, y' \in X \) such that \( u(x) + \delta u(y) = u(x') + \delta u(y') \), i.e. the act \( (xAx', yAy') \) is effectively certain. Since \( A \in P \in \mathcal{P} \) it follows from Lemma 7 that

\[
W(xAx', yAy') = \pi(A)u(x) + \pi(A^c)u(x') + \delta \pi(A)u(y) + \delta \pi(A^c)u(y') \\
= \pi(A) [u(x) + \delta u(y)] + \pi(A^c) [u(x') + \delta u(y')] \\
= u(x) + \delta u(y),
\]

since \( (xAx', yAy') \) is effectively certain. Therefore, \( (xAx', yAy') \) is subjectively certain, so \( A \) is foreseen as desired.

Conversely, take any foreseen \( A \in \Sigma \), and suppose toward a contradiction that there is no \( P \in \mathcal{P} \) such that \( A \in P \). Fix \( x, x', y, y' \in X \) as above, and assume without loss of generality that \( x > x' \) and \( y' > y \). Take any \( P \in \varphi(xAx') \) and \( Q \in \varphi(yAy') \). Then there exists \( B \in \sigma(P) \) and \( E \in \sigma(Q) \) such that \( V(xAx') = \pi(B)u(x) + \pi(B^c)u(x') \) and \( V(yAy') = \pi(E)u(y) + \pi(E^c)u(y') \). Thus we have

\[
W(xAx', yAy') = \pi(B)u(x) + \pi(B^c)u(x') + \delta \pi(E)u(y) + \delta \pi(E^c)u(y').
\]

Since \( A \) is foreseen, \( (xAx', yAy') \) is subjectively certain, so it follows that \( W(xAx', yAy') = u(x) + \delta u(y) \). Combining the two expressions above, we have that

\[
\pi(B^c)u(x) + \delta \pi(E^c)u(y) = \pi(B^c)u(x') + \delta \pi(E^c)u(y'),
\]
and equivalently
\[ u(x) + \delta \frac{\pi(E^c)}{\pi(B^c)} u(y) = u(x') + \delta \frac{\pi(E^c)}{\pi(B^c)} u(y'). \]
Moreover, notice that since \( A \notin P \) for any \( P \in \mathcal{P} \), we must have \( B \subset A \subset E \), so by the strict monotonicity of \( \pi \) it follows that \( \pi(E) > \pi(B) \) and hence \( \pi(B^c) > \pi(E^c) \), so \( \frac{\pi(E^c)}{\pi(B^c)} \neq 1 \). However, since \((xA\sigma', yA\sigma')\) is effectively certain, the previous displayed equation simplifies to
\[ \delta \left[ \frac{\pi(E^c)}{\pi(B^c)} - 1 \right] u(y) = \delta \left[ \frac{\pi(E^c)}{\pi(B^c)} - 1 \right] u(y'), \]
which contradicts the fact that \( y' > y \) since \( \left[ \frac{\pi(E^c)}{\pi(B^c)} - 1 \right] \neq 0 \).

**Proof of Proposition 7.** We begin with (i). Suppose \( A \in \sigma(\Omega) \), and take any \( E \in \Sigma \). First, notice that since \( E \cap A \) and \( E \cap A^c \) are disjoint, the convexity of \( \nu \) implies that
\[ \nu(E \cap A) + \nu(E \cap A^c) \leq \nu([E \cap A] \cup [E \cap A^c]) = \nu(E). \]
As such, suppose toward a contradiction that \( \nu(E) > \nu(E \cap A) + \nu(E \cap A^c) \). Take any
\[ B^* \in \arg \max_{B \in \sigma(\Omega) : B \subseteq E} \pi(B), \]
so \( \nu(E) = \pi(B^*) \). Notice that since \( A, B^* \in \sigma(\Omega) \) and \( \pi \) is additive on \( \sigma(\Omega) \), \( \pi(B^*) = \pi(B^* \cap A) + \pi(B^* \cap A^c) \). Moreover, by the definition of \( \nu \), \( B^* \subseteq E \), so \( B^* \cap A \subseteq E \cap A \) and \( B^* \cap A^c \subseteq E \cap A^c \). These facts together imply that
\[ \pi(B^* \cap A) + \pi(B^* \cap A^c) = \pi(B^*) = \nu(E) = \nu(E \cap A) + \nu(E \cap A^c) \geq \pi(B^* \cap A) + \pi(B^* \cap A^c), \]
a contradiction (the rightmost inequality follows from the definition of \( \nu \)).

Conversely, take any \( A \in \Sigma \) and suppose that for every \( E \in \Sigma \), \( \nu(E) = \nu(E \cap A) + \nu(E \cap A^c) \). Suppose toward a contradiction that \( A \notin \sigma(\Omega) \). Let \( \{A_i\}_{i=1}^n \) enumerate the events such that \( A_i \cap A \neq \emptyset \) and \( A_i \in \Omega \). Let \( E = \bigcup_{i=1}^n A_i \), so \( A \subset E \) (\( A \neq E \) since \( E \in \sigma(\Omega) \)). Notice that since each \( A_i, A_j \) are pairwise disjoint, \( E \cap A^c = E \setminus A = \bigcup_i [A_i \setminus A] \). As such, \( \nu(E \cap A^c) = 0 \), for if not, then there must exist \( B' \subseteq \bigcup_i [A_i \setminus A] \) and \( B \subseteq B' \) such that \( B \in \Omega \), which is impossible since \( \Omega \) is a partition. Moreover, notice that since \( A \notin \sigma(\Omega) \) and \( E \supset A \), we must have \( \nu(E) > \nu(E \cap A) = \nu(A) \) (recall that \( \pi \) is strictly monotone). Thus, we have constructed an \( E \in \Sigma \) such that \( \nu(E) > \nu(E \cap A) + \nu(E \cap A^c) \), a contradiction.

---

29If \( \pi(B) = 0 \), then since \( A \) is foreseen, we have \( W(xA\sigma', yA\sigma') = u(x') + \delta \pi(E)u(y) + \delta \pi(E^c)u(y') = u(x') + \delta u(y') \). This implies that \( \pi(E)u(y) = \pi(E)u(y') \), again a contradiction to \( y' > y \), since \( \pi(E) > 0 \) follows since \( E \supset B \).
As for (ii), if $\sigma(\Omega) = \Sigma$, then $\Sigma = \sigma(\Omega) = \mathcal{E}(\mathcal{P})$, so $\pi$ is a probability on $(S, \Sigma)$. Therefore, by the definition of $\nu$, $\nu(E) = \pi(E)$ for all $E \in \Sigma$ and hence $\nu$ is additive. Conversely, if $\nu$ is additive, $\langle u, \nu \rangle$ is a SEU representation for $\succeq$, so $\succeq$ satisfies Independence on $\mathcal{F} = \mathcal{F}_{\{s\}_{s \in S}}$, so $\Omega = S$ and hence $\sigma(\Omega) = \Sigma$. ■


