# Preferences for Partial Information and Ambiguity*

Jian Li†

April 8, 2015

**Abstract**

This paper studies intrinsic preferences for how information is revealed. We enrich the standard dynamic choice model in two dimensions. First, we introduce a novel choice domain that allows preferences to depend on how information is revealed. Second, conditional on a given information partition, we allow preferences over state-contingent outcomes to depart from expected utility axioms. In particular, we accommodate ambiguity sensitive preferences. We establish that a dynamically consistent decision maker (DM) is averse to partial information if and only if her static preferences satisfy a property called Event Complementarity. We show that Event Complementarity is closely related to ambiguity aversion in popular families of ambiguity preferences.

**Keywords:** Value of information, ambiguity aversion, dynamic choice, recursive utilities, learning.

*This paper is adapted from the first chapter of my dissertation and has been previously circulated under the title "Preferences for Information and Ambiguity". I am especially indebted to my advisor Chris Shannon for her continuous guidance and support. I am grateful to David Ahn for his constant input on the project, and Haluk Ergin, Matthew Rabin, and Fabian Lange for their valuable suggestions on improving the paper. I have benefited from discussions with Simone Cerreia-Vioglio, Larry Epstein, Shachar Kariv, Faruk Gul, Fabio Maccheroni, Juan Sebastián Lleras, Aniko Oery, Wolfgang Pesendorfer, Marciano Siniscalchi, and Junjie Zhou, and seminar and conference participants at McGill, Bonn, Syracuse, Bocconi, University of Montreal, U Penn Decision Theory Day, Decision Theory at Caltech Mini-Conference, Kansas Theory Workshop, RUD 2012, CETC 2013. All errors are my own.

†Department of Economics, McGill University. Email: jian.li7@mcgill.ca

# 1 Introduction

Bob is an economics PhD student who is on the job market. In December, Bob has submitted many job applications and is concerned about his prospects the following March. His future job outcomes depend not only on his own quality and performance, but also on uncertain factors like the quality of other candidates as well as tastes of different employers. Starting in late December and early January, online forums like the Blu-Wiki and Economics Job Market Rumors post interview and fly-in schedules for different schools, which provide partial information about the uncertainties. We observe that Bob avoids (with an obvious effort) ever looking at these rumors. One could say that Bob is *averse to partial information* (API): he prefers uncertainties to be revealed in one-shot.

Bob's behavior cannot be rationalized by the classic Bayesian model (Blackwell, 1951, 1953), where more information is always weakly preferable. Anecdotes and empirical evidence suggest that information aversion is prevalent in important life decisions such as financial investment or health choices. For example, some investors dislike overly frequent feedback information on their portfolio performance, and higher feedback frequency tends to lower willingness to take risk.[1] In the medical literature, it is well-documented that some patients are averse to receive test information.[2] In this case, a patient with family history of hereditary breast-ovarian cancer might choose not to take a genetic test and find out whether she carries the high-risk BRCA1 gene.

In this paper, we explore how API is related to the choice pattern illustrated by the Ellsberg paradox and known as ambiguity aversion. We show that Ellsbergian behaviors can be explained by a property of preferences that we call event complementarity (EC). We then show that API is equivalent to EC. Following this we demonstrate that many of the leading preference formulations for ambiguity aversion are consistent with EC, including maxmin EU, Choquet EU, and multiplier preferences.

[1] In a dynamic portfolio choice experiment, Bellemare et al. (2005) find that when a DM is committed to some given portfolio, higher frequency of feedback information leads to a lower ex-ante willingness to invest in risky assets. Gneezy et al. (2003) cite the event that in February 1999 an Israeli mutual fund, Bank Hapoalim, announced to lower the frequency of performance updates from monthly to quarterly in order to encourage the investors to hold more assets.

[2] Lerman et al. (1996) documented that when asked in an education session, 40% of the adults from families of hereditary breast-ovarian cancer refuse BRCA1-genetic test results. Lerman et al. (1999, 1996) find that 57% of the family members of hereditary nonpolyposis colon cancer do not want to learn about genetic test results during a phone interview, and 48% (more at-risk adults) decline in a structure interview. Besides insurance, quoted barriers to testing include test accuracy and emotional reaction. See Koszegi (2003) and Hoy et al. (2014) on the decision models of patient information aversion.

Understanding information preferences in an ambiguous environment is important for a number of reasons. First, from a theoretical perspective, decisions to acquire information will depend on preferences for partial information and will in general depart from those of standard expected utility maximizing agents. In particular, our work relates to a common criticism levelled at the ambiguity models: in a stationary environment, we would expect that ambiguity will eventually be learnt away. However, when learning is endogenous, ambiguity aversion implies intertemporal preferences that can undermine the incentive to collect new information, and hence ambiguity may persist in the long-run steady state. Second, of more direct policy relevance, recent work illustrates the importance of ambiguity in finance and macroeconomics for providing more accurate and robust dynamic measures of risk in financial positions.[3] Our results suggest that the nature and timing of information could be an important additional component to include in the design of risk measures that account for ambiguity. Third, our main result sheds new light on the empirical testing of ambiguity aversion. In a laboratory setting, an experimenter has the freedom to carefully design a set of static choice questions, which allows her to clearly test ambiguity aversion. Yet with naturally occurring data, it is often harder to isolate the effect of ambiguity aversion from the other confounding factors. Our main result shows that API is equivalent to EC in a fairly general setting. Hence evidence of API could also be considered as evidence of ambiguity aversion.

To accommodate these information preferences, we enrich the standard dynamic choice model in two dimensions. First, we introduce a novel choice domain that allows for preferences over state-contingent outcomes to depend on the way information is dynamically revealed. Second, conditional on a given information partition, we allow preferences to depart from expected utility axioms (Anscombe and Aumann, 1963; Savage, 1954). And more specifically, we look at a setting where uncertainties are resolved in two stages and the intermediate information could be partial. State-contingent outcomes are only realized in the second stage. Hence our model primitives, which could be ambiguity sensitive, are the ex-ante preferences over both intermediate information and state-contingent outcomes.

The first main result (Theorem 1, Theorem 3) shows that a DM is averse to partial information *if and only if* her static preferences display ambiguity aversion in a way characterized by a property called event complementarity. A DM is *averse to partial information* if, fixing any state-contingent outcomes, she would rather have all uncertainties resolved all at once in

[3] For applications of ambiguity in finance and macroeconomics, see Epstein and Wang (1994), Hansen and Sargent (2001), Cao et al. (2005), and Ju and Miao (2012). In addition, Epstein and Schneider (2010) survey applications of ambiguity preferences in finance, and Backus et al. (2005) survey applications of ambiguity preferences in macroeconomics. For work on dynamic risk measures under ambiguity, see Riedel (2004) and Acciaio et al. (2012) and references therein.

the end than gradually learning intermediate partial information. For example, the job candidate Bob, or a patient who does not want to find out whether she is a carrier of the BRCA1 gene mutation, displays partial information aversion. The property *event complementarity* is motivated by the literature on the behavioral definition of ambiguity (Epstein and Zhang, 2001; Zhang, 2002).[4] Intuitively, a DM perceives an event to be ambiguous if she does not know its probability. For two overlapping events $E$ and $F$ that are both unambiguous, it is possible that $E \cap F$ is ambiguous. Particularly, let $A = E \cap F$ and $B = E^c \cap F$. Suppose the DM could place bets on these events. Then event complementarity captures the intuition that

$$U(\text{bet on event } F) \geq U(\text{bet on event } A) + U(\text{bet on event } B)$$

The above inequality reflects ambiguity aversion: the DM is more averse to betting on the ambiguous events $A$ and betting on $B$ than betting on their disjoint union $F$, if the latter is unambiguous. Hence the joint evaluation of two events can potentially create a complementary effect and boost valuation. In the motivating example, Bob might have an idea about the overall probability of himself "finding a good job", and the overall probability that "the market for his research expertise is weak", but he might not know very well his likelihood of finding a good job when the market for his research expertise is weak. Thus Bob becomes more pessimistic about his job prospects if he *anticipates* to find out whether the market is weak or not before his final job outcomes, which forces him to think through the difficult question of how his job prospect will be if the market is weak.

The second set of results explores how preferences for partial information are connected with ambiguity attitudes in four families of ambiguity preferences commonly used in applications. We find that for the maxmin expected utility (MEU) preferences (Gilboa and Schmeidler, 1989) and the Choquet expected utility (CEU) preferences (Schmeidler, 1989), ambiguity aversion implies aversion to partial information and ambiguity loving implies attraction to partial information. The multiplier preferences (MP) (Hansen and Sargent, 2001; Strzalecki, 2011) exhibit partial information neutrality. Finally, we test this connection in the most general variational preferences (VP) family (Maccheroni et al., 2006a), with the help of a simple updating rule (Theorem 4) that we characterize in this paper. Under this updating rule, the connection between ambiguity aversion and partial information aversion is more complex and an ambiguity averse variational preferences agent might display partial information loving. In this case, we identify a necessary and sufficient global condition for partial information aversion and sufficient local conditions for partial information aversion.

[4] Epstein and Zhang (2001); Zhang (2002) consider the set of events endogenously revealed to be unambiguous, which has the structure of a $\lambda$-system: particularly, the intersection of two unambiguous events might be ambiguous. Earlier papers on this topic include Epstein (1999); Ghirardato et al. (2004). More recent critiques include Amarante and Filiz (2007); Kopylov (2007).

The third set of results (Theorem 1, Theorem 2) is two nested representation theorems. The main innovation here is to model dynamic preferences under purely subjective uncertainty by *relaxing reduction*.[5] Let $S$ denote the set of states of the world and $X$ denote the set of outcomes. An act $f : S \mapsto X$ describes a list of state-contingent outcomes. Let $E \subseteq S$ denote an event and a partition $\pi = \{E_1, \ldots, E_n\}$ of $S$ denote intermediate partial information.[6] Then the first utility representation says the DM's ex-ante utility of an act $f$ when anticipating information $\pi$ is

$$V(\pi, f) = I_S \begin{pmatrix} V_{E_1}(f) & \text{if } s \in E_1 \\ \vdots & \vdots \\ V_{E_n}(f) & \text{if } s \in E_n \end{pmatrix}, \tag{1}$$

where $V_{E_i}(\cdot)$ is the interim utility from $f$ conditional on event $E_i$, and $I_S : u(X)^{|S|} \mapsto \mathbb{R}$ describes how conditional utilities are aggregated.

Representation (1) describes a folding-back evaluation procedure: when a DM anticipates to learn intermediate partial information $\pi$, her utility from a state-contingent act $f$ is calculated by backward induction—she first accesses her interim utility of the act $f$ conditional on knowing an event $E_i$, and then aggregates these interim conditional utilities $\{V_{E_1}(f), \ldots, V_{E_n}(f)\}$ into the ex-ante utility $V(\pi, f)$ based on the perceived "likelihoods" of the events $\{E_1, \ldots, E_n\}$. In this way, utilities from acts are recursively calculated and dependent on the anticipated information partitions. This could be roughly viewed as a subjective analogue of Segal's (1990) model of two-stage compound lotteries.[7]

The second representation builds on (1) by further requiring that all interim conditional preferences are updated from some given one-shot unconditional preferences. Specifically, given act $f$ and event $E_i$, the interim $E_i$-conditional utility $V_{E_i}(f) := u(x_i)$ is determined by the unique solution to

$$u(x_i) = I_0 \begin{pmatrix} u(f(s)) & \text{if } s \in E_i \\ u(x_i) & \text{if } s \in E_i^c \end{pmatrix}, \tag{2}$$

where $x_i$ is the $E_i$-conditional certainty equivalent to $f$, $u : X \mapsto \mathbb{R}$ is an affine utility function over sure outcomes, and $I_0 (= I_S) : u(X)^{|S|} \mapsto \mathbb{R}$ is an unconditional aggregator of

[5] As noted by Siniscalchi (2011), a modeler of dynamic choice under ambiguity necessarily faces a tradeoff among four assumptions: dynamic consistency, consequentialism, unrestricted ambiguity, and reduction. It is almost a folk theorem in this literature. This paper fits into the big picture by relaxing reduction and accommodating the other three assumptions. Papers relaxing other assumptions are discussed in the literature review.

[6] A partition of the set $S$ is a collection of mutually exclusive events whose union is $S$.

[7] See Remark 4 for detailed comparison between our model and Segal (1990).

state-contingent utilities that is continuous, strongly monotone, normalized, and vertically invariant.

Intuitively, the updating rule (2) says conditional on event $E_i$, utility from act $f$ is determined by finding a conditional certainty equivalent $x_i$ such that the DM is indifferent between (i) getting the state-contingent outcomes of $f$ if the event is true and getting the certainty equivalent $x_i$ otherwise, and (ii) getting the certainty equivalent $x_i$ no matter what. It naturally generalizes the classic Bayesian updating rule.[8] We show that (2) is well-defined for all unconditional aggregator $I_0$ that is strongly monotone and Lipchitz continuous of rank 1, and hence works as a natural updating rule for all strongly monotone and weakly certainty independent one-shot preferences. As a result, in our second representation all conditional preferences and ex-ante preferences are pined down by the one-shot unconditional preferences representation $(u, I_0)$.[9]

Beyond the main results, Section 6 provides an extension on the value of information under ambiguity. Although perfect information is always preferred, the value of an information partition is neither monotone in its fineness nor the DM's degree of ambiguity aversion. We apply our model to study information acquisition in portfolio choice problem under ambiguity (Dow and Werlang, 1992). Section 7 discusses the case of second-order belief model (Klibanoff et al., 2005).

## 1.1 Related Literature

This paper belongs to the literature on dynamic decision making under ambiguity. Epstein and Schneider (2003) axiomatize recursive preferences over adapted consumption processes where all conditional preferences are maxmin expected utility (MEU), and find that dynamic consistency (our $\pi$-Recursivity) implies that the prior belief set has to satisfy a "rectangularity" restriction. Later work axiomatizes recursive preferences for other static ambiguity preferences and finds similar restrictions (Klibanoff et al., 2008; Maccheroni et al., 2006b).

Siniscalchi (2011) shows that within a given filtration, dynamic consistency implies Savage's Sure-Thing Principle and Bayesian updating. Together with reduction, dynamic consistency rules out modal Ellsberg preferences and thus ambiguity.[10] To allow for ambiguity, Siniscalchi studies preferences over a richer domain of decision trees, and relaxes dynamic consistency

[8] The updating rule was introduced by Pires (2002) for updating MEU and used by Eichberger et al. (2007) to update CEU.

[9] And $u$ is unique up to positive affine transformations and $I_0$ is unique for a given $u$.

[10] See also earlier work by Epstein and LeBreton (1993).

by introducing a weaker axiom called Sophistication. Together with auxiliary axioms, he proposes a general approach where preferences can be dynamically inconsistent, and the DM addresses these inconsistencies through Strotz-type Consistent Planning.

In this paper, we start from the observation that the noted tension between dynamic consistency and ambiguity relies on reduction, that is, on the assumption that the DM is indifferent to the temporal resolution of uncertainties. However, experimental evidence suggests that reduction is often violated in environments with objective risk. For example, Halevy (2007) finds evidence for non-reduction of compound lotteries and ambiguity aversion, as well as a positive association between the two. In a dynamic portfolio choice experiment, Bellemare et al. (2005) find that when a DM is committed to some ex-ante portfolio, higher frequency of information feedback leads to lower willingness to invest in risky assets. In this paper, we explore how dynamic consistency and unrestricted ambiguity preferences can be reconciled by relaxing reduction.

Thus this paper is also related to a rich literature that relaxes reduction and studies intrinsic preferences for early or late resolution of uncertainty. It was pioneered by Kreps and Porteus (1978) who introduced a novel domain of objective temporal lotteries to study preferences for temporal resolution of uncertainties. Epstein and Zin (1989, 1991) subsequently extend the Kreps-Porteus preferences to study asset pricing. Grant et al. (1998) link time preferences to intrinsic information preferences.[11][12] While these earlier literature directly models information preferences, more recent literature suggest information preferences can emerge *endogenously* as a consequence of how preferences deviate from the canonical Baysian EU assumptions. In a purely subjective domain, Strzalecki (2013) shows that even with standard discounting most models of ambiguity aversion display some preference with regard to the timing of resolution of uncertainty, with the notable exception of the MEU model.[13] Closest to our paper, recent work studies preferences for one-shot versus gradual resolution of (objective) uncertainty. In the domain of objective two-stage compound lotteries,[14] Dillenberger (2010) identifies a link between preferences for one-shot resolution of uncertainty and Allais-type behaviors. In their reference-dependent utility model, Koszegi and Rabin (2009) find preferences for getting information “clumped together rather than apart.” In

[11]Information preferences are *intrinsic* if the DM can only take one action, i.e., information has no decision value.

[12]Two applied models are also relevant. Koszegi (2003) finds patient anxiety as an explanation for information aversion. Hoy et al. (2014) suggest ambiguity aversion can account for the puzzle that take-up rate for genetic tests is low.

[13]Specifically, Strzalecki (2013) looks at an environment with identical and indistinguishable (i.i.d) ambiguity. We abstract away from the time preferences he identifies by imposing Time Neutrality. Our focus is on learning and preferences for sequential resolution of uncertainties under ambiguity.

[14]Segal (1990) was the first to study two-stage compound lotteries without reduction.

contrast, here we identify a link between ambiguity attitudes and intrinsic preferences for partial information in a purely subjective environment.

Finally, our work is also related to the literature on updating under ambiguity.[15] Pires (2002) introduces a coherence property that characterizes the prior-by-prior Bayesian updating rule for MEU preferences. Eichberger et al. (2007) then apply this coherency property to characterize full Bayesian updating for CEU preferences. Here we apply this property to general vertical invariant preferences to connect unconditional and conditional preferences. We then show that this characterizes a simple updating rule for variational preferences, which nests previous results for Bayesian updating in the MEU and multiplier preferences cases.

The rest of the paper is summarized as follows. Section 2 illustrates the main intuition in a dynamic three-color Ellsberg example. Section 3 introduces notation. Section 4 provides two axiomatic representation models and establishes the key equivalence between Event Complementarity and partial information aversion in both. Section 5 relates our key result to four families of ambiguity preferences. Section 6 extends the model to an information acquisition problem with a menu of actions. Section 7 discusses the main result in the smooth ambiguity model. Proofs not provided in the main text are relegated to the Appendix.

# 2 An Illustrating Example

In this section, we revisit the classic three-color Ellsberg Paradox (Ellsberg, 1961), to illustrate the main connection between Event Complementarity and Aversion to Partial Information. Ellsberg considered an urn with 90 balls. Among them 30 balls are red, and 60 balls are either green or yellow, with the exact proportion unknown. The decision maker (DM) places bets on the color of a ball drawn from the urn. In this section, a bet that pays \$1 on event $E \subseteq \{R, G, Y\}$ and \$0 otherwise is denoted $f_E$. Let $\succcurlyeq_E$ denote the updated preferences conditional on learning event $E$; particularly, $\succcurlyeq_S$ denotes the atemporal/unconditional preferences.

In an atemporal setting, a typical Ellsbergian DM strictly prefers betting on red to betting on green, but strictly prefers betting on the event that the ball is either green or yellow ($\{G, Y\}$), to betting on the event that the ball is either red or yellow ($\{R, Y\}$), i.e.,

$$f_R \succ_S f_G \quad \text{and} \quad f_{RY} \prec_S f_{GY}$$

[15] Alternatively, Hanany and Klibanoff (2007, 2009) relax consequentialism, and characterize dynamically consistent updating rules for ambiguity preferences. They use a weaker notion of dynamic consistency than ours.

Thus by adding the same event yellow to both red and green, the DM reverses her preference ranking between them. Obviously, this reflects a preference for betting on events with known probabilities ($p(\{R\}) = \frac{1}{3}$ and $p(\{G,Y\}) = \frac{2}{3}$) to betting on events with unknown probabilities ($p(\{G\}) \in \{0, \frac{1}{90}, \ldots, \frac{60}{90}\}$ and $p(\{R,Y\}) \in \{\frac{30}{90}, \frac{31}{90}, \ldots, 1\}$), which is the standard interpretation of ambiguity aversion.

To crisply illustrate the key concept of Event Complementarity, we assume the DM (i) is risk-neutral; (ii) has MEU preferences (Gilboa and Schmeidler, 1989); and (iii) conditional on an event $E$ her preferences are updated prior-by-prior (Jaffray, 1994; Pires, 2002). Our main result does not rely on any of these assumptions. Particularly, the DM has vNM utility index over money $u(x) = x$; prior set $\mathcal{P} = \{(p_R, p_G, p_Y) \in \mathbb{R}^3_+ : p_R = \frac{1}{3}, p_R + p_G + p_Y = 1\}$; and conditional on an event $F \subseteq \{R, G, Y\}$, her preferences $\succsim_F$ are represented by $V_F(f) = \min_{p \in \mathcal{P}} E_p[f|F]$. Simple calculations show that

$$\begin{aligned} V_S(f_{GY}) &> V_S(f_G) + V_S(f_Y), \\ V_S(f_{RY}) &= V_S(f_R) + V_S(f_Y) \end{aligned}$$

The first inequality says the DM would rather bet on the joint event $\{G, Y\}$ than having two separate bets on the singleton event $\{G\}$ and on event $\{Y\}$. The intuition is as follows: the singleton events $\{G\}$ and $\{Y\}$ have unknown probabilities (in $\{0, \frac{1}{90}, \ldots, \frac{60}{90}\}$) and an Ellsberg-type DM dislikes betting on them; when assessed jointly, state green and state yellow complement each other and betting on the joint event $\{G, Y\}$, which has a precise probability $\frac{2}{3}$, becomes more attractive. The second equality says the DM is indifferent between betting on the joint event $\{R, Y\}$ and betting separately on event $\{R\}$ and on event $\{Y\}$. This is because there is no complementarity between event $\{R\}$ and event $\{Y\}$. The point is an MEU DM will never strictly prefer separate bets on two disjoint events to a single bet on their union.

We also need to check the informational preferences of an Ellsergian DM, suppose her dynamic preferences conforms to our model. We compare two interim information structures: (i) null information partition ($\pi_0 = \{\{R, G, Y\}\}$); (ii) partial information that reveals whether the ball is yellow or not ($\pi = \{\{R, G\}, \{Y\}\}$). Under prior-by-prior updating, it is straightforward to see that $V_{RG}(f_{GY}) = 0$ and $V_{RG}(f_{RY}) = \frac{1}{3}$; obviously $V_Y(f_{GY}) = V_Y(f_{RY}) = 1$. By the folding back procedure (1),

$$\begin{aligned} V(\pi, f_{GY}) &< V(\pi_0, f_{GY}), \\ V(\pi, f_{RY}) &= V(\pi_0, f_{RY}). \end{aligned}$$

Thus the DM also weakly prefer null information $\pi_0$ to partial information $\pi$, displaying partial information aversion.

Furthermore, this example suggests that the act at which the DM displays a preference for event complementarity is also the act at which she has partial information aversion, thus these two preference patterns might be closely related in an ambiguity averse DM. Section 3 provides a general characterization of this connection.

# 3 The Model

## 3.1 Preliminaries

We consider an environment with subjective uncertainty, where $S$ is a *finite* set of possible states of the world. Let $s$ denote a generic element of $S$. An event $E$ is a subset of the state space $S$ and $E^c$ denotes its complement. $\Sigma$ is the collection of all non-empty subsets of $S$. For all nonempty $E \subseteq S$, let $\Delta(E)$ denote the set of probabilities on $E$.

Set $X$ describes all consequences. We assume $X$ is a connected and convex subset of a general topological vector space. An act $f : S \to X$ is a mapping that assigns every state $s$ an outcome $x$.[16] Let $\mathcal{F}$ be the set of all such acts.[17] An act $f$ is *constant* if it maps every state to the same consequence $x$; in this case $f$ is identified with $x$. For $f, g \in \mathcal{F}$ and $E \in \Sigma$, let $fEg$ denote the composed act such that $(fEg)(s) = f(s)$ if $s \in E$ and $(fEg)(s) = g(s)$ if $s \notin E$. For $f, g \in \mathcal{F}$ and $\alpha \in (0, 1)$, $\alpha f + (1 - \alpha)g$ denotes the pointwise mixture of $f$ and $g$: $(\alpha f + (1 - \alpha)g)(s) = \alpha f(s) + (1 - \alpha)g(s)$ for all $s$.

The DM faces two-stage resolution of uncertainty, described by the following simple information structures. In stage 0, the DM has no ex-ante information about the true state. In intermediate stage 1, she receives some partial information that the true state lies in some event $E_i \subseteq S$. Finally, the true state $s \in E_i$ is revealed in stage 2. In particular, the intermediate stage partial information, denoted $\pi = \{E_1, \ldots, E_n\}$, is modeled as a partition of the state space $S$.[18] Let $\Pi$ be the set of all such partitions. In particular, $\pi_0 = \{S\}$ denotes the coarsest partition, corresponding to the case when no information is learned in stage 1, and $\pi^* = \{\{s_1\}, \ldots, \{s_{|S|}\}\}$ denotes the finest partition, corresponding to the case when all

[16] The act here is a slight generalization of the classic Anscombe and Aumann (1963) act, which can be viewed as a special case when $X$ is a set of lotteries over some riskless consequence set $Z$. All results in this paper will hold in the Anscombe-Aumann domain. We choose a general convex set $X$ for a conceptual advantage— there is no explicit discussion about the second-stage objective randomization.

[17] We endow $\mathcal{F}$ with the product topology.

[18] A partition of the set $S$ is a collection of disjoint subsets whose union is $S$.

relevant uncertainties are resolved in stage 1. For all $\pi$, let $\mathcal{F}_\pi$ be the subset of $\pi$-measurable acts in $\mathcal{F}$.[19]

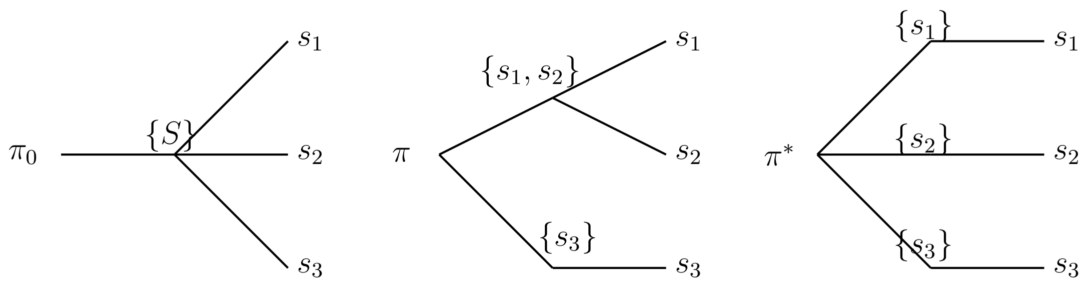


Figure 1: Event trees.

We consider as primitive the DM's ex-ante preferences in stage 0 over the product space of anticipated information $\Pi$ and acts $\mathcal{F}$. Denote the preferences and this enriched domain by $\succcurlyeq$ and $\Pi \times \mathcal{F}$.[20] If $(\pi, f) \succcurlyeq (\pi', g)$, then the DM prefers act $f$ with anticipated information $\pi$ to act $g$ with anticipated information $\pi'$. In this enriched domain, preferences can be "indexed" in two ways: If $\pi = \pi'$, then $\succcurlyeq$ describes the DM's ex-ante preferences over acts for fixed interim information $\pi$; If $f = g$, then $\succcurlyeq$ describes the DM's ex-ante preferences over anticipated information at a given act $f$.[21] In particular, since the DM's preferences over acts depend on how information is revealed over time, our model allows for non-reduction.

# 4 Axiomatic Characterization

## 4.1 Recursive Model

We impose axioms on ex-ante preferences $\succcurlyeq$ on $\Pi \times \mathcal{F}$. For notational simplicity, we sometimes use $\succcurlyeq_\pi$ to denote the restriction of $\succcurlyeq$ on $\{\pi\} \times \mathcal{F}$.

[19] An act is $\pi$-measurable if it is constant on every event in the partition $\pi$.

[20] Space $\Pi$ is endowed with the discrete topology, and space $\Pi \times \mathcal{F}$ is endowed with the product topology.

[21] A similar kind of double-indexing was used in Gajdos et al. (2008).

**Axiom 1** (Weak Order)**.** For all $\pi, \pi', \pi'' \in \Pi$ and $f, g, h \in \mathcal{F}$,

1. (Completeness) $(\pi, f) \succcurlyeq (\pi', g)$ or $(\pi, f) \preccurlyeq (\pi', g)$;
2. (Transitivity) If $(\pi, f) \succcurlyeq (\pi', g)$ and $(\pi', g) \succcurlyeq (\pi'', h)$, then $(\pi, f) \succcurlyeq (\pi'', h)$.

**Axiom 2.** For all $\pi \in \Pi$, $\succcurlyeq_\pi$ on $\{\pi\} \times \mathcal{F}$ satisfies

1. (Continuity) For all $f \in \mathcal{F}$, sets $\{(\pi, g) \in \{\pi\} \times \mathcal{F} : (\pi, g) \succcurlyeq (\pi, f)\}$ and $\{(\pi, g) \in \{\pi\} \times \mathcal{F} : (\pi, f) \succcurlyeq (\pi, g)\}$ are closed.
2. (Strong Monotonicity) If $(\pi, f(s)) \succcurlyeq (\pi, g(s))$ for all $s$, then $(\pi, f) \succcurlyeq (\pi, g)$.[22] If in addition one of the preference rankings is strict, then $(\pi, f) \succ (\pi, g)$.
3. (Risk Independence) For all $\pi \in \Pi$, $x, y, z \in X$ and all $\alpha \in (0, 1)$,

$$(\pi, x) \succcurlyeq (\pi, y) \Leftrightarrow (\pi, \alpha x + (1 - \alpha)z) \succcurlyeq (\pi, \alpha y + (1 - \alpha)z)$$

4. (Non-degeneracy) $(\pi, f) \succ (\pi, g)$ for some $f, g \in \mathcal{F}$.

These axioms are standard. Weak Order is a basic assumption of rationality. Continuity is a common technical assumption needed for the existence of a real-valued utility representation. Strong Monotonicity says both more is better and improvements in every state "count", where the latter restriction is added to simplify our discussions on conditional preferences elicitation. Risk Independence imposes Independence axiom on the constant acts. Non-degeneracy rules out the trivial case that all acts are in the same indifference class.

**Axiom 3** (Stable Risk Preferences)**.** For all $\pi, \pi' \in \Pi$ and $x, y \in X$:

1. $(\pi, x) \succcurlyeq (\pi, y)$ if and only if $(\pi', x) \succcurlyeq (\pi', y)$.
2. $(\pi, x) \sim (\pi', x)$.

Intuitively, Stable Risk Preferences says the DM's preferences over constant acts are not affected by the anticipated information partitions.[23]

The next axiom is substantial: It requires that the ex-ante preferences anticipating information $\pi$, $\succcurlyeq_\pi$, are separable with respect to every event in the partition $\pi$.

[22] Here $f(s)$ and $g(s)$ are constant acts that give in every state outcomes $f(s)$ and $g(s)$, respectively.

[23] Moreover, as shown in Lemma 4 in Appendix B with Stable Risk Preferences, Continuity of every $\{\succcurlyeq_\pi\}$ implies continuity of $\succcurlyeq$ on $\Pi \times \mathcal{F}$.

**Axiom 4** (Anticipated Partition Separability)**.** For all $\pi, \pi' \in \Pi$ and event $E \in \pi \cap \pi'$, and all $f, g, h, h' \in \mathcal{F}$,

$$(\pi, fEh) \succcurlyeq (\pi, gEh) \Leftrightarrow (\pi', fEh') \succcurlyeq (\pi', gEh')$$

Specifically, if the DM anticipates an information partition $\pi$ containing the event $E$ and compares two acts $fEh$ and $gEh$ that agree with each other on $E^c$, then two types of modifications will not affect relative rankings: (i) the common component of $fEh$ and $gEh$ on $E^c$, $h$, is substituted by another common component $h'$; (ii) the anticipated partition $\pi$ is modified to $\pi'$, yet changes only occur to subevents of $E^c$. Note that Anticipated Partition Separability resembles Savage's (1954) Sure-Thing-Principle. The key difference is Anticipated Partition Separability only requires the ex-ante preferences $\succcurlyeq_\pi$ to be separable with respect to events in $\pi$; while the Sure-Thing-Principle looks at static preferences over acts that must be separable with respect to all events in $\Sigma$.

With Anticipated Partition Separability, we can elicit from the ex-ante preferences $\succcurlyeq$ the stage-1 conditional preferences upon learning event $E$, which we denote by $\succcurlyeq_E$. The definition below follows Savage (1954).[24]

**Definition 1** (Conditional Preferences Elicitation)**.** For all $E \in \Sigma$ and $f, g \in \mathcal{F}$, $f \succcurlyeq_E g$ if $(\pi, fEh) \succcurlyeq (\pi, gEh)$ for some $h \in \mathcal{F}$ and some $\pi$ such that $E \in \pi$.

*Remark* 1. Anticipated Partition Separability ensures that the elicited conditional preferences $\{\succcurlyeq_E\}_{E\in\Sigma}$ are well defined and satisfy two well-known properties. First, outcomes in states that have been ruled out should not matter for the conditional preferences, i.e., $fEg \sim_E fEh$ for all acts $f, g, h$ and all event $E$.[25] This property is *Consequentialism*(Gumen and Savochkin, 2013; Machina, 1989). Second, fix any anticipated partition $\pi$, if all interim conditional preferences $\{\succcurlyeq_E\}_{E\in\Pi}$ prefer act $f$ to act $g$, then the ex-ante preferences $\succcurlyeq_\pi$ prefer act $f$ to act $g$. That is, $\succcurlyeq_\pi$ and $\{\succcurlyeq_E\}_{E\in\Pi}$ are *Dynamically Consistent* (Epstein and Schneider, 2003; Maccheroni et al., 2006b).[26]

The next two axioms are intuitive and we impose them for simplicity.

**Axiom 5** (Indifference to Redundant Information)**.** For all $\pi$, $f \in \mathcal{F}_\pi$, $(\pi, f) \sim (\pi^*, f)$.

[24]Epstein and LeBreton (1993) and Machina and Schmeidler (1992) also use this definition to derive the conditional preferences.

[25]To see this, let $fEg = f'$ and $fEh = g'$. Then $f'Ef = g'Ef = f$. Therefore $(\pi, f'Ef) \sim (\pi, g'Ef)$ for $\pi = \{E, E^c\}$. By Definition 1, $f' \sim_E g'$.

[26]To see this, let $\pi = \{E_1, \cdots, E_n\}$ and $f \equiv f^0$. Then by Definition 1, $f \succcurlyeq_{E_1} g \Rightarrow f \succcurlyeq_\pi gE_1f$. Let $gE_1f \equiv f^1$. Then by Consequentialism $f \succcurlyeq_{E_2} g \Rightarrow f^1 \succcurlyeq_{E_2} g$, and by Definition 1 $f^1 \succcurlyeq_\pi gE_2f^1$. Let $gE_2f^1 \equiv f^2$. Repeat this for all $E_i \in \pi$ $(i = 1, \cdots, n)$, let $f^i = gE_if^{i-1}$ and we have $f^{i-1} \succcurlyeq_\pi f^i$. By transitivity, $f = f^0 \succcurlyeq_\pi f^n = g$.

Intuitively, if an act $f$ is $\pi$-measurable, then all outcome uncertainties about $f$ are resolved in stage 1 after learning which event in $\pi$ realizes. Thus the additional information in the perfect information partition $\pi^*$ relative to that in $\pi$ should not matter at the $\pi$-measurable act $f$.

The last axiom, Time Neutrality, assumes the DM does not care when uncertainties are revealed, as long as that happens in one shot. It allows us to focus on preferences for the sequences of information revelation.[27]

**Axiom 6** (Time Neutrality)**.** For all $f$, $(\pi^*, f) \sim (\pi_0, f)$.

Finally, Time Neutrality implies $\succsim_{\pi^*}=\succsim_{\pi_0}$. Both preferences are identified as the one-shot unconditional preferences and denoted by $\succsim_0$ in the following text.

#### 4.1.1 Folding-back Procedure

Axiom 1 to 6 characterizes a recursive framework between the ex-ante preferences and the elicited conditional preferences, which we call the *folding-back procedure*. Loosely speaking, the DM evaluates every partition-act pair $(\pi, f)$ by backward induction: she first evaluates for every stage-1 event $E_i \in \pi$ what is the (elicited) conditional certainty equivalent of $f$, which we denote by $x_{f,E_i}$, and then she goes back to stage 0 and evaluates the $\pi$-*conditional certainty equivalent of* $f$, that is,

$$f_\pi = \begin{pmatrix} x_{f,E_1} & \text{if } s \in E_1 \\ \vdots & \vdots \\ x_{f,E_n} & \text{if } s \in E_n \end{pmatrix}. \tag{3}$$

Before providing a formal representation of the folding-back procedure, we first introduce some useful notation. For a real interval $K \subseteq \mathbb{R}$, let $K^{|S|}$ be the set of $|S|$-dimensional $K$-ranged vectors, and we denote its generic elements by $\xi, \phi$. For all $k \in K$, let $\bar{k}$ be the constant vector equals to $k$ in every dimension. Fix any nonempty $E \in \Sigma$. A composed vector $\xi E \phi$ equals $(\xi E \phi)(s) = \xi(s)$ if $s \in E$, and $(\xi E \phi)(s) = \phi(s)$ if $s \notin E$. For every $f$ and $\xi$, let $f_E$ and $\xi_E$ be the restrictions of $f$ and $\xi$ on $E$, respectively. Denote the set of all such restricted acts $f_E$ by $\mathcal{F}_E$. For a functional $I_E : K^{|E|} \mapsto \mathbb{R}$, we say $I_E$ is *monotone* if $\xi_E \geq \phi_E$ implies $I_E(\xi_E) \geq I_E(\phi_E)$, and *strongly monotone* if in addition $\xi_E > \phi_E$ implies $I_E(\xi_E) > I_E(\phi_E)$. We say $I_E$ is *normalized* if $I_E(\bar{k}) = k$ for all $k \in K$. Finally, we say $I_E$ is *vertically invariant* if $I_E(\xi_E + \bar{k}) = I_E(\xi_E) + k$ for all $\xi$ and $k$ such that $\xi_E + \bar{k} \in K^{|E|}$. [28]

[27] In a risk environment, Segal (1990) and Dillenberger (2010) make a similar assumption.

[28] The term "vertically invariant" is synonymous with "translation invariant" used in the literature.

**Theorem 1.** *$\succsim$ satisfies Axiom 1 to 6 if and only if there exists a nonconstant affine function $u : X \mapsto \mathbb{R}$, and a continuous, strongly monotone, and normalized functional $I_E : u(X)^{|E|} \mapsto \mathbb{R}$ for every nonempty $E \in \Sigma$ such that $\succsim$ is represented by a utility function $V : \Pi \times \mathcal{F} \mapsto \mathbb{R}$ such that*

$$V(\pi, f) = I_S \begin{pmatrix} V_{E_1}(f) & \text{if } s \in E_1 \\ \vdots & \vdots \\ V_{E_n}(f) & \text{if } s \in E_n \end{pmatrix}$$

*for $\pi = \{E_1, \ldots, E_n\}$ and*

$$V_E(f) = I_E(u(f_E)) \quad \text{for all } E \in \Sigma.$$

*Moreover, if both $(u, \{I_E\}_{E\in\Sigma})$ and $(u', \{I'_E\}_{E\in\Sigma})$ represent $\succsim$, then there exists $a > 0$ and $b \in \mathbb{R}$ such that $u' = au + b$ and $I'_E(\xi_E) = aI_E(\frac{\xi_E - \bar{b}}{a}) + b$ for all $\xi_E \in (u'(X))^{|E|}$.*

### 4.1.2 Event Complementarity

Next we introduce the novel behavioral concept that we call Event Complementarity. Recall the Ellsberg example introduced in Section 2. A typical Ellsbergian DM prefers betting on red over betting on green ($f_R \succ_0 f_Y$), and prefers betting on green or yellow over betting on red or yellow ($f_{GY} \succ_0 f_{RY}$). The classic theory explaining the Ellsbergian behaviors is a preference for hedging (Schmeidler, 1989): for all $f, g \in \mathcal{F}$ and all $\alpha \in (0, 1)$, $f \sim_0 g$ implies $\alpha f + (1 - \alpha)g \succsim_0 f \sim_0 g$. In other words, when facing subjective uncertainties, the DM can always weakly benefit from randomization. For example, if the DM in the Ellsberg example can also toss a fair coin, then the mixed act of choosing (the dispreferred) bet $f_Y$ if the coin lands head and choosing (the dispreferred) bet $f_{RY}$ if the coin lands tail gives a winning probability $\frac{1}{2}$ no matter which color is the drawn ball, which gives exactly the same state-by-state outcome as an analogous mixing of the two preferred bets, $f_R$ and $f_{GY}$.[29] Thus randomization (or mixing) helps the DM hedge against ambiguity. This characterization of ambiguity aversion builds the foundation of most existing models of ambiguity averse preferences that we will discuss in Section 5.

Our main concept, Event Complementarity, is directly motivated by the idea that DM dislikes betting on an event with unknown probability relative to a comparable event with known probability. Particularly, for any event $F$ (event $\{G, Y\}$ in the Ellsberg example), any partition $\{E, E^c\}$ (partition $\{\{R, G\}, \{Y\}\}$ in the Ellsberg example) can divide it into two sub-events: $F_1 = F \cap E$ and $F_2 = F \cap E^c$ (events $\{G\}$ and $\{Y\}$ in the Ellsberg example). The

[29] This argument modifies Raiffa 's (1961) critique to the three-color Ellsberg urn scenario.

DM weakly prefers having a single bet on the event $F$ rather than having two separate bets on event $F_1$ and $F_2$. The intuition is that while event $F$ can be unambiguous, its separate components $F_1$ and $F_2$ are subjected to ambiguity.[30] Hence the DM weakly prefers the joint bet on event $F$.

Below we formalizes this "preference to bet on a whole event rather than its separate components" interpretation as an axiom on preferences.

**Axiom 7** (Event Complementarity)**.** For all $E \in \Sigma$, $f \in \mathcal{F}$, and $x \in X$, if $f \sim_E x$, then $f \succcurlyeq_0 xEf$.

Intuitively, the axiom describes an "ironing" procedure: Fix an event $E$, if an act $f$ yields uncertain payoffs on $E$, then we "iron out" this variation by replacing the restriction of $f$ on $E$ by its $E$-conditional certainty equivalent, that is, the constant act $x$ such that $f \sim_E x$. In this way, any informational complementarity between states in events $E$ and its complement $E^c$ that is embedded in $f$ is eliminated; yet the ironed-out act, denoted $xEf$, is equivalent to $f$ conditional on event $E$ and is equal to $f$ on event $E^c$. From $f$ to $xEf$, nothing has changed except for the elimination of informational complementarity. Thus the DM's attitude towards this complementarity should be revealed by her unconditional preferences $\succcurlyeq_0$ between $f$ and $xEf$. In particular, a strict preference $f \succ_0 xEf$ reveals a DM who strictly values such complementarity.

### 4.1.3 Aversion to Partial Information

The second new concept is aversion (attraction) to partial information. Recall that in the motivating example, Bob, who would rather wait for his final job outcome to be revealed in the end than get rumour information in the interim stage, is averse to partial information. This behavioral pattern is formalized as follows.

**Definition 2.** We say $\succcurlyeq$ exhibits *aversion to partial information at act $f$* if $(\pi_0, f) \succcurlyeq (\pi, f)$ for all $\pi$. We say $\succcurlyeq$ exhibits *aversion to partial information* if $\succcurlyeq$ exhibits aversion to partial information at all acts.

We can similarly define attraction to partial information when $(\pi_0, f) \preccurlyeq (\pi, f)$ for all $\pi$ and $f$. If $\succcurlyeq$ displays both aversion and attraction to partial information, then we can say $\succcurlyeq$ is neutral to partial information.

[30]This is motivated by the idea that the intersection of two unambiguous event might be ambiguous (Zhang, 2002).

Our concept of aversion to partial information is reminiscent of the Preferences for One-Shot Resolution of Uncertainty studied by Dillenberger (2010), who considers a DM disprefers a two-stage compound lottery relative to its reduced one-shot lottery. The main difference is we consider a subjective framework and the DM might not be probabilistic sophisticated. In a different set-up, Koszegi and Rabin (2009) show that loss aversion over variations in beliefs imply a similar preference for getting information "clumped together rather than apart".[31]

Note that in our definition the DM only compares the null information $\pi_0$ with some partial information $\pi \neq \pi_0$. We remain ignorant about how the DM compares two partial information partitions $\pi$ and $\pi'$, even if partition $\pi'$ is finer than $\pi$. In this sense, we consider a weaker revealed information ranking than those studied by Grant et al. (1998) and Skiadas (1998), which require that coarser information is always preferred to finer information.[32]

#### 4.1.4 Main Theorem

The main theorem of this paper shows that in the folding-back model, aversion to partial information is equivalent to Event Complementarity.

**Theorem 2.** *Suppose $\succcurlyeq$ satisfies Axiom 1–6. Then the following statements are equivalent:*

1. *$\succcurlyeq$ satisfies Event Complementarity.*
2. *$\succcurlyeq$ exhibits aversion to partial information.*

The proof is elementary and given below.[33]

*Proof.* Suppose $\succcurlyeq_0$ satisfies Event Complementarity. Fix a finite partition $\pi = \{E_1, \cdots, E_n\}$, and an act $f$. For each $i = 1, \cdots, n$, let $x_i \in X$ be the $E_i$-conditional certainty equivalent of $f$, i.e., $x_i \sim_{E_i} f$. Let $f_0 := f$, $f_1 = x_1 E_1 f_0$, $f_2 = x_2 E_2 f_1$, $\cdots$, $f_n = x_n E_n f_{n-1} = (x_1 E_1 x_2 E_2 \cdots x_{n-1} E_{n-1} x_n)$. Note that $f_n$ is $\pi$-measurable. By definition 1, $x_i \sim_{E_i} f_{i-1}$ for all $i = 1, \cdots, n$. Repeatedly applying Event Complementarity yields $(\pi_0, f_0) \succcurlyeq (\pi_0, f_1) \succcurlyeq \ldots \succcurlyeq (\pi_0, f_n)$. Furthermore, Anticipated Partition Separability implies $(\pi, f_0) \sim (\pi, f_1) \sim \cdots \sim (\pi, f_n)$. Putting these results together yields

$$(\pi, f) \sim (\pi, f_n) \sim (\pi^*, f_n) \sim (\pi_0, f_n) \preccurlyeq (\pi_0, f),$$

[31] In Koszegi and Rabin (2009), an information structure is a sequence of discrete-value signals about a binary consumption outcome. Their setting is also purely objective.

[32] In Grant et al. (1998), finer information corresponds to higher Blackwell's informativeness ranking.

[33] The construction of the proof is similar to that of Theorem 1 in Dillenberger (2010).

where the second indifference follows from Indifference to Redundant Information and the third indifference is due to Time Neutrality. Since this is true for an arbitrary act $f$ and partition $\pi$, $\succcurlyeq$ exhibits aversion to partial information.

We prove the converse by contradiction. Suppose not, so $\succcurlyeq$ exhibits aversion to partial information but there exists some $\pi$, $E \in \pi$, $f$, and $x$ such that $f \sim_E x$, but $(\pi_0, xEf) \succ (\pi_0, f)$. Let $n_1, \cdots, n_m$ be labels for states in $E^c$, i.e., $E^c = \{s_{n_1}, \cdots, s_{n_m}\}$. Then consider the finer partition $\pi' = \{E, \{s_{n_1}\}, \cdots, \{s_{n_m}\}\}$. Thus $xEf$ is $\pi'$-measurable, and by Indifference to Redundant Information and Time Neutrality, $(\pi', xEf) \sim (\pi^*, xEf) \sim (\pi_0, xEf)$. By Anticipated Partition Separability, $(\pi', f) \sim (\pi', xEf)$. By transitivity, $(\pi', f) \sim (\pi_0, xEf) \succ (\pi_0, f)$. This violates partial information aversion, a contradiction. □

*Remark* 2. If define Event Substitution as $f \sim_E x \Rightarrow f \preccurlyeq_0 xEf$, then a parallel equivalence between attraction to partial information and Event Substitution also holds.[34] In our model, the patterns of revealed preferences for partial information are tightly linked to systematic deviations from expected utility.

*Remark* 3. So far we have not imposed any restriction among conditional preferences $\{\succcurlyeq_E\}_{E\in\Sigma}$. Under Axiom 1–6, any continuous, consequentialist, and strongly monotone preferences over acts $\mathcal{F}$ can be a candidate for $\succcurlyeq_E$. This means even if events $E \subsetneq F$, $\succcurlyeq_E$ and $\succcurlyeq_F$ can rank most nonconstant acts differently. One interpretation is that the general recursive model has too many degrees of freedom. This is subjected to the criticism that conditional preferences should reflect beliefs updated from the same underlying belief about the state space $S$. Hence some belief consistency requirement connecting $\{\succcurlyeq_E\}_{E\in\Sigma}$ (and thus $\{\succcurlyeq_\pi\}_{\pi\in\Pi}$) can be desirable. We will address this in Section 4.2.

*Remark* 4. Our general model (Theorem 1) is reminiscent of the two-stage compound lottery model of Segal (1990), and our main result bears similarity to Dillenberger's (2010) equivalence result between preferences for one-shot resolution of uncertainty and Allais-type behaviors in the Segal domain. In fact, if $\{\succcurlyeq_E\}_{E\in\Sigma}$ are also probabilistically sophisticated (Machina and Schmeidler, 1995) with conditional probabilities $\{p_E \in \Delta(E)\}_{E\in\Sigma}$, our general representation model can be viewed as a special case of Segal (1990). An informal argument is as follows. Segal's domain corresponds to the set $\Delta(\Delta X)$. In our model, fix an arbitrary pair of partition and act $(\pi, f)$. Due to probabilistic sophistication, at every event $E \in \pi$, the pair induces a conditional probability distribution: $\hat{p}_E(B) := p_E(f^{-1}(B) \cap E)$ for all measurable set $B$ in the outcome space $X$. The pair also induces a probability $\hat{P}$ over these conditional lotteries, given by $\hat{P}(\hat{p}_E) := p_S(E)$, where $p_S$ is the unconditional probability over the state space $S$. Thus each pair $(\pi, f)$ induces a compound lottery $\hat{P} \in \Delta(\Delta X)$, and

[34]Proof is analogous to that of Theorem 2 and thus omitted.

our domain can be mapped to Segal's domain. Yet these two domains are not equivalent under probabilistic sophistication, because the mapping we described is not onto. An obvious counterexample is when the support of $P$ contains more distinct lotteries than the number of states $|S|$.[35]

## 4.2 Cross-partition Recursive Model

To address the concern raised in Remark 3, in this part we impose axioms to ensure that the elicited conditional preference $\{\succsim_E\}_{E\in\Sigma}$ reflect some common consistent beliefs. We then derive a more restrictive recursive utility representation (Definition 3) where the conditional utilities are determined via a fixed point formula from the one-shot unconditional utilities. The key equivalence relation between EC and API is still valid in this new representation.

### 4.2.1 Belief Consistency and Updating

For most of our results, we impose on $\succsim_0$ Maccheroni et al.'s (2006a) Weak Certainty Independence. It generalizes standard Independence axiom in the EU model, Certainty Independence in the MEU model (Gilboa and Schmeidler, 1989), and Comonotomic Independence in the CEU model (Schmeidler, 1989).

**Axiom 8** (Weak Certainty Independence)**.** For all $f, g \in \mathcal{F}$, $x, y \in X$, and $\alpha \in (0, 1)$,

$$\alpha f + (1-\alpha)x \succsim_0 \alpha g + (1-\alpha)x \Rightarrow \alpha f + (1-\alpha)y \succsim_0 \alpha g + (1-\alpha)y$$

In words, Weak Certainty Independence says that if mixing an act $f$ with a constant act $x$ at a weight $\alpha$ is preferred to mixing another act $g$ with the same constant act $x$ at the same weight $\alpha$, then replacing the common constant act $x$ by another constant act $y$ and maintaining the same weight $\alpha$ should not change the ranking. Geometrically, Weak Certainty Independence implies that the indifference curves in the utility profile space are parallel around the certainty line. We refer interested readers to Maccheroni et al. (2006a) for further interpretation of Weak Certainty Independence. In this paper, we impose Weak Certainty Independence because, together with other standard axioms on $\succsim_0$, it ensures that the representation $(u, I_0)$ satisfy *vertical invariance*, i.e., $I_0(\xi + \bar{k}) = I_0(\xi) + k$ for all $\xi, \xi + \bar{k} \in u(X)^{|S|}$. Particularly, as we will discuss below, a vertically invariant aggregator $I_0$ is

[35]The connection between an infinite-state-space extension of our model and Segal's model remains an open question.

also Lipschitz continuous of rank 1, which is a condition we need to ensure the identification of our updating rule.[36]

The next axiom impose belief consistency across anticipated partitions.

**Axiom 9** (Belief Consistency)**.** For all $f$, $x$, and $E$, let $\pi_E = \{E, E^c\}$, then

$$(\pi_E, fEx) \sim (\pi_E, x) \Rightarrow (\pi_0, fEx) \sim (\pi_0, x) \tag{4}$$

Intuitively, Belief Consistency requires the following. Suppose the DM expects to learn about $\pi_E = \{E, E^c\}$ in stage 1 and is ex-ante indifferent between the act of getting $f$ on $E$ and $x$ on $E^c$ and the constant act $x$, in which case $f$ must be equivalent to $x$ conditional on $E$. Then the act $fEx$ has the same value as $x$ conditional on $E$ and conditional on $E^c$. In this case, knowing whether $E$ or $E^c$ happens in stage 1 is not informative about the conditional value of final outcomes. Therefore a DM with act $fEx$ will be indifferent between the indiscriminating partition $\pi_E$ and the trivial partition $\pi_0$, that is, $(\pi_0, fEx) \sim (\pi_E, fEx)$.

As shown in the proof of Theorem 3, the converse of (4) is implied by other axioms. Particularly, this specifies an updating rule from $\succcurlyeq_0$ to $\succcurlyeq_E$:

$$fEx \sim_0 x \Rightarrow f \sim_E x. \tag{5}$$

That is, for all act $f \in \mathcal{F}$, the outcome $x$ that solves the fixed point equation $fEx \sim_0 x$ is the $E$-conditional certainty equivalent of $f$. Here the certainty equivalent $x$ "calibrates" the utility the DM receives from act $f$ conditional on $E$. We say there is an essentially unique solution to the fixed point equation $fEx \sim_0 x$ if

$$fEx \sim_0 x \text{ and } fEy \sim_0 y \Rightarrow x \sim_0 y. \tag{6}$$

The counterexample in Appendix B.2 demonstrates that if $\succcurlyeq_0$ only satisfies continuity and strong monotonicity, then the solution to relation $fEx \sim_0 x$ might not be essentially unique, leading to the troubling case of multiple conditional certainty equivalents. Lemma 5 shows that if $\succcurlyeq_0$ is also Lipschitz continuous of rank one, then the solution to $fEx \sim_0 x$ is essentially unique. Since Weak Certainty Independence is stronger than Lipschitz continuous of rank 1, Axiom 9 is sufficient for a well-defined updating rule in our main model. Yet the question of what is the largest preference family in which such updating rule is well-defined remains open.

A frequently used axiom for updating ambiguous preferences is called Conditional Certainty Equivalent Consistency.

[36]See Appendix B.2 for detailed discussion.

**Axiom** (Conditional Certainty Equivalent Consistency, CCEC)**.** *For all $f \in \mathcal{F}$, $x \in X$, and $E \in \Sigma$,*

$$f \sim_E x \Leftrightarrow fEx \sim_0 x$$

It is first introduced by Pires (2002) as a coherence condition between unconditional and conditional preferences for MEU preferences, and later called Conditional Certainty Equivalent Consistency by Eichberger et al. (2007) as an updating condition for CEU preferences. In our model, Belief Consistency is equivalent to CCEC.

**Corollary 1.** *Suppose $\succcurlyeq_0$ satisfies Weak Order, Continuity, Strong Monotonicity, and Weak Certainty Independence and $\succcurlyeq$ satisfies Weak Order, Stable Risk Preferences and Partition Separability. Then $\succcurlyeq$ satisfies Belief Consistency if and only if conditional preferences $\{\succcurlyeq_E\}_{E \in \Sigma \setminus \emptyset}$ satisfy CCEC.*

Here we extend prior axiomatic models on updating ambiguous preferences in two ways. First, CCEC has been mostly viewed as mathematical generalization of the Bayesian updating rule. Its fixed point structure typically provides clean formulas for conditional utility representations.[37] For example, applying CCEC to the MEU preferences leads to a conditional utility representation with a prior-by-prior updated posterior set. In our model, CCEC can be equivalently reformulated as a Belief Consistency axiom, which also has a natural behavioral interpretation of an indifference to indiscriminating interim information. Second, we explore CCEC's full potential as a useful updating axiom for ambiguity sensitive preferences. Particularly, all preferences satisfying Weak Certainty Independence, including MEU and CEU preferences, are compatible with CCEC. In Section 5, we apply CCEC (and equivalently, Belief Consistency) to variational preferences, and derive a new updating formula there (Theorem 4).

### 4.2.2 Cross-partition Recursive Representation

Now we are ready to introduce the second representation result.

**Definition 3.** We say preferences $\succcurlyeq$ on $\Pi \times \mathcal{F}$ *admit a cross-partition recursive representation* $(u, I_0)$ if

[37] See Eichberger et al. (2007) for a detailed comparison between CCEC and the Bayesian updating axiom used in Savage (1954).

1. There exists a nonconstant affine function $u : X \mapsto \mathbb{R}$, and a continuous, strongly monotone, normalized, and vertical invariant $I_0 : u(X)^S \mapsto \mathbb{R}$ such that for all $f, g \in \mathcal{F}$,

$$f \succsim_0 g \Leftrightarrow I_0(u \circ f) \geq I_0(u \circ g) \tag{7}$$

2. For all $E \in \Sigma$, $\succsim_E$ elicited according to (1) is represented by $V_E : \mathcal{F} \mapsto \mathbb{R}$, where $V_E(f) := k$ is the unique solution to

$$k = I_0((u \circ f)E\bar{k}) \tag{8}$$

3. $\succsim$ is represented by a continuous function $V : \Pi \times \mathcal{F} \mapsto \mathbb{R}$, where

$$V(\pi, f) = I_0 \begin{pmatrix} V_{E_1}(f) & \text{if } s \in E_1 \\ \vdots & \vdots \\ V_{E_n}(f) & \text{if } s \in E_n \end{pmatrix} \tag{9}$$

In this case, we also say $\succsim$ *is recursively generated by* $\succsim_0$.

This representation is more restrictive than that in Theorem 1. In addition to the recursive utility structure in Theorem 1, the updating rule (8) here specifies how the conditional preferences $\{\succsim_E\}_{E\in\Sigma}$ are derived from unconditional one-shot preferences $\succsim_0$. As discussed before, this updating rule is consistent with many families of ambiguity sensitive preferences. Importantly, the representation features consistent ex-ante beliefs across partitions: for two different information partitions $\pi$ and $\pi'$, $\succsim_\pi$ and $\succsim_{\pi'}$ are related by the same unconditional $\succsim_0$ and thus reflect the same perceived likelihoods of events in $S$. In fact, knowing $\succsim_0$ is sufficient for inferring all $\{\succsim_\pi\}_{\pi\in\Pi}$. Thus if one observes any difference between $\succsim_\pi$ and $\succsim_{\pi'}$, it must be due to differences in the "framing" effects of the anticipated information partitions $\pi$ and $\pi'$, instead of differences in their embedded beliefs about events in $S$.

**Theorem 3.** *The following statements are equivalent:*

1. $\succsim$ *satisfies Axiom 1 to 6 and Axiom 8 to 9.*

2. $\succsim$ *has a cross-partition recursive representation with* $(u, I_0)$.

*Moreover, if both* $(u, I_0)$ *and* $(u', I_0')$ *represent* $\succsim_0$*, then there exists* $a > 0$ *and* $b \in \mathbb{R}$ *such that* $u' = au + b$ *and* $I_0'(\xi) = aI_0(\frac{\xi}{a})$ *for all* $\xi \in (u'(X))^{|S|}$.

### 4.2.3 Static Event Complementarity

In our model the updating axiom implies $f \sim_E x \Leftrightarrow fEx \sim_0 x$. Hence we could rewrite Event Complementarity as an axiom on the one-shot preference $\succcurlyeq_0$.

**Axiom 10** ($\succcurlyeq_0$-Event Complementarity)**.** For all $E$ and $f$, if $fEx \sim_0 x$ for some $x$, then $f \succcurlyeq_0 xEf$.

Here we only rewrite Event Complementarity. Yet we think this could be meaningful in two ways. First, it is much easier to design an experiment and test $\succcurlyeq_0$-Event Complementarity, since one only need observations about $\succcurlyeq_0$. Second, $\succcurlyeq_0$-Event Complementarity is directly comparable to other well-known axioms. For instance, it is a systematic deviation from Savage's Sure-Thing-Principle[38], which implies $f \sim_E x \Rightarrow f \sim_0 xEf$. Hence it clarifies the connection between our model and the classic SEU model. Moreover, rewriting Event Complementarity makes it easier to compare it with Schmeidler's (1989) Uncertainty Aversion axiom. We will fully explore this comparison in the Section 5. Before that we provide a useful sufficient condition for $\succcurlyeq_0$-Event Complementarity.

**Proposition 1.** *Suppose $\succcurlyeq_0$ is represented by $(u, I_0)$ where $I_0$ is vertically invariant. Then $\succcurlyeq_0$ satisfies Event Complementarity if and only if*

$$I_0(u \circ f) \geq I_0(u \circ (fEx)) + I_0(0E(u \circ f - u \circ x)) \tag{10}$$

*for all $f$ and $x$ such that $fEx \sim_0 x$. In particular, if $I_0$ is superadditive, then $\succcurlyeq_0$ satisfies Event Complementarity.*

Inequality (10) describes Event Complementarity of $\succcurlyeq_0$ in terms of the functional form property of its representation $(u, I_0)$. This gives us another way to understand this axiom. Given an act $f$ and a constant act $x$ such that $fEx \sim_0 x$, notice that the utility profile $u \circ f$ corresponding to $f$ can be decomposed into

$$u \circ f = u \circ (fEx) + 0E(u \circ f - u \circ x)$$

By construction $u \circ (fEx)$ is constant on $E^c$, and $0E(u \circ f - u \circ x)$ is constant on $E$. Thus $u \circ f$ is decomposed into the sum of two utility profiles, one capturing the variation of $u \circ f$ on $E$ and one capturing the variation of $u \circ f$ on $E^c$. Proposition 1 shows that Event Complementarity holds if and only if the value of utility profile $u \circ f$, $I_0(u \circ f)$, is greater than or equal to the sum of the values of these two pieces, $I_0(u \circ fEx) + I_0(0E(u \circ f - u \circ x))$.

[38] The Sure-Thing-Principle says $fEh \succcurlyeq_0 gEh \Rightarrow fEh' \succcurlyeq_0 gEh'$ for all $f, g, h, h' \in \mathcal{F}$ and all $E \in \Sigma$.

Notice that if $I_0$ is superadditive, then Event Complementarity always holds. However, the converse is not true generally.

Finally, the following corollary re-establishes our key equivalence result in the cross-partition recursive model.

**Corollary 2.** *Suppose $\succcurlyeq$ is recursively generated by $\succcurlyeq_0$. Then the following statements are equivalent:*

1. *$\succcurlyeq$ satisfies $\succcurlyeq_0$-Event Complementarity.*
2. *$\succcurlyeq$ exhibits aversion to partial information.*

*Proof.* Follows immediately from Theorem 2. □

# 5 Ambiguity Preferences

In this section, we investigate further the link between ambiguity aversion and aversion to partial information. In particular, we examine whether partial information aversion is implied by ambiguity aversion for four well-known classes of vertical invariant ambiguity preferences: MEU, multiplier preferences, variational preferences, and CEU. Another popular class of ambiguity preferences, the second-order belief model, does not satisfy vertical invariance and thus does not belong to our model. We will discuss the second-order belief model separately in Section 7.

Recall Schmeidler's (1989) Uncertainty Aversion axiom, which we call Ambiguity Aversion in this paper:

**Axiom** (Ambiguity Aversion)**.** *For all $f, g \in \mathcal{F}$ and $\alpha \in (0,1)$,*

$$f \sim_0 g \Rightarrow \alpha f + (1-\alpha) g \succcurlyeq_0 f$$

Ambiguity Aversion captures a preference for state-by-state hedging. If $\succcurlyeq_0$ is represented by $(u, I_0)$ and $I_0$ is continuous, monotone, normalized, and vertically invariant, then $\succcurlyeq_0$ is ambiguity averse if and only if $I_0$ is concave.

## 5.1 Maxmin EU

MEU is the most popular model that captures ambiguity aversion. The static MEU model is axiomatized by Gilboa and Schmeidler (1989), and a recursive MEU model is axiomatized by Epstein and Schneider (2003).[39]

We say $\succsim_0$ has an MEU representation $(u, \mathcal{P})$ if it can be represented by a function $V_0 : \mathcal{F} \to \mathbb{R}$ of the form

$$V_0(f) = \min_{p \in \mathcal{P}} \int_S u(f) dp$$

where $\mathcal{P}$ is a closed and convex subset of $\Delta(S)$.

Following Epstein and Schneider (2003), for any convex and closed prior set $\mathcal{P}$ and any partition $\pi$, we define the *$\pi$-rectangular hull of $\mathcal{P}$* to be $rect_\pi(\mathcal{P}) = \{p = \sum_{i=1}^{k} p^i(\cdot|E_i)q(E_i) | \forall p^i, q \in \mathcal{P}\}$. The set $rect_\pi(\mathcal{P})$ is the largest set of probabilities that have the same marginal probabilities and conditional probabilities for events in $\pi$ as elements of $\mathcal{P}$. By definition, $\mathcal{P} \subseteq rect_\pi(\mathcal{P})$ for any $\mathcal{P}$ and $\pi$. The set $\mathcal{P}$ is called *$\pi$-rectangular* if $rect_\pi(\mathcal{P}) = \mathcal{P}$. Whether $\mathcal{P}$ is $\pi$-rectangular is closely related to whether a DM with belief set $\mathcal{P}$ is strictly averse to partial information $\pi$. The next proposition summarizes the link between MEU preferences and aversion to partial information.

**Proposition 2.** *Suppose $\succsim$ is recursively generated by $\succsim_0$. Suppose $\succsim_0$ has a MEU representation $(u, \mathcal{P})$, and $\succsim_E$ has a MEU representation $(u, \mathcal{P}_E)$, for all $E \in \Sigma$. Then*

1. *$\succsim$ exhibits aversion to partial information at all acts.*
2. *For any partition $\pi$, there exists some act $f$ such that $\succsim$ is strictly averse to $\pi$ at $f$, i.e., $(\pi_0, f) \succ (\pi, f)$, if and only if $\mathcal{P}$ is not $\pi$-rectangular.*

*Remark* 5. MEU has an intuitive interpretation as a malevolent Nature playing a zero-sum game against the DM (Maccheroni et al., 2006b). In this interpretation, Nature has a constraint set $\mathcal{P}$, and chooses a probability in order to minimize the DM's expected utility. In our recursive model without reduction, the information $\pi$ turns this into a sequential game. In period 0, Nature chooses a probability from $\mathcal{P}$ for events in $\pi$. In period 1, Nature chooses a (possibly different) probability from $\mathcal{P}$ over states for every event in $\pi$, conditional on that event. In this way, information $\pi$ expands Nature's constraint set from $\mathcal{P}$ to $rect_\pi(\mathcal{P})$. On the other hand, the DM has committed ex-ante to a fixed act $f$. So introducing information $\pi$ helps Nature and hurts the DM. Part (2) of Proposition 2 shows that if information strictly expands Nature's constraint set, that is, if $\mathcal{P} \subsetneq rect_\pi(\mathcal{P})$, then Nature can make the DM strictly worse off at some act.

[39] In contrast with our model, Epstein and Schneider (2003) assume reduction.

*Remark* 6. Epstein and Schneider (2003) develop a recursive MEU model in which they maintain reduction. They show that $\succcurlyeq$ is dynamically consistent with respect to $\pi$ if and only if $\mathcal{P}$ is $\pi$-rectangular. Part (2) of Proposition 2 shows that if we instead maintain dynamic consistency but relax reduction, then information neutrality at $\pi$ is equivalent to $\pi$-rectangularity of $\mathcal{P}$.

*Remark* 7. When the prior set $\mathcal{P}$ is a singleton (so the DM has SEU), or when $\mathcal{P} = \Delta(S)$, the DM is intrinsically information neutral.

## 5.2 Choquet EU

The CEU model is axiomatized by Schmeidler (1989). It is of particular interest because it allows for both ambiguity averse and ambiguity loving preferences, so this provides a framework for studying the relationship between information preferences and ambiguity attitudes more generally.

We say $\succcurlyeq_0$ has a CEU representation $(u, \nu)$ if it can be represented by a function $V_0 : \mathcal{F} \to \mathbb{R}$ of the form

$$V_0(f) = \int u(f) d\nu$$

where $\nu : \Sigma \to [0, 1]$ is a capacity, that is, $\nu(S) = 1$, $\nu(\emptyset) = 0$, and $\nu(E) \leq \nu(F)$ for all $E \subseteq F$.

If $\succcurlyeq_0$ satisfies Ambiguity Aversion, then $\nu$ is a convex capacity.[40] In this case, CEU preferences become a special case of MEU preferences, with the set of priors $\mathcal{P}$ being the core of the convex capacity $\nu$.[41] So for CEU preferences, ambiguity aversion implies aversion to partial information.

For CEU preferences, we can say a bit more about the connection between ambiguity attitudes and information preferences. We can also define ambiguity loving.[42]

**Axiom** (Ambiguity Loving)**.** *For all $f, g \in \mathcal{F}$ and $\alpha \in [0, 1]$,*

$$f \sim_0 g \Rightarrow f \succcurlyeq_0 \alpha f + (1 - \alpha) g$$

We show that within the CEU model, ambiguity aversion implies partial information aversion, and ambiguity loving implies partial information loving.

[40] A capacity $\nu$ is convex if $\nu(E \cup F) + \nu(E \cap F) \geq \nu(E) + \nu(F)$ holds for all $E, F \in \Sigma$.

[41] For a convex capacity $\nu$, its core is $\{p \in \Delta(S) | p(E) \geq \nu(E)$ for all $E \in \Sigma\}$.

[42] This is called "uncertainty appeal" in Schmeidler (1989).

**Proposition 3.** *Suppose $\succsim_0$, $\{\succsim_E\}_{E\in\Sigma}$ have CEU representations, and $\succsim$ is recursively generated by $\succsim_0$.*

1. *If $\succsim_0$ satisfies Ambiguity Aversion, then $\succsim$ exhibits partial information aversion at all acts.*
2. *If $\succsim_0$ satisfies Ambiguity Loving, then $\succsim$ exhibits attraction to partial information at all acts.*

## 5.3 Multiplier Preferences

Introduced by Hansen and Sargent (2001) to capture concerns about model misspecification, and later axiomatized by Strzalecki (2011), multiplier preferences have found broad applications in macroeconomics.[43] We say $\succsim_0$ has a multiplier preferences representation $(u, q, \theta)$ if it can be represented by a function $V_0 : \mathcal{F} \to \mathbb{R}$ of the form

$$V_0(f) = \min_{p\in\Delta(S)} [\int u(f)dp + \theta R(p||q)]$$

where $q \in \Delta(S)$ is the reference probability, $R(p||q) = \int \ln \frac{p}{q} dp$ is the relative entropy distance between probability $p$ and the reference probability $q$, and $\theta \in \mathbb{R}_+$ is a measure of the degree of ambiguity aversion.

**Proposition 4.** *Suppose $\succsim_0$ has a multiplier preferences representation $(u, q, \theta)$, and $\succsim$ is recursively generated by $\succsim_0$. Then $\succsim$ exhibits intrinsic information neutrality.*

## 5.4 Variational Preferences

Variational preferences are axiomatized by Maccheroni et al. (2006a,b). Formally, we say $\succsim_0$ has a variational representation $(u, c)$ if it can be represented by a function $V_0 : \mathcal{F} \to \mathbb{R}$ of the form

$$V_0(f) = \min_{p\in\Delta(S)} \int u(f)dp + c(p)$$

where $c : \Delta(S) \to [0, +\infty]$ is a convex, lower semicontinuous and grounded (there exists $p$ such that $c(p) = 0$) function. The domain of $c$ is $dom(c) = \{p : c(p) < +\infty\}$. In this section, we also assume that $u(X)$ is unbounded either above or below, so that function $c$ is uniquely

[43] See Hansen and Sargent (2007) and references therein.

identified.[44] The function $c$ is interpreted as the cost of choosing a probability. In particular, the variational preferences nest MEU and the multiplier preferences as special cases: If $c$ is an indicator function that equals to 0 on a set $\mathcal{P}$ and $+\infty$ elsewhere, then it is an MEU representation $(u, \mathcal{P})$; If $c$ is a positively scaled relative entropy with $c(p) = \theta R(p||q)$, then it is a multiplier representation $(u, q, \theta)$. Variational preferences are the most general class of ambiguity averse preferences that satisfy vertical invariance.

### 5.4.1 Updating Variational Preferences

For any non-empty $E \in \Sigma$, the conditional preferences $\succsim_E$ admits a variational representation $(u_E, c_E)$ if it can be represented by a function $V_E : \mathcal{F} \to \mathbb{R}$ of the form

$$V_E(f) = \min_{p_E \in \Delta(E)} \int_S u_E(f_E) dp_E + c_E(p_E)$$

where $f_E$ is the restriction of $f$ on $E$, $c_E : \Delta(E) \to [0, +\infty]$ is a convex, lower-semi-continuous, and grounded *conditional cost function.*

The next theorem shows that within the variational preferences family, Stable Risk Preferences and Conditional Certainty Equivalent Consistency characterize the following updating rule for conditional cost functions: for all $p_E \in \Delta(E)$,

$$c_E(p_E) = \inf_{\{p \in \Delta(S): p(\cdot|E) = p_E\}} \frac{c(p)}{p(E)} \tag{11}$$

where $p(\cdot|E)$ denotes the Bayesian posterior of probability $p$. Taking the infimum over all probabilities with posterior $p_E$ controls for any concern for model mis-specification outside event $E$, which is irrelevant to $\succsim_E$ due to consequentialism; normalization by $\frac{1}{p(E)}$ captures a maximum likelihood intuition: probabilities $p$ assigning a higher probability on the event that occurred are more likely to be selected and determine $c_E$. Since we imposed Strong Monotonicity on $\succsim_0$, every event $E$ is $\succsim_0$-non-null. In particular, $p(E) > 0$ for all $p \in dom(c)$. Then by Lemma 6 in the Appendix, the minimum in (11) attains at some $p$.

**Theorem 4.** *Suppose $\succsim_0$ admits a variational representation $(u, c)$ and satisfies Strong Monotonicity. Suppose for any non-empty $E \in \Sigma$, $\succsim_E$ admits a variational representation $(u_E, c_E)$. Then the following statements are equivalent:*

[44] The unboundedness assumption could be relaxed if we restrict the analysis to the minimal cost function $c^*$ (Maccheroni et al., 2006a, Theorem 3). Our analysis only relies on that in the chosen representation, function $c$ is uniquely identified for a fixed cardinalization of $u$.

1. *$\succsim_E$ and $\succsim_0$ satisfy Stable Risk Preferences and Conditional Certainty Equivalent Consistency.*

2. *$\succsim_E$ admits a variational representation $(u_E, c_E)$ such that $u_E = u$ and*

$$c_E(p_E) = \min_{\{p \in \Delta(S): p(\cdot|E) = p_E\}} \frac{c(p)}{p(E)}$$

This generalizes well-known updating rules for the two important subclasses of variational preferences: prior-by-prior updating in the MEU class, and Bayesian updating in the multiplier preferences class.

**Corollary 3.** *Suppose the assumptions and Statement 1 in Theorem 4 hold.*

1. *If $\succsim_0$ also admits a MEU representation $(u, \mathcal{P})$, then for any non-empty $E$, $\succsim_E$ has a MEU representation $(u, \mathcal{P}_E)$, where $\mathcal{P}_E$ is the set of prior-by-prior posteriors of $\mathcal{P}$, i.e.,*

$$\mathcal{P}_E = \{p(\cdot|E) | p \in \mathcal{P}\}.$$

2. *If $\succsim_0$ admits a multiplier preference representation $(u, q, \theta)$, then for any non-empty $E$, $\succsim_E$ has a multiplier preference representation $(u, q_E, \theta)$, where $q_E$ is the Bayesian posterior of $q$.*

### 5.4.2 Variational Preferences and Preferences for Partial Information

In general, recursive variational preferences might not exhibit aversion to partial information at all acts. This can be explained by the following intuition. Similar to the MEU model, variational preferences also has the intuitive interpretation of a malevolent Nature playing a zero-sum game against the DM (Maccheroni et al., 2006b). With variational preferences, Nature's constraint set is the domain of the cost function $c$, $dom(c)$. In addition, Nature has to pay a non-negative cost (or transfer) of $c(p)$ to the DM if it chooses a probability $p$ in $dom(c)$. Nature seeks to minimize the DM's expected utility plus the transfer. In our recursive model without reduction, information $\pi$ turns this into a sequential game, affecting both Nature's constraint set and how often Nature has to pay the DM a transfer. Similar to the MEU model, in period 0, Nature chooses a probability from $dom(c)$ for events in $\pi$. In period 1, Nature chooses a (possibly different) probability from $dom(c)$ over states for every event in $\pi$, conditional on that event. So information $\pi$ expands Nature's constraint set from $dom(c)$ to $rect_\pi(dom(c))$. On the other hand, with information $\pi$, Nature also needs to pay a non-negative transfer to the DM at every node where it chooses a probability. The total transfer can be higher or lower than what Nature would have paid in the static game, depending on the cost function $c$. If the total transfer is higher, then this helps the DM. So the overall effect from information $\pi$ is indeterminate. Below is an example in which when the transfer effect dominates and the DM strictly prefers information $\pi$ at an act $f$.

**Example 1** (Attraction to Partial Information in VP)**.** Suppose $S = \{s_1, s_2, s_3\}$. Let $u(x) = x$ (where $X = \mathbb{R}$). Consider the partition $\pi = \{\{s_1, s_2\}, \{s_3\}\}$. Let $E = \{s_1, s_2\}$. Let $\bar{p} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, and $\mathcal{P} = \{p \in \Delta(S) : p(s_i) \geq \delta, \forall i = 1, 2, 3\}$, for some $\delta \in (0, \frac{1}{5}]$.

Let $\alpha_{\bar{p}} = 0$. For all $p \in \mathcal{P} \backslash \bar{p}$, in the probability simplex illustrated by Figure 2, we connect $\bar{p}$ to $p$ by a line segment and extend it to a point $p'$ on the boundary of $\mathcal{P}$. Let $\alpha_p$ be the ratio of the length of line segment $\bar{p}p$ to the length of line segment $\bar{p}p'$. Consider the cost function

$$c(p) = \begin{cases} \alpha_p & \text{if } p \in P, \\ +\infty & \text{otherwise.} \end{cases}$$

Note that $c$ is convex, lower semicontinuous, and grounded, so $(u, c)$ characterizes some VP.

Consider the act $f = (0, 3K, 1K)$, where $K$ is a large number in $\mathbb{R}_+$ and $K\delta > 10$. Without information, $V(\pi_0, f) = 4\delta K + 1$. Suppose the DM now gets partial information $\pi$. Then

$$V_E(f) = \min_{p_E \in \Delta(E)} 3Kp_E(s_2) + \min_{p(\cdot|E) = p_E} \frac{c(p)}{p(E)} = \frac{1}{1-\delta}(3\delta K + 1)$$

$$V(\pi, f) = \min_p p(E) \frac{1}{1-\delta}(3\delta K + 1) + p(s_3)K + c(p) = 3\delta K + 1 + \delta K + 1 = 4\delta K + 2$$

Then $V(\pi, f) = 4\delta K + 2 > 4\delta K + 1 = V(\pi_0, f)$, so the DM has a strict preference for partial information $\pi$ at $f$.

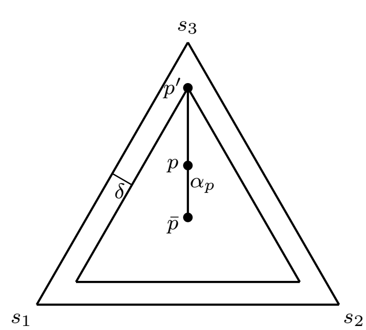


Figure 2: The probability simplex

The following proposition identifies a necessary and sufficient condition on the unconditional cost function $c$ under which aversion to partial information holds at all acts. In the zero-sum game against Nature interpretation, this condition ensures that the total transfer Nature pays under information $\pi$ does not exceed that in the static game. To formalize this, we need some additional notation. For all $p_E \in \Delta(E)$ and $p' \in \Delta(S)$, define $p_E \otimes_E p'$ by

$$(p_E \otimes_E p')(B) = p'(E)p_E(B) + p'(B \cap E^c) \qquad \text{for all } B \in \Sigma$$

That is, in $p_E \otimes_E p'$, we substitute $p'(\cdot|E)$ by $p_E$ for probability conditional on $E$, while measuring probabilities of events in $E^c$ (including $E^c$) by $p'$.

**Proposition 5.** *Suppose $\succcurlyeq_0$ has a variational representation $(u, c)$, and $\succcurlyeq$ is recursively generated by $\succcurlyeq_0$. Then $\succcurlyeq$ exhibits intrinsic aversion to partial information at all $f$ if and only if for any non-empty $E \in \Sigma$,*

$$c(p) \geq \min_{q_E \in \Delta(E)} c(q_E \otimes_E p) + p(E) \min_{q \in \Delta(S)} \frac{c(p_E \otimes_E q)}{q(E)}$$

*for all $p$ such that $p(E) > 0$. Here $p_E$ is the Bayesian posterior of $p$.*

It is straightforward to verify that this condition holds for MEU and multiplier preferences.

The above condition restricts the cost function $c$ so that $\succcurlyeq$ exhibits partial information aversion at all acts. As shown in Example 1, this can be violated by some variational preferences, where attraction to partial information at some act is possible. So this condition might be too strong for some purposes.

The next proposition characterizes a local sufficient condition on the cost function $c$ such that $\succcurlyeq$ exhibits aversion to partial information locally at $f$. This does not preclude the possibility that $\succcurlyeq$ exhibits attraction to partial information at some other act $g$.

**Proposition 6.** *Suppose $\succcurlyeq_0$ has a variational representation $(u, c)$, and $\succcurlyeq$ is recursively generated by $\succcurlyeq_0$. Then for any act $f$ such that*

$$c^{-1}(0) \cap \arg\min_{p \in \Delta}\left[\int_S u(f)dp + c(p)\right] \neq \emptyset \tag{12}$$

*$\succcurlyeq$ exhibits aversion to partial information at $f$.*

If $\succcurlyeq_0$ has MEU representation $(u, \mathcal{P})$, then the cost function is an indicator function where

$$c(p) = \delta_{\mathcal{P}}(p) = \begin{cases} 0 & \forall p \in \mathcal{P} \\ +\infty & \text{otherwise} \end{cases}$$

In this case, for any act $f$, $\arg\min_{p\in\Delta}[\int_S u(f)dp + c(p)] \subseteq \mathcal{P} = c^{-1}(0)$. So the result that an MEU DM is averse to partial information at all acts follows as a natural corollary of Proposition 6.

Condition (12) has an intuitive interpretation in terms of comparative ambiguity. Following the notion of comparative ambiguity aversion in Ghirardato and Marinacci (2002) and Epstein (1999), given two static preferences $\succcurlyeq_1$ and $\succcurlyeq_2$ over $\mathcal{F}$, we say $\succcurlyeq_1$ *is more ambiguity averse than* $\succcurlyeq_2$ if for all $f \in \mathcal{F}$ and $x \in X$,

$$f \succcurlyeq_1 x \Rightarrow f \succcurlyeq_2 x$$

By Maccheroni et al. (2006a) Proposition 8, if $\succcurlyeq_1$ has a variational representation $(u_1, c_1)$ and $\succcurlyeq_2$ has a variational representation $(u_2, c_2)$, then $\succcurlyeq_1$ is more ambiguity averse than $\succcurlyeq_2$ if and only if $u_1 \approx u_2$[45] and, provided $u_1 = u_2$, $c_1 \leq c_2$. We normalize risk utilities so that $u_1 = u_2$.

Fixing $\succcurlyeq$, we say an act $f$ can be locally approximated by an SEU preference that is less ambiguity averse than $\succcurlyeq_0$, if there exists a preference relation $\geq'$ on $\mathcal{F}$ that admits an SEU representation

$$U'(f) = \int_S u'(f)dq$$

such that (i) $\geq'$ is less ambiguity averse than $\succcurlyeq_0$ and (ii) $V(f) = U'(f)$.

**Proposition 7.** *Suppose $\succcurlyeq_0$ has a variational representation $(u, c)$. Condition (12) holds at some act $f$ if and only if $f$ can be locally approximated by an SEU preference that is less ambiguity averse than $\succcurlyeq_0$.*

An immediate implication of the proposition is that if $f$ can be locally approximated by an SEU preference that is less ambiguity averse than $\succcurlyeq_0$, then $\succcurlyeq$ exhibits aversion to partial information at $f$.

Finally, we provide a two-agent comparative version of the sufficient condition for local partial information aversion.

**Proposition 8.** *Suppose $\succcurlyeq_0^1$ has a variational representation $(u^1, c^1)$ and $f$ can be locally approximated by some SEU preference $\geq'$ that is less ambiguity averse than $\succcurlyeq_0^1$. Suppose $\succcurlyeq_0^2$ also has a variational representation $(u^2, c^2)$, and let $\succcurlyeq^2$ be recursively generated by $\succcurlyeq_0^2$. If $\succcurlyeq_0^2$ is less ambiguity averse than $\succcurlyeq_0^1$ and more ambiguity averse than $\geq'$, then $\succcurlyeq^2$ exhibits partial information aversion at $f$.*

[45] $u_1 \approx u_2$ if $u_1 = au_2 + b$, for some $a > 0$, $b \in \mathbb{R}$.

To sum up, in this section we explore the link between ambiguity attitudes and partial information attitudes within four families of ambiguity preferences. The key takeaway is a close yet somewhat delicate link between ambiguity aversion and aversion to partial information. When restricting to ambiguity preferences with rich structures, such as MEU, CEU, and MP preferences, we can draw a sharp conclusion that ambiguity aversion implies (at least weak) partial information aversion. Yet when we look at preferences with less structure, such as the variational preferences, the connection becomes less clear cut. This impression is also reinforced by our discussion of the Second-order Belief model in Section 7. A natural question to ask is what is the corresponding connection for other ambiguity preference families, especially the most general Uncertainty Averse Preferences (Cerreia-Vioglio et al., 2011). We think the main obstacle in this direction is the lack of understanding of what is a good updating rule for those preferences.

# 6 Multi-action Menus and Information Acquisition

In this section we consider when the DM can also choose from a menu of acts, and study the value of information under ambiguity. We model the scenario of a two-stage information acquisition problem. The DM can choose an act $f$ from a compact menu $F \subseteq \mathcal{F}$ and information partition $\pi \in \Pi$. The timing goes as follows. At the ex-ante stage, the DM first chooses to acquire some partial information $\pi$ and pays a cost $c(\pi)$. At the intermediate stage, she learns which event $E \in \pi$ has realized, and then chooses a act from the menu $F$. In the end, a state $s$ realizes and the DM receives the corresponding outcome.

The DM faces a standard trade-off between the value and cost of information. Her information acquisition decision can be expressed as

$$\max_{\pi \in \Pi} V(\pi, F) - c(\pi)$$

where $V(\pi, F)$ is the value she gets from having information $\pi$ and menu $F$, and $c : \Pi \to \mathbb{R}$ is the cost of acquiring $\pi$. Cost $c(\pi)$ is deterministic, and hence ambiguity will only affect the decision problem through value $V(\pi, F)$. So we can without loss of generality focus our analysis on the effect of ambiguity on the DM's ex-ante preferences over information and choice menus, when information is costless.

We model such preferences on an extended domain of information and menus. Let $\mathcal{M}$ be the collection of compact subsets of $\mathcal{F}$. We are interested in the DM's ex-ante preferences, denoted as $\succsim^+$, over the set of information and menus $\Pi \times \mathcal{M}$. For fixed menu $F \in \mathcal{M}$ and information $\pi = \{E_1, \cdots, E_n\}$, the DM can choose her most preferred act from menu $F$

contingent on which event in $\pi$ has realized. Hence the strategy space under a given $(\pi, F)$ is

$$F^\pi = \left\{ \begin{pmatrix} f_1 & \text{if } s \in E_1 \\ \vdots & \vdots \\ f_n & \text{if } s \in E_n \end{pmatrix} : f_i \in F, \forall i = 1, \cdots, n \right\}.$$

Note $F \subseteq F^\pi \subseteq \mathcal{F}$, and $F = F^\pi$ whenever $F$ is a singleton.

We assume the menu preferences $\succcurlyeq^+$ *is extended from* $\succcurlyeq$ in the following way:

$$(\pi, F) \succcurlyeq^+ (\pi', G) \Longleftrightarrow \forall g \in G^{\pi'}, \exists f \in F^\pi, (\pi, f) \succcurlyeq (\pi', g)$$

**Lemma 1.** *Suppose $V : \Pi \times \mathcal{F} \to \mathbb{R}$ represents $\succcurlyeq$. If $\succcurlyeq^+$ is extended from $\succcurlyeq$, then $\succcurlyeq^+$ is represented by $\tilde{V} : \Pi \times \mathcal{M} \to \mathbb{R}$ where*

$$\tilde{V}(\pi, F) = \max_{f \in F^\pi} V(\pi, f) = I_0 \begin{pmatrix} \max_{f \in F} V_{E_1}(f) & \textit{if } s \in E_1 \\ \vdots & \vdots \\ \max_{f \in F} V_{E_n}(f) & \textit{if } s \in E_n \end{pmatrix}$$

Since $\tilde{V}$ and $V$ agree on $\Pi \times \mathcal{F}$, we abuse notation and denote them by $V$ when there is no confusion. One remark is this extension is straightforward mainly because the single-act preferences $\succcurlyeq$ have the folding-back structure characterized in Theorem 1.

First we link information preferences in the single-action case to information preferences under menus. We say that $\succcurlyeq^+$ exhibits *a preference for perfect information* if $(\pi^*, F) \succcurlyeq^+ (\pi, F)$ for all $F \in \mathcal{M}$ and $\pi \in \Pi$. Proposition 9 says a preference for perfect information at all menus is equivalent to partial information aversion at all singleton acts.[46]

**Proposition 9.** *Suppose (i) $\succcurlyeq$ is recursively generated by $\succcurlyeq_0$; (ii) $\succcurlyeq^+$ is extended from $\succcurlyeq$. Then the following statements are equivalent:*

1. *$\succcurlyeq^+$ exhibits a preference for perfect information.*
2. *$\succcurlyeq$ exhibits partial information aversion.*
3. *$\succcurlyeq_0$ satisfies Event Complementarity.*

Next we explore some comparative static exercises. Particularly, we ask how the value of a decision problem, $V(\pi, F)$, will change if (i) the menu becomes larger and so the DM has greater flexibility; (ii) the DM is more ambiguity averse; (iii) the DM acquires finer information. For question (iii), define ranking $\pi_2 \geq (>) \pi_1$ where $\pi_2$ is (strictly) more informative than $\pi_1$, i.e., the partition $\pi_2$ is (strictly) finer than the partition $\pi_1$.

[46]This result is reminiscent of Proposition 2 in Dillenberger (2010).

**Proposition 10.** *The value of decision problem $V(\pi, F)$ has the following comparative statics:*

1. *If $F \subseteq F'$, then $V(\pi, F) \leq V(\pi, F')$.*
2. *Suppose $\succsim^1$ and $\succsim^2$ are recursively generated by variational preferences $\succsim_0^1$ and $\succsim_0^2$. If $\succsim_0^1$ is more ambiguity averse than $\succsim_0^2$, then $V^1(\pi, F) \leq V^2(\pi, F)$ for all $\pi, F$.*
3. *There exists cases when $\pi_2 > \pi_1$ yet $V(\pi_2, F) < V(\pi_1, F)$.*

Intuitively, part (1) says that in our model the DM always weakly prefers bigger menus, and hence displays a preference for flexibility. This prediction is distinct from that in Siniscalchi (2011), where the DM might prefer a smaller menu due to dynamic inconsistency and desire for commitment. Hence one potential way to test our model versus that of Siniscalchi (2011) is by testing preferences for flexibility versus commitment. Part (2) says that the more ambiguity averse the DM is, the less she values any information and menu pair $(\pi, F)$. Part (3) suggests with non-trivial ambiguity $V(\pi, F)$ is not monotone in information $\pi$. So more information can make the DM worse off.

Next we explore how ambiguity aversion can affect the value of information. Define the value of information $\pi$ at menu $F$ as the difference between the utility DM could get if she acquires information $\pi$ and her utility if she has the null information $\pi_0$. That is, the value of acquiring information $\pi$ is $\Delta V(\pi, F) := V(\pi, F) - V(\pi_0, F)$. In Appendix B.15, we also discuss the marginal value of increasing information, i.e., $V(\pi_2, F) - V(\pi_1, F)$ for $\pi_2 \geq \pi_1$.

The value of information $\pi$ in a decision problem $F$ can be decomposed into two parts. First, information allows the DM to fine-tune her strategy: anticipating $\pi$, the DM conditions her choice of optimal action on the event realized in $\pi$, so her strategy set expands from $F$ to $F^\pi$. This is traditionally called the *instrumental value* of information and is always non-negative (Blackwell, 1951). Second, as ambiguity aversion implies non-neutral attitude towards information, anticipating information also directly affects the DM's utility at any given act. We call this part the *intrinsic value* of information. Hence

$$\begin{aligned}\Delta V(\pi, F) &= V(\pi, F) - V(\pi_0, F) \\ &= [\max_{f \in F^\pi} V(\pi, F) - \max_{f \in F} V(\pi, f)] + [\max_{f \in F} V(\pi, f) - \max_{f \in F} V(\pi_0, f)]\end{aligned}$$

The first term reflects the instrumental value of information and the second term reflects the intrinsic value of information.

Next we look for conditions under which the value of information is non-negative, that is, the DM is still willing to acquire information $\pi$ when it is free, regardless of ambiguity.

Under ambiguity, the sign of $\Delta V(\pi, F)$ is unclear. It depends on the relative strength of instrumental versus intrinsic value of information. Nevertheless, we can still identify in some rather restrictive sufficient condition under which the value of information is non-negative. In that case, a DM will never reject free information, even if she is ambiguity averse.

Let $F_0 = \arg\max_{f \in F} V(\pi_0, f)$ be the set of uninformed optimal acts from the menu $F$.

**Proposition 11.** *If there exists an uninformed optimal act $f_0$ from menu $F$ that is $\pi$-measurable, then $\Delta V(\pi, F) \geq 0$.*

This implies the following interpersonal comparative statics on the value of information.

**Corollary 4.** *Suppose $\succcurlyeq_0^1$ and $\succcurlyeq_0^2$ admit VP representations $(u, c_1)$ and $(u, c_2)$ such that $c_1 \geq c_2$. Also $\succcurlyeq_0^{1+}$ and $\succcurlyeq_0^{2+}$ are extended as in Lemma 1. Fix any menu $F$ that contains at least one constant act. If there exists some constant act $x \in \arg\max_{f \in F} V^1(\pi_0, f)$, then $\Delta V^2(\pi, F) \geq 0$ for all $\pi \in \Pi$.*

In words, the corollary says that suppose DM 1 and DM 2 both have VP. Suppose for DM 1 some constant act (which is measurable to any information) is optimal in menu $F$ when uninformed. If DM 2 is more ambiguity averse than DM 1, then DM 2 will never reject any free information at menu $F$.

This raises the question whether a more ambiguity averse DM values information more. Appearance, the answer is no. As illustrated by Example 2, the value of information is non-monotone in the degree of ambiguity aversion.

**Example 2.** Suppose DM 1 has SEU preferences with belief $p \in \Delta(S)$. DM 2 has MEU preferences with non-singleton prior set $\mathcal{P} \subsetneq \Delta(S)$, and $\mathcal{P}$ is not rectangular with respect to some partition $\pi$ (therefore $\pi > \pi_0$). DM 3 has MEU preferences with prior set $\mathcal{Q} = rect_\pi(\mathcal{P})$. Assume further that these three DMs have the same risk preferences, so DM 3 is more ambiguity averse than DM 2, and DM 2 is more ambiguity averse than DM 1.

Since $\mathcal{P} \subsetneq \mathcal{Q}$, there exists $f \in \mathcal{F}$ such that $V^2(\pi_0, f) > V^3(\pi_0, f)$. Also $V^2(\pi, f) = V^3(\pi_0, f) = V^3(\pi, f)$.[47] Therefore

$$V^3(\pi, f) - V^3(\pi_0, f) > V^2(\pi, f) - V^2(\pi_0, f).$$

Increasing ambiguity aversion *increases* the value of information $\pi$ in this case.

[47] The argument is similar to that in the proof of Proposition 2.

Alternatively, DM 1 is intrinsically neutral to information, so $V^1(\pi, f) = V^1(\pi_0, f)$. Therefore

$$V^1(\pi, f) - V^1(\pi_0, f) = 0 > V^2(\pi, f) - V^2(\pi_0, f).$$

Increasing ambiguity aversion *decreases* the value of information $\pi$ in this case.

Finally, we provide an application of our model to a portfolio choice problem.

**Example 3** (Portfolio Choice)**.** Consider the portfolio choice example in Dow and Werlang (1992). Suppose there is a risk-neutral DM with wealth $W$. There is a risky asset with unit price $P$ and present value that is either high, $H$, or low, $L$. The DM has MEU preferences and believes the probability of $H$ belongs to the interval $[\underline{p}, \bar{p}]$. For simplicity, we assume the DM could choose to buy a unit of the risky asset ($B$), short-sell a unit of the risky asset ($S$), or not do anything ($N$). So $F = \{B, S, N\}$. The DM's optimal portfolio choice is

$$f_0^*(P) = \begin{cases} B & \text{if } \underline{p}H + (1-\underline{p})L > P; \\ N & \text{if } \bar{p}H + (1-\bar{p})L \geq P \geq \underline{\pi}H + (1-\underline{p})L; \\ S & \text{if } P > \bar{p}H + (1-\bar{p})L. \end{cases}$$

We now add an information acquisition stage before the portfolio choice. The DM can acquire a binary signal, $\pi = \{h, l\}$, which is correlated with the state of the risky asset, with $p(h|H) = p(l|L) = q > \frac{1}{2}$. We want to know if the DM will collect information $\pi$ if it is costless.

Suppose the DM's uninformed optimal choice is $B$. Then $V(\pi_0, B) = \underline{p}H + (1-\underline{p})L - P$, and $V(\pi, B) = [\underline{p}qH + (1-\underline{p})(1-q)L + \underline{p}(1-q)H + (1-\underline{p})qL - P] = V(\pi_0, B)$. By Lemma 9 in the appendix, $\pi$ is valuable. The other two cases could be calculated similarly. Without the need to compute the informed optimal strategies and $V(\pi, F)$, we can conclude that in this portfolio choice problem the DM will want to collect information $\pi$ if it is costless.

# 7 Discussion: Second-Order Belief Models

Another important class of ambiguity preferences is the second-order belief model (Klibanoff et al., 2005; Nau, 2006; Seo, 2009). We say $\succsim_0$ has a second-order belief representation if

$$V(\pi_0, f) = \int_{\Delta(S)} \phi[\int_S u(f) dp_\theta] d\mu$$

where $\mu \in \Delta(\Delta(S))$ is a second-order belief over the space of distributions $\Delta(S)$, and $\phi$ is a non-decreasing function capturing ambiguity attitude. When $\phi$ is smooth and concave (convex), the DM is ambiguity averse (loving).

For the second-order belief models, vertical invariance fails, and thus Conditional Certainty Equivalent Consistency cannot provide a well-defined update rule. Instead we adopt Bayes rule for the second-order belief $\mu$ as our update rule.

**Assumption 1.** Suppose $\succcurlyeq_0$ has a second-order belief representation $(u, \phi; \Theta, \mu)$. Then for any non-null event $E$, $\succcurlyeq_E$ has a second-order belief representation $(u_E, \phi_E; \Theta_E, \mu_E)$ satisfying

1. Risk and ambiguity attitudes are not updated: $u_E = u$, $\phi_E = \phi$.
2. Prior-by-prior updating of first order belief: $\Theta_E = \{p_\theta(\cdot|E)|p_\theta \in \Theta\}$.
3. Bayes rule for second-order belief:

$$\mu_E(\theta) = \frac{\mu(\theta)p_\theta(E)}{\int_\Theta p_{\theta'}(E)d\mu(\theta')} \tag{13}$$

In general, second-order belief models exhibit no systematic relation between ambiguity aversion and information aversion, as the following example illustrates.

**Example 4.** Consider the standard three color Ellsberg urn. Let $S = \{R, G, Y\}$ and $\Theta = \{(\frac{1}{3}, \frac{2}{3}\theta, \frac{2}{3}(1-\theta))|\theta = \frac{1}{3}, \frac{2}{3}\}$. Suppose the second-order prior $\mu$ puts equal probability on $p_{\frac{1}{3}} = (\frac{1}{3}, \frac{2}{9}, \frac{4}{9})$, and $p_{\frac{2}{3}} = (\frac{1}{3}, \frac{4}{9}, \frac{2}{9})$. Assume the DM is risk neutral with $u(x) = x$, and ambiguity averse with $\phi(y) = log(y)$. Information is given by the partition $\pi = \{\{R, G\}, \{Y\}\}$. Let $E = \{R, G\}$. Suppose the above update rule captures conditional preferences, so $\mu_E(p_{\frac{1}{3}}) = \frac{5}{12}$, and $\mu_E(p_{\frac{2}{3}}) = \frac{7}{12}$. By computation we can show that the DM is strictly averse to $\pi$ ($V(\pi, f) < V(\pi_0, f)$) at acts $f = (1, 0, 0)$ and $(0, 1, 1)$, and strictly loves $\pi$ ($V(\pi, f) > V(\pi_0, f)$) at acts $f = (0, 1, 0)$ and $(1, 0, 1)$.

Observe that the partition $\pi' = \{\{R\}, \{G, Y\}\}$ contains only events with known probabilities. The two acts $(1, 0, 0)$ and $(0, 1, 1)$, at which the DM is strictly averse to partial information $\pi$, are measurable with respect to $\pi'$ and thus unambiguous. This suggests that a DM with second-order belief preferences will be averse to partial information at acts where she has local ambiguity neutrality. The next proposition formalizes this idea.

Following Definition 4 in Klibanoff et al. (2005), we say $\succcurlyeq_0$ displays (local) smooth ambiguity neutrality at act $f$ if $V(\pi_0, f) = \phi[\int_{\Delta(S)} \int_S u(f)dp_\theta d\mu]$. In second-order belief models, ambiguity aversion only implies partial information aversion at the subclass of locally ambiguity neutral acts.

**Proposition 12.** *Suppose $\succsim_0$ and $\{\succsim_E\}_{E\in\Sigma}$ are second-order belief preferences, with update rule satisfying Assumption 1. If $\succsim_0$ is ambiguity averse (loving), then $\succsim$ exhibits partial information aversion (loving) at all acts where $\succsim_0$ displays (local) smooth ambiguity neutrality.*

# A Appendix: $\succsim_0$-non-null Events

This subsection clarifies the concept of a $\succsim_0$-non-null event for defining conditional preferences.

The literature normally adopts the condition of a non-null event from Savage. An event $E$ is *Savage $\succsim_0$-non-null* if there exists $f, g, h$, such that $fEh \succ_0 gEh$.

We consider a stronger condition: an event $E$ is $\succsim_0$*-non-null* if there exist constant acts $x^*, x_*$ such that $x^* \succ_0 x_*$ and $x^*Ex_* \succ_0 x_*$. An event $E$ is Savage $\succsim_0$-non-null if it is $\succsim_0$-non-null, but not vice versa. The next lemma compares how these two definitions differ in the variational preference family.

**Lemma 2.** *Suppose $\succsim_0$ has a variational representation $(u, c)$. An event $E$ is $\succsim_0$-non-null if and only if $p(E) > 0$ for all $p \in c^{-1}(0)$. An event $E$ is Savage $\succsim_0$-non-null if and only if there exists some act $f$ and some $p \in \arg\min_{p'\in\Delta(S)} \int u(f)dp' + c(p')$ such that $p(E) > 0$.*

*Proof.* For the first claim, we prove $E$ is $\succsim_0$-non-null iff $\exists p \in c^{-1}(0)$ such that $p(E) = 0$. Choose constant acts $x^*, x_*$ such that $x^* \succ_0 x_*$. First, suppose $\exists p \in c^{-1}(0)$ such that $p(E) = 0$. Then

$$V_0(x^*Ex_*) = u(x^*)p(E) + u(x_*)p(E^c) + c(p) = u(x_*) = V_0(x_*)$$

The first equality holds because $c(p) = 0$ and $p(E) = 0$. Next, suppose instead $p(E) > 0$ for all $p \in c^{-1}(0)$. Then let $p^* \in \arg\min_{p'} u(x^*)p'(E) + u(x_*)p'(E^c) + c(p')$. Either $p^* \in c^{-1}(0)$ and $p^*(E) > 0$, or $c(p^*) > 0$. In either case,

$$V_0(x^*Ex_*)u(x^*)p^*(E) + u(x_*)p^*(E^c) + c(p^*) > u(x_*)$$

so $x^*Ex_* \succ_0 x^*$.

For the second claim, suppose there exists some act $f$ and some $p \in argmin_{p'\in\Delta(S)} \int u(f)dp' + c(p')$ such that $p(E) > 0$. Then we can construct an act $f'$ such that $f'(s) = f(s)$ for all

$s \in E^c$, and $u(f'_s) = u(f_s) - \epsilon$ for all $s \in E$, and some $\epsilon > 0$. Since $p(E) > 0$,

$$\begin{aligned} V_0(f) &= \int_E u(f)dp + \int_{E^c} u(f)dp + c(p) \\ &> \int_E u(f)dp - \epsilon p(E) + \int_{E^c} u(f)dp + c(p) \\ &= \int_S u(f')dp + c(p) \geq V_0(f') \end{aligned}$$

So $f \succ_0 f'$. For the converse, suppose there exists $f, g, h$ such that $fEh \succ_0 gEh$. Let

$$p \in \arg\min_{p' \in \Delta(S)} \int u(gEh)dp' + c(p').$$

We argue that $p(E) > 0$. If instead $p(E) = 0$, then

$$V_0(gEh) = \int_E u(g)dp + \int_{E^c} u(h)dp + c(p) = \int_E u(f)dp + \int_{E^c} u(h)dp + c(p) \geq V_0(fEh)$$

This contradicts $fEh \succ_0 gEh$. □

Suppose $\succcurlyeq_0$ has an MEU representation $(u, \mathcal{P})$. As a corollary, $E$ is $\succcurlyeq_0$-non-null if and only if $p(E) > 0$ for all $p \in \mathcal{P}$. In contrast, $E$ is Savage $\succcurlyeq_0$-non-null if and only if there exists $f$ and $p \in \arg\min_{p \in \mathcal{P}} \int u(f)dp$ such that $p(E) > 0$.

For the results about updating, the stronger $\succcurlyeq_0$-non-null condition is needed. Pires (2002) shows that if the unconditional preferences $\succcurlyeq_0$ have an MEU representation $(u, \mathcal{P})$ and all priors give positive probability to event $E$, then Conditional Certainty Equivalence Consistency is satisfied if and only if $\succcurlyeq_E$ has an MEU representation $(u, \mathcal{P}_E)$, where $\mathcal{P}_E$ is the prior-by-prior updated posteriors from $\mathcal{P}$. In section 4.3, we show that if the unconditional preferences $\succcurlyeq_0$ have a variational representation $(u, c)$ and $p(E) > 0$ for all $p \in c^{-1}(0)$, then Conditional Certainty Equivalence Consistency is satisfied if and only if $\succcurlyeq_E$ has a variational representation $(u, c_E)$, where $c_E$ is obtained from $c$ using update rule (11). In both cases, $E$ has to be $\succcurlyeq_0$-non-null instead of Savage $\succcurlyeq_0$-non-null.

In the text, we impose Strong Monotonicity on $\succcurlyeq_0$ to ensure that updating is always well-defined. The following lemma follows directly by definition.

**Lemma 3.** *If $\succcurlyeq_0$ satisfies Strong Monotonicity, then every event $E$ in $\Sigma$ is $\succcurlyeq_0$-non-null.*

# B Appendix: Proofs

## B.1 Theorem 1

*Proof.* First, we verify the continuity of $\succcurlyeq$ on $\Pi \times \mathcal{F}$.

**Lemma 4.** *If $\succcurlyeq$ satisfies Stable Risk Preferences and for every $\pi$ the preference $\succcurlyeq_\pi$ is continuous on $\mathcal{F}$, then $\succcurlyeq$ is continuous on $\Pi \times \mathcal{F}$.*

*Proof.* Fix $(\pi, f)$. We want to show sets $U = \{(\pi', g) : (\pi', g) \succcurlyeq (\pi, f)\}$ and $L = \{(\pi', g) : (\pi, f) \succcurlyeq (\pi', g)\}$ are closed. Let $\{(\pi'_n, g_n)\}$ be a convergent sequence in the set $U$, with limit $(\pi', g)$. It suffices to show $(\pi', g)$ is also in $U$. Suppose not, then $(\pi, f) \succ (\pi', g')$. Since $\pi'_n \to \pi'$ in the discrete topology on $\Pi$, there exists some $N$ such that for all $n > N$, $\pi'_n = \pi'$. Continuity of $\succcurlyeq_\pi$ and convexity of $X$ ensure there exists a constant act $x_f$ such that $(\pi, f) \sim (\pi, x_f) \sim (\pi', x_f)$, where the last statement follows from Stable Risk Preferences. If $(\pi', x_f) \sim (\pi, f) \succ (\pi', g)$, then by continuity of $\succcurlyeq_{\pi'}$, there exists $M(> N)$ such that for all $n > M$, $(\pi', x_f) \succ (\pi', g_n) = (\pi'_n, g_n)$. So $(\pi, f) \succ (\pi'_n, g_n)$ for sufficiently large $n$, a contradiction to the assumption $\{(\pi'_n, g_n)\} \subseteq U$. Following similar arguments we can show set $L$ is closed. □

Second, we prove the sufficiency of Axiom 1 to 6 in Theorem 1.

Risk preferences. Fix arbitrary partition $\pi$, then the restriction of $\succcurlyeq$ on $\{\pi\} \times X$ is a continuous and independent preference relation on a mixture space $X$, thus by Herstein and Milnor (1953)'s Mixture Space Theorem it can be represented by an affine function $u : X \mapsto \mathbb{R}$. By Stable Risk Preferences, $(\pi, x) \succcurlyeq (\pi, y) \Leftrightarrow (\pi', x) \succcurlyeq (\pi', y)$ for all $\pi' \in \Pi$, thus $u$ also represents the restriction of $\succcurlyeq$ on $\{\pi'\} \times X$. And $u(X) \subseteq \mathbb{R}$ is a real interval since $X$ is connected.

Conditional Preferences. Fix arbitrary partition $\pi$ and event $E \in \pi$. Recall that $f \succcurlyeq_E g \Leftrightarrow (\pi, fEh) \succcurlyeq (\pi, gEh)$ for some $h \in \mathcal{F}$. The elicited conditional preferences $\succcurlyeq_E$ inherits continuity from that of $\succcurlyeq_\pi$. By definition of $\succcurlyeq_E$ and Partition Separability, $\succcurlyeq_E$ satisfies Consequentialism, i.e., $fEh \sim_E fEh'$ for all $f, h, h' \in \mathcal{F}$, and thus without loss of generality we can only look at the restriction of $\succcurlyeq_E$ on $\mathcal{F}_E$. Since $\succcurlyeq_\pi$ is strongly monotone on $\mathcal{F}$, $\succcurlyeq_E$ is strongly monotone on $\mathcal{F}_E$. Finally, $\succcurlyeq_E$ agrees with $\succcurlyeq_\pi$ on $X$ and thus can be represented by $u$.[48] Let $\succcurlyeq^*_E$ be the preference relation on $(u(X))^E$ induced by $\succcurlyeq_E$: $\xi_E \succcurlyeq^*_E \phi_E$ if and only if

[48]To see this, for all $x, y \in X$, $x \succcurlyeq_E y \Leftrightarrow (\pi, x) \succcurlyeq (\pi, yEx) \Leftrightarrow (\pi, x) \succcurlyeq (\pi, y)$. The first $\Leftrightarrow$ is by definition and Anticipated Partition Separability and the second $\Leftrightarrow$ follows from strong monotonicity of $\succcurlyeq_\pi$.

$f \succsim_E g$ for some $f, g \in \mathcal{F}$ such that $u(f_E) = \xi_E$ and $u(g_E) = \phi_E$. Then $\succsim_E^*$ is continuous and strongly monotone on $\succsim_E$. Thus there exists a continuous and strongly monotone functional $I_E : (u(X))^E \mapsto \mathbb{R}$ that represents $\succsim_E^*$. By Risk Independence, all $\succsim_E$ and $\succsim_\pi$ agree on $X$. Hence we can normalize them to $I_E(\bar{k}) = k$ for all $k \in u(X)$. Define conditional utility $V_E : \mathcal{F} \mapsto \mathbb{R}$ by $V_E(f) = I_E(u(f_E))$. Then

$$f \succsim_E g \Leftrightarrow u(f_E) \succsim_E^* u(g_E) \Leftrightarrow I_E(u(f_E)) \geq I_E(u(g_E)) \Leftrightarrow V_E(f) \geq V_E(g).$$

Moreover, for all $f \in \mathcal{F}$ there exists $x_{f,E} \in u^{-1}(I_E(u(f_E)))$, which is the $E$-conditional certainty equivalent of $f$.

Ex-ante preferences. Fix any $\pi \in \Pi$ and $f \in \mathcal{F}$, we can find conditional certainty equivalent $x_{f,E_i} \sim_{E_i} f$ for all $E_i \in \pi$. Denote $f^0 = f$, $f^1 = x_{f,E_1} E_1 f^0$, $f^2 = x_{f,E_2} E_2 f^1$, …, and $f^n = x_{f,E_n} E_n f^{n-1} = f_\pi$. Then by Definition 1 and Partition Separability, $(\pi, f) = (\pi, f^0) \sim (\pi, f^1) \sim \ldots \sim (\pi, f^n) = (\pi, f_\pi)$. Moreover, $(\pi, f_\pi) \sim (\pi^*, f_\pi) \sim (\pi_0, f_\pi)$, where the first indifference relation is by Indifference to Redundant Information and the second by Time Neutrality. By transitivity of $\succsim$, $(\pi, f) \sim (\pi_0, f_\pi)$. Thus for all $\pi = \{E_1, \ldots, E_n\}, \pi' = \{E'_1, \ldots, E'_m\}$ and $f, g \in \mathcal{F}$, let $f_\pi$ and $g_{\pi'}$ be the partition conditional certainty equivalents constructed as above. Then

$$\begin{aligned}
(\pi, f) \succsim (\pi', g) &\iff (\pi_0, f_\pi) \succsim (\pi_0, g_{\pi'}) \\
&\iff f_\pi \succsim_S g_{\pi'} \\
&\iff I_S(u(f_\pi)) \geq I_S(u(g_{\pi'})) \\
&\iff I_S \begin{pmatrix} V_{E_1}(f) & E_1 \\ V_{E_2}(f) & E_2 \\ \ldots & \\ V_{E_n}(f) & E_n \end{pmatrix} \geq I_S \begin{pmatrix} V_{E'_1}(g) & E'_1 \\ V_{E'_2}(f) & E'_2 \\ \ldots & \\ V_{E'_m}(f) & E'_m \end{pmatrix}
\end{aligned}$$

Third, we verify the necessity of Axiom 1 to 6. Suppose $\succsim$ is represented by $V : \Pi \times \mathcal{F} \mapsto \mathbb{R}$ as stated. The only axiom that is not straightforward to verify is Partition Separability. For all partitions $\pi = \{E, F_1, \ldots, F_m\}$ and $\pi' = \{E, G_1, \ldots, G_n\}$ that both contain $E$ and all acts $f, g, h, h' \in \mathcal{F}$,

$$\begin{aligned}
(\pi, fEh) \succcurlyeq (\pi, gEh) \;&\Leftrightarrow\; I_S\begin{pmatrix} V_E(f_E) & E \\ V_{F_1}(h_{F_1}) & F_1 \\ \cdots & \\ V_{F_m}(h_{F_m}) & F_m \end{pmatrix} \geq I_S\begin{pmatrix} V_E(g_E) & E \\ V_{F_1}(h_{F_1}) & F_1 \\ \cdots & \\ V_{F_m}(h_{F_m}) & F_m \end{pmatrix} \\
&\Leftrightarrow\; V_E(f_E) \geq V_E(g_E) \\
&\Leftrightarrow\; I_S\begin{pmatrix} V_E(f_E) & E \\ V_{G_1}(h'_{G_1}) & G_1 \\ \cdots & \\ V_{G_n}(h'_{G_n}) & G_n \end{pmatrix} \geq I_S\begin{pmatrix} V_E(g_E) & E \\ V_{G_1}(h'_{G_1}) & G_1 \\ \cdots & \\ V_{G_n}(h'_{G_n}) & G_n \end{pmatrix} \\
&\Leftrightarrow\; (\pi', fEh') \succcurlyeq (\pi', gEh').
\end{aligned}$$

where the second and third $\Leftrightarrow$ follow from strong monotonicity of $I_S$.

Finally, suppose both $(u, I_E)$ and $(u', I'_E)$ represent $\succcurlyeq_E$. Since both $u$ and $u'$ are affine representations of $\succcurlyeq_E$ on $X$, by the Mixture Space Theorem (Herstein and Milnor, 1953), $u' = au + b$ for some $a > 0$ and $b \in \mathbb{R}$. For all $f$, let $x_f$ be the constant act that $f \sim_E x_f$. Then

$$I_E(u(f_E)) = u(x_f)$$

$$I'_E(u'(f_E)) = u'(x_f)$$

Substituting $u' = au + b$, we get

$$I'_E(u'(f_E)) = I'_E(au(f_E) + \bar{b}) = u'(x_f) = au(x_f) + b$$

and thus $I'_E(au(f_E) + \bar{b}) = aI_E(u(f_E)) + b$. Since $f$ is arbitrary, we have for all $\xi \in (u'(X))^{|E|}$, $I'_E(\xi) = aI_E(\frac{\xi - \bar{b}}{a}) + b$.

$\square$

## B.2 Uniqueness of Conditional Certainty Equivalence

Recall the updating axiom

**Axiom** (Conditional Certainty Equivalent Consistency, CCEC)**.** *For all $f \in \mathcal{F}$, $x \in X$, and $E \in \Sigma$,*

$$f \sim_E x \Leftrightarrow fEx \sim_0 x \tag{14}$$

In Section 4.2, we commented that if $\succcurlyeq_0$ only satisfies continuity and strong monotonicity, then solution to $fEx \sim_0 x$ might not be essentially unique. For a counterexample, suppose $\succcurlyeq_0$ is represented by $(u, I_0)$. For given $f$ and $E$, let $\phi : u(X) \mapsto \mathbb{R}$ be $\phi(k) = I_0(u(f)E\bar{k})$. Find $x^* \succcurlyeq_0 f(s) \succcurlyeq_0 x_*$ for all $s \in E$ so that $k^* := u(x^*) \geq I_0(u(fEx^*)) = \phi(k^*)$ and $k_* := u(x_*) \leq I_0(u(fEx_*)) = \phi(k_*)$.

As depicted by Figure 3, $\phi$ is continuous and strictly increasing, thus it will cross the 45-degree line at least once, i.e., there exists a fixed point of $\phi$ on $[k_*, k^*]$. However, there might be more than one fixed point, as a result there might be $x$ and $y$ both are $E$-conditional certainty equivalent to $f$ yet $x \succ_0 y$.

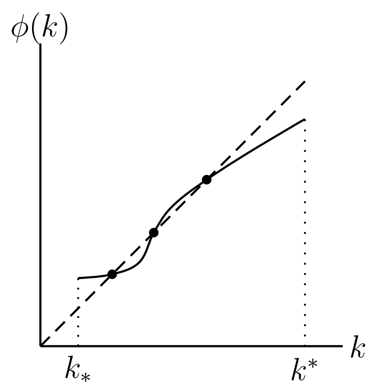


Figure 3: A continuous and monotone function with more than one fixed points.

A natural question is what additional structural assumption on $I_0$ can guarantee that there is only one solution. For all $E \subseteq S$, let $\mathbb{1}_E(s) = \begin{cases} 1 \text{ if } s \in E, \\ 0 \text{ otherwise} \end{cases}$. Below are sufficient conditions under which CCEC is well-defined.

**Lemma 5.** *If $I_0$ is Lipschitz continuous of rank 1, strongly monotone, and normalized, then $\phi$ is a contraction mapping. Hence for all $E \in \Sigma$ and $\xi \in u(X)^S$, there is a unique $k^* \in u(X)$ that solves*

$$k = I_0(\xi E\bar{k}).$$

*Proof.* If $I_0$ is Lipschitz continuous of rank 1, strongly monotone, and normalized, then

$$I_0(\xi + \epsilon\mathbb{1}_{E^c}) - I_0(\xi) < I_0(\xi + \epsilon\mathbb{1}) - I_0(\xi) \leq \epsilon$$

for all $\xi \in u(X)^S$ and $\epsilon \in u(X)_+$ such that $\xi + \epsilon\mathbb{1}_{E^c}, \xi + \epsilon\mathbb{1} \in u(X)^S$. The first $<$ is by strong monotonicity and the second $\leq$ is by rank-1 Lipschitz continuity of $I$.

To see $\phi$ is a contraction mapping, for all $k_1, k_2 \in u(X)$ such that $k_1 > k_2$ and $k_1 - k_2 \in u(X)$,

$$I_0(u(f)E\bar{k}_1) = I(u(f)E\bar{k}_2 + (k_1 - k_2)\mathbb{1}_{E^c}) < I_0(u(f)E\bar{k}_2) + k_1 - k_2$$

By definition of $\phi$, $0 < \phi(k_1) - \phi(k_2) < k_1 - k_2$ so it is a contraction mapping. Thus there exists a unique fixed point $k \in [k_*, k^*]$ such that $\phi(k) = k$. □

Suppose $\succcurlyeq_0$ is continuous, strongly monotone, and weak certainty independent, then it has a representation $(u, I_0)$ where $I_0$ is continuous, strongly monotone, and vertical invariant. It is straightforward to verify that $I_0$ satisfies conditions in Lemma 5 and CCEC is well-defined. This includes variational preferences, Maxmin EU, Choquet EU, and multiplier preferences.

## B.3 Theorem 3

We first cite a lemma from Maccheroni et al. (2006a).

**Lemma 28, Maccheroni et al. (2006a)** *A binary relation $\succcurlyeq_0$ on $\mathcal{F}$ satisfies Weak Order, Weak Certainty Independence, Continuity, Monotonicity, and Non-degeneracy if and only if there exists a nonconstant affine function $u : X \mapsto \mathbb{R}$ and a normalized, monotone, and vertical invariant $I_0 : (u(X))^S \mapsto \mathbb{R}$ such that*

$$f \succcurlyeq_0 g \Leftrightarrow I_0(u(f)) \geq I_0(u(g))$$

Below we will apply this result to prove our representation Theorem 3.

*Proof of Theorem 3.* Direction (2) $\Rightarrow$ (1). Suppose $\succcurlyeq$ admits a cross-partition recursive representation. By Theorem 1, Axiom 2 to 6 are satisfied. Weak Certainty Independence of $\succcurlyeq_0$ follows from Lemma 28 in Maccheroni et al. (2006a). To verify Belief Consistency, suppose $(\pi_E, fEx) \sim (\pi_E, x)$. Then by part 3 of the representation,

$$V(\pi_E, fEx) = I_0 \begin{pmatrix} V_E(f) & E \\ u(x) & E^c \end{pmatrix} = u(x) \Longrightarrow V_E(f) = u(x),$$

where the last step is due to strong monotonicity of $I_0$. By part 2 of the representation $V_E(f)$ also solves $k = I_0(u(f)E\bar{k})$. This implies $I_0(u(x)) = u(x) = I_0(u(f)Eu(x)) = I_0(u(fEx))$, i.e., $(\pi_0, x) \sim (\pi_0, fEx)$.

Direction (1) $\Rightarrow$ (2).

Step 1: Verify part 1 and part 3 of the cross-partition recursive utility representation.

Suppose $\succcurlyeq$ satisfies Axiom 2 to 6 and Axiom 8. Let $u : X \mapsto \mathbb{R}$ be the nonconstant affine utility index and let $\{I_E : u(X)^E \mapsto \mathbb{R}\}_{E\in\Sigma\setminus\emptyset}$ be the continuous, strongly monotone, normalized aggregators constructed in the proof of Theorem 1. Then conditional preferences $\succcurlyeq_E$ elicited via Definition 1 are represented by $V_E(f) = I_E(u(f_E))$. Moreover, $\succcurlyeq_{\pi_0}$ is represented by $V(\pi_0, f) = I_S\Big(V_S(f) \quad S\Big) = I_S(I_S(u(f)) = I_S(u(f))$, where the last equality follows from normalization of $I_S$. Let $I_0 = I_S$, then $\succcurlyeq_0=\succcurlyeq_{\pi_0}=\succcurlyeq_{\pi^*}$ is represented by $V_0 = I_0 \circ u$. Moreover, by Lemma 28 in Maccheroni et al. (2006a), Weak Certainty Independence of $\succcurlyeq_0$ implies that $I_0$ is vertically invariant. This proves part 1 of the utility representation. Part 3 of the utility representation follows directly from Theorem 1.

Step 2: Show that for all $f$ and nonempty $E \in \Sigma$, $V_E(f)$ uniquely solves $k = I_0[(u \circ f)E\bar{k}]$.

Let $x$ be the $E$-conditional certainty equivalent to $f$ so $V_E(f) = u(x)$. By Definition 1 and Partition Separability, $(\pi_E, fEx) \sim (\pi_E, x)$. Thus $(\pi_0, fEx) \sim (\pi_0, x)$ by Belief Consistency and $u(x)$ solves $k = I_0[(u \circ f)E\bar{k}]$.

It remains to check that $k = I_0[(u \circ f)E\bar{k}]$ has a unique solution in $u(X)$.

Existence of solution. Fix $f$ and nonempty $E$. Define $G(k) = I_0[(u \circ f)E\bar{k}] - k = I_0[(u \circ f - \bar{k})E0]$, for all $k \in u(X)$. Since $f$ is finite-ranged, we can find $x^*$, $x_*$ such that $x^* \succ_0 f(s) \succ_0 x_*$ for all $s$. Let $k^* = u(x^*)$, and $k_* = u(x_*)$. Then $G(k^*) \geq 0$ and $G(k_*) \leq 0$ by monotonicity of $I_0$. Since $I_0$ is continuous, $G$ is a continuous function of $k$ on $u(X)$. By the intermediate value theorem, there exists $k_0 \in [k_*, k^*]$ such that $G(k_0) = 0$.

Uniqueness of solution. Suppose $k_1$ and $k_2$ both solve $k = I_0[(u \circ f)E\bar{k}]$, and $k_1 \neq k_2$. Without loss of generality, let $k_1 > k_2$. By vertical invariance of $I_0$,

$$I_0[(u \circ f - \bar{k}_1)E0] = I_0[u(fE\bar{k}_1)] - k_1 = 0 = I_0[u(fE\bar{k}_2)] - k_2 = I_0[(u \circ f - \bar{k}_2)E0]$$

Then $(u \circ f - \bar{k}_1)E0 < (u \circ f - \bar{k}_2)E0$, since $E$ is non-empty. Since $I_0$ is strongly monotone, $I_0[(u \circ f - \bar{k}_1)E0] < I_0[(u \circ f - \bar{k}_2)E0]$. A contradiction.

Finally, suppose both $(u, I_0)$ and $(u', I_0')$ represent $\succcurlyeq_0$. Following the same argument as that in the proof of Theorem 1, there are some $a > 0$ and $b \in \mathbb{R}$ such that $u' = au + b$ and $I_0'(\xi) = aI_0(\frac{\xi - \bar{b}}{a}) + b$ for all $\xi \in (u'(X))^{|S|}$. Since $I_0$ is also vertically invariant, we have $I_0'(\xi) = aI_0(\frac{\xi}{a})$ for all $\xi \in (u'(X))^{|S|}$.

$\square$

## B.4 Proof of Corollary 1

*Proof.* For the if statement, suppose $(\pi_E, fEx) \sim (\pi_E, x)$. Then Definition 1 implies $f \sim_E x$, and $(\pi_0, fEx) \sim (\pi_0, x) \sim (\pi_E, x) \sim (\pi_E, fEx)$ by CEEC and Stable Risk Preferences.

For the only if statement, first suppose $f \sim_E x$, then by Definition 1 $(\pi, fEh) \sim (\pi, xEh)$ for some $h$ and $\pi$ such that $E \in \pi$. By Partition Separability, $(\pi_E, fEx) \sim (\pi_E, x)$. Thus $(\pi_0, fEx) \sim (\pi_E, fEx) \sim (\pi_E, x)$ by Belief Consistency and $(\pi_E, x) \sim (\pi_0, x)$ by Stable Risk Preferences. For the opposite direction, suppose $fEx \sim_0 x$ and $f \sim_E y$. We want to show $f \sim_E x$. By Belief Consistency, $fEy \sim_0 y$. By Continuity, Strong Monotonicity, and Weak Certainty Independence of $\succcurlyeq_0$, it can be represented by some nonconstant affine utility index $u$ and some strongly monotone, normalized, and vertically invariant aggregator $I_0$. Thus $I_0(u(f)Eu(x)) = u(x)$ and $I_0(u(f)Eu(y)) = u(y)$. By the same argument as that in the proof of Theorem 3, for given $f$ and $E$ the solution $k$ to $I_0(u(f)E\bar{k}) = k$ is unique. Hence $u(x) = u(y)$ and $f \sim_E y \sim_E x$. □

## B.5 Proof of Proposition 1

*Proof.* Fix $f, x, E$ such that $fEx \sim_0 x$. By vertical invariance of $I_0$,

$$I_0(u \circ (xEf)) = I_0(0E(u \circ f - u \circ x)) + u(x).$$

Since $fEx \sim_0 x$, $I_0(u \circ (fEx)) = u(x)$, thus

$$I_0(u \circ (xEf)) = I_0(0E(u \circ f - u \circ x)) + I_0(u \circ (fEx))$$

Thus $\succcurlyeq_0$ satisfies Event Complementarity, i.e., for all $f, E, x$ such that $fEx \sim_0 x$,

$$\begin{aligned} f \succcurlyeq_0 xEf \quad &\Leftrightarrow \quad I_0(u \circ f) \geq I_0(u \circ xEf) \\ &\Leftrightarrow \quad I_0(u \circ f) \geq I_0(u \circ fEx) + I_0(0E(u \circ f - u \circ x)) \end{aligned}$$

Note that the last inequality holds whenever $I_0$ is a supperadditive function: $I_0(\xi + \eta) \geq I_0(\xi) + I_0(\eta)$ for all $\xi, \eta \in u(X)^{|S|}$. □

## B.6 Proposition 2

*Proof.* For part (1), by Theorem 2, it suffices to show that $\succcurlyeq_0$ satisfies Event Complementarity. Since $\succcurlyeq_0$ belongs to the MEU class, by Lemma 3.3 in Gilboa and Schmeidler (1989), $I_0$ is superadditive. Event Complementarity follows from that.

For part (2), if $\succcurlyeq_0$ has an MEU representation $(u, \mathcal{P})$ and $\succcurlyeq$ is recursively generated by $\succcurlyeq_0$, then $\succcurlyeq$ can be represented by

$$\begin{aligned} V(\pi, f) &= \min_{p \in \mathcal{P}} \sum_{i=1}^{n} [\min_{p^i \in \mathcal{P}} \int u(f) dp^i(\cdot|E_i)] p(E_i) \\ &= \min_{p \in \mathcal{P}} \min_{p^i \in \mathcal{P}} \sum_{i=1}^{n} [\int u(f) dp^i(\cdot|E_i)] p(E_i) \\ &= \min_{p' \in rect_\pi(\mathcal{P})} \int u(f) dp' \end{aligned}$$

Suppose $\mathcal{P}$ is not $\pi$-rectangular, so there exists $q \in rect_\pi(\mathcal{P}) \backslash \mathcal{P}$. Since $\mathcal{P}$ is convex and compact, by the strict separating hyperplane theorem, there exists a nonzero, bounded and measurable map $\xi \in B(\Sigma, \mathbb{R})$ such that

$$\int \xi dq < \int \xi dp, \forall p \in \mathcal{P}$$

Without loss of generality, let $0 \in int(u(X))$. There exists $f \in \mathcal{F}$ such that $u(f) = \alpha\xi$, for some $\alpha > 0$. Thus without loss of generality we can replace $\xi$ by $u(f)$ in above inequality. By compactness of $\mathcal{P}$, $\min_{p \in \mathcal{P}} \int u(f) dp$ attains at some $p^* \in \mathcal{P}$, so using above

$$V(\pi, f) = \min_{q' \in rect_\pi(\mathcal{P})} \int u(f) dq' \leq \int u(f) dq < \int u(f) dp^* = V(\pi_0, f)$$

Thus $\succcurlyeq$ is strictly averse to partition $\pi$ at $f$.

For the converse, suppose $\mathcal{P}$ is $\pi$-rectangular, so $\mathcal{P} = rect_\pi(\mathcal{P})$. Then $V(\pi, f) = V(\pi_0, f), \forall f$, and $\succcurlyeq$ is intrinsically neutral to information $\pi$.

□

## B.7 Proposition 3

*Proof.* (1) Suppose $\succcurlyeq_0$ has a CEU representation $(u, \nu)$ and satisfies Uncertainty Aversion. By the Proposition in Schmeidler (1989) the corresponding functional $I_0$ is concave and superadditive. By Proposition 1, this implies that Event Complementarity holds. By Theorem 2, $\succcurlyeq$ exhibits aversion to partial information.

(2) Suppose $\succcurlyeq_0$ has a CEU representation $(u, \nu)$ and satisfies Uncertainty Loving. By Schmeidler (1989) Remark 6 the corresponding functional $I_0$ is convex and subadditive. By Proposition 1 and Theorem 2, $\succcurlyeq$ exhibits attraction to partial information. □

## B.8 Proposition 4

*Proof.* By Theorem 1 in Strzalecki (2011), if $\succcurlyeq_0$ has a multiplier representation, then Savage's Sure-Thing-Principle is satisfied. So for all $f \in F$ and $x$ such that $fEx \sim_0 x$, we have $f \sim_0 xEf$. By step 1 of our proof for Theorem 2, this yields information neutrality. $\square$

## B.9 Theorem 4 and Corollary 3

According to Appendix 2, if $\succcurlyeq_0$ satisfies Strong Monotonicity, then all nonempty events $E \in \Sigma$ are $\succcurlyeq_0$-non-null. That is, for all nonempty $E \in \Sigma$, $p(E) > 0$ for all $p \in c^{-1}(0)$.

**Lemma 6.** *For any non-empty event $E \in \Sigma$, suppose $p(E) > 0$ for all $p \in c^{-1}(0)$. Then the conditional cost function*

$$c_E(p_E) = \inf_{\{p \in \Delta(S): p(\cdot|E) = p_E\}} \frac{c(p)}{p(E)}$$

*attains a minimum value.*

*Proof.* Let $Q(p_E) := \{p \in \Delta(S) : p(\cdot|E) = p_E\}$. If $c(p) = +\infty$ for all $p \in Q(p_E)$, then $c_E(p_E) = +\infty$ and the minimum value attains at any $p \in Q(p_E)$. Otherwise, $c(p) < +\infty$ for some $p \in Q(p_E)$ and thus $c_E(p_E) < +\infty$. Define function $\phi : \Delta(S) \mapsto [0, +\infty]$ as $\phi(p) = \frac{c(p)}{p(E)}$.

We first show that $\phi(p)$ is a lower semicontinuous function on $\Delta(S)$. For all $r \in [0, +\infty]$, let $L_r = \{q \in \Delta(S) : \frac{c(q)}{q(E)} \leq r\}$. Take any sequence $q^n$ in $L_r$, which satisfies $c(q^n) \leq rq^n(E)$ for all $n$. This implies

$$c(q) \leq \lim_m \inf_{n \geq m} c(q^n) \leq r \lim_n q^n(E) = rq(E)$$

where the first inequality follows from lower semicontinuity of function $c$, the second inequality follows from that $q^n \in L_r$, and the last equality from $q^n \to q$.

Second, we show that $\overline{Q(p_E)} = Q(p_E) \cup \Delta(E^c)$ is compact in $\Delta(S)$, and hence $\inf_{p \in \overline{Q(p_E)}} \phi(p)$ attains its minimum on $\overline{Q(p_E)}$. To see this, $\overline{Q(p_E)}$ is obviously bounded and it remains to show that it is also closed. For all sequence $\{q^n\} \subseteq \overline{Q(p_E)}$, there exists a convergent subsequence $q^l \to q^* \in \Delta(S)$ since $\Delta(S)$ is compact. We want to show $q^* \in \overline{Q(p_E)}$. If $q^* \in \Delta(E^c)$, then we are done. Otherwise, $q^*(E) > 0$. Thus for sufficiently large $l$, $q^l(E) > 0$ and hence $q^l \in Q(p_E)$. This implies

$$\frac{q^*(B \cap E)}{q^*(E)} = \lim_l \frac{q^l(B \cap E)}{q^l(E)} = p_E(B) \qquad \text{for all } B \in \Sigma,$$

and thus $q \in Q(p_E)$.

Finally, we show if $c_E(p_E) = \frac{c(p^*)}{p^*(E)}$, that is, $\inf_{p \in \overline{Q(p_E)}} \phi(p)$ attains its minimum at $p^*$, then $p^* \in Q(p_E)$. Suppose not, then $p^* \in \Delta(E^c)$. Since $p^*(E) = 0$, we have $p^* \notin c^{-1}(0)$, i.e., $c(p^*) > 0$. Hence for all $p \in Q(p_E)$,

$$+\infty = \frac{c(p^*)}{p^*(E)} = \min_{p \in \overline{Q(p_E)}} \frac{c(p)}{p(E)} \leq \frac{c(p)}{p(E)} < +\infty$$

A contradiction.

□

We then verify that the conditional cost function $c_E(p_E) = \min_{\{p \in \Delta(S): p(\cdot|E) = p_E\}} \frac{c(p)}{p(E)}$ is convex, lower semicontinuous, and grounded.

**Lemma 7.** *The function $c_E : \Delta(E) \to [0, \infty]$ defined in (11) is (i) convex, (ii) lower semicontinuous, and (iii) grounded.*

*Proof.* Convexity. By the lower semicontinuity of $c$, for all $p_E, q_E \in \Delta(E)$ and $\alpha \in [0, 1]$, we can find $p^*, q^* \in \Delta(S)$ such that

$$p^*(\cdot|E) = p_E, q^*(\cdot|E) = q_E, \qquad c_E(p_E) = \frac{c(p^*)}{p^*(E)}, c_E(q_E) = \frac{c(q^*)}{q^*(E)}.$$

Fix $\alpha \in [0, 1]$. Pick $\gamma \in [0, 1]$ that solves $\frac{\gamma p^*(E)}{\gamma p^*(E) + (1-\gamma) q^*(E)} = \alpha$. Set $p' := \gamma p^* + (1 - \gamma) q^*$. Then

$$\begin{aligned} c_E(\alpha p_E + (1 - \alpha) q_E) &\leq \frac{c(p')}{p'(E)} \leq \frac{\gamma c(p^*) + (1 - \gamma) c(q^*)}{\gamma p^*(E) + (1 - \gamma) q^*(E)} \\ &= \frac{\gamma p^*(E)}{\gamma p^*(E) + (1 - \gamma) q^*(E)} c_E(p_E) + \frac{(1 - \gamma) q^*(E)}{\gamma p^*(E) + (1 - \gamma) q^*(E)} c_E(q_E) \\ &= \alpha c_E(p_E) + (1 - \alpha) c_E(q_E). \end{aligned}$$

The first inequality is by definition and that $p'(\cdot|E) = \alpha p_E + (1 - \alpha) q_E$. The second inequality follows from convexity of $c$.

Lower semicontinuity. We want to show that for all $r \in [0, +\infty]$, $L_r = \{p_E \in \Delta(E) : c_E(p_E) \leq r\}$ is closed in $\Delta(E)$. To that end, take any sequence $\{p_E^n\}$ in $L_r$ and $p_E^n \to p_E \in \Delta(E)$. Then $c_E(p_E^n) \leq r$ for all $n$. By Lemma 6, there exists a corresponding sequence $p^n \in \Delta(S)$ such that $p^n(\cdot|E) = p_E^n$ and $\frac{c(p^n)}{p^n(E)} = c_E(p_E^n)$. This implies $c(p^n) \leq r p^n(E)$

for all $n$. Since $\Delta(S)$ is compact, there exists a convergent subsequence of $p^k$ such that $p^k \to p^* \in \Delta(S)$. Again we have

$$c(p^*) \leq \liminf c(p^k) \leq r \lim p^k(E) = rp^*(E).$$

If $p^*(E) > 0$, then its Bayesian posterior, denoted $p^*_E$, exists and $p^*_E = \lim_k p^k_E = p_E$. So $c_E(p^*_E) \leq \frac{c(p^*)}{p^*(E)} \leq r$ and we are done. If $p^*(E) = 0$, then the above inequalities imply $c(p^*) = 0$. This contradicts our assumption.

Groundedness. $c$ is grounded, so there exists $p^*$ such that $c(p^*) = 0$. By assumption, $p^*(E) > 0$, so $c_E(p^*(\cdot|E)) = 0$. □

For vectors $\phi, p \in \mathbb{R}^{|S|}$, $\langle \phi, p \rangle$ denotes the inner product.

**Lemma 8.** *Consider two variational functionals $I(\phi) = \min_{p\in\Delta}\langle \phi, p\rangle + c(p)$, and $I'(\phi) = \min_{p\in\Delta}\langle \phi, p\rangle + c'(p)$. If $c(p_0) < c'(p_0)$ for some $p_0$, then there exists $\xi \in B(\Sigma)$ such that $I(\xi) < I'(\xi)$.*

*Proof.* Consider the epigraph of $c'$:

$$epi(c') = \{(p, r) \in \Delta \times \mathbb{R} | r \geq c'(p)\} \subseteq \Delta \times [0, +\infty]$$

Since $c'$ is nonnegative, convex, lower semicontinous, and grounded, $epi(c')$ is nonempty, closed and convex. Let $r_0 = c(p_0)$. Since $c(p_0) < c'(p_0)$, $(p_0, r_0) \notin epi(c')$. By the strict separating hyperplane theorem there exists $(\xi_0, r^*) \in \mathbb{R}^{|S|+1}$, $(\xi_0, r^*) \neq 0$, that strictly separates $(p_0, r_0)$ from the set $epi(c')$, i.e.,

$$\langle \xi_0, p_0\rangle + r_0 \cdot r^* < \min_{r' \geq c'(p')} \langle \xi_0, p'\rangle + r' \cdot r^* \tag{15}$$

[49] Note that we cannot have $r^* < 0$, otherwise we could take $r' = +\infty$ in the right hand side and the inequality fails. Also we cannot have $r^* = 0$, otherwise we can pick $(p_0, r') \in epi(c')$ and get $\langle \xi_0, p_0\rangle < \langle \xi_0, p_0\rangle$. Thus $r^* > 0$. Multiplying both sides of (15) by $\frac{1}{r^*}$ (take $\xi = \frac{1}{r^*}\xi_0$) yields

$$\langle \xi, p_0\rangle + r_0 < \langle \xi, p'\rangle + r' \qquad \forall (p', r') \in epi(c')$$

Using the fact that $r_0 = c(p_0)$ yields

$$\langle \xi, p_0\rangle + r_0 = \langle \xi, p_0\rangle + c(p_0) \geq \min_{p\in\Delta}\langle \xi, p\rangle + c(p) = I(\xi).$$

By definition,

$$\min_{r' \geq c'(p')} \langle \xi, p'\rangle + r' = \min_{p'\in\Delta}\langle \xi, p'\rangle + c'(p') = I'(\xi)$$

Thus $I(\xi) \leq \langle \xi, p_0\rangle + r_0 < \min_{r' \geq c'(p')}\langle \xi, p'\rangle + r' = I'(\xi)$. □

[49] Since $\inf_{r' \geq c'(p')}\langle \xi, p'\rangle + r' = \min_{p'\in\Delta}\langle \xi, p'\rangle + c'(p')$, the right hand side always attains its minimum.

*Proof of Theorem 4.* (2) $\Rightarrow$ (1). Suppose (2) holds. It is straightforward to verify Stable Risk Preferences and Consequentialism. We prove Conditional Certainty Equivalent Consistency also holds.

Fix $f \in \mathcal{F}$ and $x \in X$ such that $x \sim_E f$. We must prove $fEx \sim_0 x$. Suppose $c$ and $c_E$ satisfy update rule (11). Then

$$\begin{aligned} x \sim_E f \Rightarrow u(x) &= \min_{p_E \in \Delta(E)} \int_E u(f) dp_E + c_E(p_E) \\ &= \min_{p_E \in \Delta(E)} \left( \int_E u(f) dp_E + \inf_{\{p \in \Delta(S): p(\cdot|E) = p_E\}} \frac{c(p)}{p(E)} \right) \end{aligned} \tag{16}$$

The infimum above attains at some $p^* \in \Delta$.[50] Substituting $p^*(E)u(x) + p^*(E^c)u(x)$, we have

$$\begin{aligned} u(x) &= p^*(E) \left[ \int_E u(f) dp^*(\cdot|E) + \tfrac{c(p^*)}{p^*(E)} \right] + p^*(E^c)u(x) \\ &= \int_E u(f) dp^* + p^*(E^c)u(x) + c(p^*) \\ &\geq \min_{p \in \Delta} \int_E u(f) dp + p(E^c)u(x) + c(p) = V_0(fEx) \end{aligned}$$

Suppose $u(x) > V_0(fEx)$. Let $\tilde{p} \in \arg\min_{p \in \Delta} \int_E u(fEx) dp + p(E^c)u(x) + c(p)$. Then

$$\begin{aligned} u(x) > V_0(fEx) &= \min_{p \in \Delta} \int_E u(f) dp + p(E^c)u(x) + c(p) \\ &= \int_E u(f) d\tilde{p} + \tilde{p}(E^c)u(x) + c(\tilde{p}) \end{aligned}$$

If $\tilde{p}(E) = 0$, then $u(x) > u(x) + c(\tilde{p})$, which contradicts the non-negativity of $c$. So $\tilde{p}(E) > 0$. Then

$$\begin{aligned} u(x) &> \frac{1}{\tilde{p}(E)} [\int_E u(f) d\tilde{p} + c(\tilde{p})] \\ &= \int_E u(f) d\tilde{p}(\cdot|E) + \frac{c(\tilde{p})}{\tilde{p}(E)} \\ &\geq \min_{p_E \in \Delta(E)} \left( \int_E u(f) dp_E + \min_{\{p \in \Delta(S): p(\cdot|E) = p_E\}} \frac{c(p)}{p(E)} \right) \\ &= V_E(f) \end{aligned}$$

This contradicts the assumption that $x \sim_E f$. So $u(x) = V_0(fEx)$ and hence $fEx \sim_0 x$.

[50] Let $I_E : u(X)^{|E|} \to \mathbb{R}$ be such that $I_E(\xi) = \inf_{p_E \in \Delta(E)} \int_E \xi dp_E + c_E(p_E)$. Then $I_E$ is also a variational functional. Applying Maccheroni et al. (2006a) Lemma 26, the infimum attains at some $p_E^*$. In addition, if $p(E) > 0$ for all $p \in c^{-1}(0)$, by the previous lemma there exists some $p^* \in \Delta(S)$, at which $p^*(\cdot|E) = p_E^*$ and the second infimum attains.

For the converse, suppose $fEx \sim_0 x$. We want to show $f \sim_E x$. Note that

$$\begin{aligned} u(x) = V_0(fEx) &= \min_{p\in\Delta(S)} \int_E u(f)dp + u(x)p(E^c) + c(p) \\ &= \int_E u(f)dp^* + u(x)p^*(E^c) + c(p^*) \end{aligned}$$

where $p^* \in \arg\min_p \int_E u(f)dp + u(x)p(E^c) + c(p)$. If $p^*(E) = 0$, then the equality above implies $c(p^*) = 0$, a contradiction to the assumption that $p(E) > 0$ for all $p \in c^{-1}(0)$. So $p^*(E) > 0$ and

$$p^*(E)u(x) = \int_E u(f)dp^* + c(p^*).$$

Thus

$$\begin{aligned} u(x) &= \int_E u(f)dp^*(\cdot|E) + \frac{c(p^*)}{p^*(E)} \\ &\geq \min_{p_E\in\Delta(E)} \left( \int_E u(f)dp_E + \min_{\{p\in\Delta(S):p(\cdot|E)=p_E\}} \frac{c(p)}{p(E)} \right) = V_E(f) \end{aligned}$$

So $x \succcurlyeq_E f$.

Following the same argument as (16), $V_E(f) = \int_E u(f)dq^*(\cdot|E) + \frac{c(q^*)}{q^*(E)}$ at some $q^* \in \Delta(S)$ such that $q^*(E) > 0$. Hence

$$\begin{aligned} & q^*(E) \cdot V_E(f) + q^*(E^c) \cdot u(x) \\ = \; & q^*(E) \left[ \int_E u(f)dq^*(\cdot|E) + \tfrac{c(q^*)}{q^*(E)} \right] + q^*(E^c)u(x) \\ \geq \; & V_0(fEx) = u(x) \end{aligned}$$

This implies $V_E(f) \geq u(x)$. Together we have $x \sim_E f$.

(1) $\Rightarrow$ (2). By assumption, $\succcurlyeq_E$ is represented by

$$V_E(f) = \min_{p\in\Delta(S)} \int_S u_E(f)dp + c_E(p)$$

By Stable Risk Preferences, $\succcurlyeq_0$ and $\succcurlyeq_E$ agree on constant acts $X$. We can normalize by setting $u_E = u$. Next we want to show only $p$ with support on $E$ can achieve the minimum defining $V_E$. For each $f \in \mathcal{F}$, choose $p^* \in \arg\min_{p\in\Delta(S)} \int_S u(f)dp + c_E(p)$. Without loss of generality, we can choose $x_* \in X$ such that $f(s) \succ_0 x_*$ for all $s$.[51] Since $(fEx_*)Ex = fEx$ for any $x$, by Conditional Certainty Equivalent Consistency, $fEx_* \sim_E f$. Then

$$V_E(f) = \int_S u(f)dp^* + c_E(p^*) = V_E(fEx_*) \leq \int_E u(f)dp^* + p^*(E^c)u(x_*) + c_E(p^*)$$

[51]If not, then $u(X)$ is bounded below and $\min_s u(f)(s)$ achieves the lower bound. By vertical invariance $p^*$ is also a minimizing probability for $f'$ such that $u(f') = u(f) + \epsilon$. Then the whole argument works for $f'$.

So $\int_{E^c}(u(f) - u(x_*))dp^* \leq 0$. Since $u(f) - u(x_*)$ is strictly positive on $E^c$, $\int_{E^c}(u(f) - u(x_*))dp^* \geq 0$, and this is an equality if and only if $p^*(E^c) = 0$. So $p^*(E) = 1$, and $p^*$ has a natural imbedding in $\Delta(E)$. Therefore $\forall f$,

$$V_E(f) = \min_{p\in\Delta(E)} \int_E u(f)dp + c_E(p)$$

Suppose $\succcurlyeq$ and $\succcurlyeq_E$ are represented by

$$V(f) = \min_{p\in\Delta(S)} \int_S u(f)dp + c(p)$$
$$V_E(f) = \min_{p_E\in\Delta(E)} \int_S u_E(f)dp_E + c_E(p_E)$$

respectively.

It remains to show that given $u = u_E$ and the unconditional cost function $c$, the conditional cost function $c_E$ coincides with $\tilde{c}_E(p_E) := \min_{\{p\in\Delta(S):p(\cdot|E)=p_E\}} \frac{c(p)}{p(E)}$. Suppose instead $c_E \neq \tilde{c}_E$. Thus there exists $p_E^*$ such that $c_E(p_E^*) \neq \tilde{c}_E(p_E^*)$. We prove a contradiction for the case $c_E(p_E^*) > \tilde{c}_E(p_E^*)$. Applying Lemma 8, we can find $\xi_E \in \mathbb{R}^{|E|}$ such that $\min_{p_E} \int_E \xi_E dp_E + \tilde{c}_E(p_E) < \min_{p_E} \int_E \xi_E dp_E + c_E(p_E)$. There is an act $f \in \mathcal{F}$ and a constant $k$ such that $(u(f) + k)(s) = \xi_E(s)$ for all $s \in E$. By vertical invariance, $\min_{p_E} \int_E u(f)dp_E + \tilde{c}_E(p_E) < \min_{p_E} \int_E u(f)dp_E + c_E(p_E)$.

By Continuity, we can find $x \in X$ that is the $E$-conditional equivalent of $f$, $x \sim_E f$, and $u(x) = V_E(f) = \min_{p_E} \int_E u(f)dp_E + c_E(p_E)$.

Then

$$\begin{aligned}
u(x) &= \min_{p_E} \int_E u(f)dp_E + c_E(p_E) \\
&> \min_{p_E} \int_E u(f)dp_E + \tilde{c}_E(p_E) \\
&= \min_{p_E} \left( \int_E u(f)dp_E + \min_{\{p\in\Delta(S):p(\cdot|E)=p_E\}} \frac{c(p)}{p(E)} \right) \\
&= \min_{p_E} \min_{\{p\in\Delta(S):p(\cdot|E)=p_E\}} \int_E u(f)dp_E + \frac{c(p)}{p(E)} \\
&= \min_{\{p\in\Delta(S):p(E)>0\}} \frac{1}{p(E)} \left[ \int_E u(f)dp + c(p) \right]
\end{aligned}$$

As argued before, we can find $\underline{p} \in \arg\min_{\{p\in\Delta(S):p(E)>0\}} \frac{1}{p(E)} \left[ \int_E u(f)dp + c(p) \right]$. Then multiplying both sides of the inequality by $\underline{p}(E)$ and adding $\underline{p}(E^c)u(x)$ to both sides yields

$$\begin{aligned} u(x) &> \underline{p}(E)\left(\tfrac{1}{\underline{p}(E)}[\textstyle\int_E u(f)d\underline{p} + c(\underline{p})]\right) + \underline{p}(E^c)u(x) \\ &= \int_E u(f)d\underline{p} + \underline{p}(E^c)u(x) + c(\underline{p}) \\ &= \int u(fEx)d\underline{p} + c(\underline{p}) > V_0(fEx) \end{aligned}$$

So $x \succ_0 fEx$, violating Conditional Certainty Equivalent Consistency.

The case of $c_E(p_E^*) < \tilde{c}_E(p_E^*)$ can be proved by the same arguments. □

*Proof of Corollary 3.* For part (1), suppose $\succcurlyeq_0$ has a MEU representation $(u, \mathcal{P})$. So it has a variational representation $(u, c)$ with cost function $c$ such that $c(p) = 0$ if $p \in \mathcal{P}$ and $c(p) = +\infty$ if $p \notin \mathcal{P}$. For any nonempty event $E$, Strong Monotonicity of $\succcurlyeq_0$ ensures that $p(E) > 0$ for all $p \in \mathcal{P}$. Let $\mathcal{P}_E = \{p(\cdot|E)|p \in \mathcal{P}\}$. Applying updating rule (11),

$$c_E(p_E) = \begin{cases} 0 & \text{if } p_E \in \mathcal{P}_E \\ +\infty & \text{otherwise} \end{cases}$$

So $\succcurlyeq_E$ admits an MEU representation $(u, \mathcal{P}_E)$.

For part (2), suppose $\succcurlyeq_0$ also has a multiplier preference representation $(u, q, \theta)$. So it has a variational representation $(u, c)$ with cost function $c(p) = \theta \int ln\frac{p}{q}dp$. For any nonempty event $E$, Strong Monotonicity of $\succcurlyeq_0$ ensures that $q(E) > 0$. Applying updating rule (11),

$$\begin{aligned} c_E(p_E) &= \min_{\{p \in \Delta(S): p(\cdot|E) = p_E\}} \frac{\theta}{p(E)} \int ln\frac{p}{q}dp \\ &= \min_{\{p \in \Delta(S): p(\cdot|E) = p_E\}} \frac{\theta}{p(E)}[\int_E ln\frac{p_E}{q_E}dp_E)p(E) + (\int_{E^c} ln\frac{p_{E^c}}{q_{E^c}}dp_{E^c})p(E^c) \\ &\qquad + (p(E)ln\frac{p(E)}{q(E)} + p(E^c)\ln\frac{p(E^c)}{q(E^c)})] \\ &= \theta \int_E ln\frac{p_E}{q_E}dp_E \end{aligned}$$

In the last step, we choose a minimizing $p$ that satisfies $p(E) = q(E)$ and $p(\cdot|E^c) = q(\cdot|E^c)$. So $\succcurlyeq_E$ admits a multiplier representation $(u, q_E, \theta)$. □

## B.10 Proposition 5

*Proof.* By Theorem 2, $\succcurlyeq$ exhibits intrinsic information aversion at all acts if and only if $\forall f$ and $x$ such that $fEx \sim_0 x$, $f \succcurlyeq_0 xEf$. By Conditional Certainty Equivalent Consistency,

$fEx \sim_0$ if and only if $x \sim_E f$.

If $x \sim_E f$, then $u(x) = \min_{p_E \in \Delta(E)} \int_E u(f) dp_E + c_E(p_E)$. So

$$
\begin{aligned}
V_0(xEf) &= \min_{p \in \Delta} p(E)u(x) + \int_{E^c} u(f)dp + c(p) \\
&= \min_{p \in \Delta} p(E)[\min_{p_E \in \Delta(E)} \int_E u(f)dp_E + \hat{c}_E(p_E)] + \int_{E^c} u(f)dp + c(p) \\
&= \min_{p \in \Delta} \min_{p_E \in \Delta(E)} p(E)[\int_E u(f)dp_E + \hat{c}_E(p_E)] + + \int_{E^c} u(f)dp + c(p) \\
&= \min_{q \in \Delta} \min_{q_E \in \Delta(E)} \int u(f)dq + q(E)\hat{c}_E(q_E) + c(q_E \otimes_E q) \\
&\text{(change of variable: } q = p_E \otimes_E p, \text{ and } q_E = p(\cdot|E)) \\
&= \min_{q \in \Delta} \int u(f)dq + q(E)\hat{c}_E(q_E) + \min_{q_E \in \Delta(E)} c(q_E \otimes_E q)
\end{aligned}
$$

Also

$$V_0(f) = \min_{q \in \Delta} \int u(f)dq + c(q)$$

"If" direction. Suppose $\inf_{q_E \in \Delta(E)} c(q_E \otimes_E p) + p(E) \inf_{q \in \Delta(S)} \frac{c(p_E \otimes q)}{q(E)} \leq c(p), \forall p$. Then for all $f$, $q$,

$$\int u(a)dq + q(E)\hat{c}_E(q_E) + \min_{q_E \in \Delta(E)} c(q_E \otimes_E q) \leq \int u(f)dq + c(q),$$

so $V_0(xEf) \leq V_0(f)$. Thus the DM is averse to partial information at all $f$.

"Only if" direction. For each $E \in \Sigma$, define

$$\tilde{c}(p) = \begin{cases} \inf_{q_E \in \Delta(E)} c(q_E \otimes_E p) + p(E)c_E(p(\cdot|E)) & \text{if } p(E) > 0 \\ +\infty & \text{otherwise} \end{cases}$$

Define $\tilde{I} : B(S, \mathbb{R}) \to \mathbb{R}$ by $\tilde{I}(\xi) = \inf_{p \in \Delta(S)} \int \xi dp + \tilde{c}(p)$. By the calculation above, we have $\forall f \in \mathcal{F}$, $x \sim_E f$, $V_0(xEf) = \tilde{I}(u(f))$.

If statement (2) fails, then there exists $p$ such that $\tilde{c}(p) > c(p)$. By Lemma 8, we can find $\xi \in B(S, \mathbb{R})$ such that $\tilde{I}(\xi) > I(\xi)$. By unboundedness, $B(S, \mathbb{R}) \subseteq B(S, u(X)) + \mathbb{R}$, so there exists $f \in F$ such that $u(f) + k = \xi$ for some constant $k$. So we can find $f \in \mathcal{F}$ such that $V_0(xEf) = \tilde{I}(u(f)) > I(u(f)) = V_0(f)$. This contradicts aversion to partial information. $\square$

## B.11 Proposition 6

*Proof.* Let $p^* \in c^{-1}(0) \cap \arg\min_{p\in\Delta}[\int_S u(f)dp + c(p)]$. Then $\forall \pi = \{E_1, \cdots, E_n\}$,

$$\begin{aligned} V(\pi, f) &= \min_{p\in\Delta} \sum p(E_i) \left[\min_{p_i\in\Delta(E_i)} \int u(f)dp_i + c_{E_i}(p_i)\right] + \min_{\{q\in\Delta(S):q=p \text{ on } \pi\}} c(q) \\ &\leq \sum p^*(E_i) \left[\int u(f)dp^*(\cdot|E_i) + c_{E_i}(p^*(\cdot|E_i))\right] + \min_{\{q\in\Delta(S):q=p^* \text{ on } \pi\}} c(q^*) \\ &= \int_S u(f)dp^* \\ &= \int_S u(f)dp^* + c(p^*) = V(\pi_0, f) \end{aligned}$$

The second equality follows from

$$c_{E_i}(p^*(\cdot|E_i)) = \min_{p(\cdot|E_i)=p^*(\cdot|E_i)} \frac{c(p)}{p(E_i)} = 0$$

and

$$\min_{\{q\in\Delta(S):q=p^* \text{ on } \pi\}} c(q^*) = 0.$$

□

## B.12 Proposition 7

*Proof.* Suppose $f$ can be locally approximated by an SEU preference $\geq'$ that is less ambiguity averse than $\succcurlyeq_0$. Let $\geq'$ be represented by $U'$ with risk utility $u'$ and belief $q \in \Delta(S)$. Since $\geq'$ is less ambiguity averse than $\succcurlyeq_0$, we can normalize $u'$ so that $u = u'$. In addition, $q \in c^{-1}(0)$ by Maccheroni et al. (2006a) Lemma 32. Since $V(f) = U'(f)$,

$$V(f) = \min_{p\in\Delta}[\int_S u(f)dp + c(p)] = U'(f) = \int_S u(f)dq = \int_S u(f)dq + c(q)$$

The last equality follows from the fact that $q \in c^{-1}(0)$. So $q \in \arg\min_{p\in\Delta}[\int_S u(f)dp + c(p)]$ by definition. Together with $q \in c^{-1}(0)$, this implies that

$$c^{-1}(0) \cap \arg\min_{p\in\Delta}[\int_S u(f)dp + c(p)] \neq \emptyset$$

Thus condition (12) holds at $f$.

Now suppose there exists some $p^* \in c^{-1}(0) \cap \arg\min_{p\in\Delta}[\int_S u(f)dp + c(p)]$. Define $U'$ by $U'(f) = \int_S u(f)dp^*$. Then by definition $U'$ represents an SEU preference $\geq'$ that is less ambiguity averse than $\succcurlyeq_0$. Also

$$V(f) = \int_S u(f)dp^* + c(p^*) = \int_S u(f)dp^* = U'(f)$$

So $f$ can be locally approximated by an SEU preference that is less ambiguity averse than $\succsim_0$. $\square$

## B.13 Proposition 8

*Proof.* By Proposition 7, $f$ can be locally approximated by an SEU preference $\geq'$ that is less ambiguity averse than $\succsim_0^1$ if and only if condition (12) holds. Then there exists $p^* \in c_1^{-1}(0) \cap argmin_{p\in\Delta}[\int_S u_1(f)dp + c_1(p)]$ such that $V_1(f) = \int_S u_1(f)dp^* + c_1(p^*)$, and $c_1(p^*) = 0$. By definition, $\succsim_0^2$ is less ambiguity averse than $\succsim_0^1$ if and only if $u_1 = u_2$ and $c_2 \geq c_1$. Since $\succsim_0^2$ is more ambiguity averse than $\geq'$, $u_2 = u'$ and $p^* \in c_2^{-1}(0)$. Let $u = u_1 = u_2 = u'$. Therefore:

$$\int_S u(f)dp^* + c_2(p^*) = \int_S u(f)dp^* + c_1(p^*) \leq \int_S u(f)dp + c_1(p) \leq \int_S u(f)dp + c_2(p), \forall p \in \Delta(S)$$

The first inequality follows from the fact that $p^* \in \arg\min_{p\in\Delta}[\int_S u_1(f)dp + c_1(p)]$, and the second from $c_1 \leq c_2$. Thus $p^* \in \arg\min_{p\in\Delta}[\int_S u(f)dp + c_2(p)]$. So

$$\arg\min_{p\in\Delta}[\int_S u(f)dp + c_2(p)] \cap c_2^{-1}(0) \neq \emptyset$$

and by Proposition 6, $\succsim^2$ exhibits aversion to partial information at $f$. $\square$

## B.14 Lemma 1

*Proof.* We first show $\succsim^+$ is represented by $\tilde{V}(\pi, F) = \max_{f\in F^\pi} V(\pi, f)$. For all $(\pi, F)$ and $(\pi', G)$,

$$(\pi, F) \succsim^+ (\pi', G) \iff \forall g \in G^{\pi'}, \exists f \in F^\pi, (\pi, f) \succsim (\pi', g)$$

Since $V : \Pi \times \mathcal{F}$ represents $\succsim$, this is equivalent to

$$\max_{f\in F^\pi} V(\pi, f) \geq \max_{g\in G^{\pi'}} V(\pi', g)$$

Thus $(\pi, F) \succsim^+ (\pi', G)$ if and only if $\tilde{V}(\pi, F) \geq \tilde{V}(\pi', G)$.

Furthermore, by definition,

$$F^\pi = \left\{ \begin{pmatrix} f_1 & \text{if } s \in E_1 \\ \vdots & \vdots \\ f_n & \text{if } s \in E_n \end{pmatrix} : f_i \in F, \forall i = 1, \cdots, n \right\}.$$

$$\begin{aligned}
\max_{f\in F^\pi} V(\pi,f) &= \max_{f_1\in F}\ldots\max_{f_n\in F} V\left(\pi, \begin{pmatrix} f_1 & \text{if } s\in E_1 \\ \vdots & \vdots \\ f_n & \text{if } s\in E_n \end{pmatrix}\right) \\
&= \max_{f_1\in F}\ldots\max_{f_n\in F} I_0 \begin{pmatrix} V_{E_1}(f_1) & \text{if } s\in E_1 \\ \vdots & \vdots \\ V_{E_n}(f_n) & \text{if } s\in E_n \end{pmatrix} \quad \text{by Theorem 1} \\
&= I_0 \begin{pmatrix} \max_{f\in F} V_{E_1}(f) & \text{if } s\in E_1 \\ \vdots & \vdots \\ \max_{f\in F} V_{E_n}(f) & \text{if } s\in E_n \end{pmatrix} \quad \text{by } I_0\text{-monotonicity}
\end{aligned}$$

□

## B.15 Proposition 9, 10, and 11

*Proof of Proposition 9.* (2) ⇔ (3) is due to Theorem 2.

To show (1) ⇒ (2), take any singleton menu $F = \{f\}$. A preference for perfect information implies $(\pi^*, f) \succcurlyeq (\pi, f), \forall\pi$. By Time Neutrality, $(\pi^*, f) \sim (\pi_0, f)$, so $(\pi_0, f) \succcurlyeq (\pi, f)$.

To show (2) ⇒ (1). Let $\pi \in \Pi$ and $F \in \mathcal{M}$. Then

$$V(\pi^*, F) - V(\pi, F) = [\max_{f\in F^{\pi^*}} V(\pi^*, f) - \max_{f\in F^\pi} V(\pi^*, f)] + [\max_{f\in F^\pi} V(\pi^*, f) - \max_{f\in F^\pi} V(\pi, f)]$$

The first term is non-negative since $F^\pi \subseteq F^{\pi^*}$. By (2) and Time Neutrality, $V(\pi^*, f) = V(\pi_0, f) \geq V(\pi, f)$, for all $\pi, f$. So

$$max_{f\in F^\pi} V(\pi, f) = V(\pi, f^*) \leq V(\pi^*, f^*) \leq \max_{f\in F^\pi} V(\pi^*, f)$$

where $f^* \in F^\pi$ is the act that maximizes $V(\pi, \cdot)$. So the second term is also non-negative. Thus $V(\pi^*, F) \geq V(\pi, F)$ and the DM has preferences for perfect information. □

*Proof of Proposition 10.* The proofs of (1) and (2) are straightforward and thus omitted.

We show (3) via a counterexample. Suppose $S = \{s_1, s_2, s_3\}$, and $\succcurlyeq_0$ has a MEU representation $(u, \mathcal{P})$ where $\mathcal{P} = \{p \in \Delta^3 | p(s_1) = \frac{1}{3}, p(s_3) \in [\frac{1}{6}, \frac{1}{2}]\}$. For simplicity assume risk neutrality, so $u(x) = x$. Suppose the DM faces menu $F = \{(0, 1, 1), (0.49, 0.49, 0.49)\}$. Then $V(\pi_0, F) = \frac{2}{3}$. Let $\pi = \{\{s_1, s_2\}, \{s_3\}\} > \pi_0$. The informed DM will choose $(0.49, 0.49, 0.49)$

given $\{s_1, s_2\}$, and $(0,1,1)$ given $\{s_3\}$. Therefore $V(\pi, F) = 0.575 < \frac{2}{3} = V(\pi_0, F)$.[52] Information hurts. $\square$

Next we prove Proposition 11. We first prove a lemma. Let $F_0 = \arg\max_{f\in F} V(\pi_0, f)$ be the set of uninformed optimal acts. By our decomposition, as long as the DM is not strictly averse to information $\pi$ at some $f_0 \in F_0$, then information is valuable.

Let $F_i^* = \arg\max_{f\in F} V_{E_i}(f)$ be the set of optimal acts in $F$ conditional on learning about $E_i$. Consider $F^* = \{f_1^* E_1 f_2^* E_2 \cdots E_{n-1} f_n^* : f_i^* \in F_i^*, \forall i\} \subseteq F^\pi$. The instrumental value of information is zero if and only if $F^* \cap F \neq \emptyset$. We collect these observations below.

**Lemma 9.** *1. If there exists an unconditional optimal act $f_0 \in F_0$ such that $V(\pi, f_0) \geq V(\pi_0, f_0)$ at $f_0$, then $V(\pi, F) - V(\pi_0, F) \geq 0$.*

*2. If there exists a conditional optimal strategy $f^* \in F^*$ such that $f^* \in F$ and $V(\pi, f^*) \leq (<) V(\pi_0, f^*)$, then $V(\pi, F) - V(\pi_0, F) \leq (<) 0$.*

*Proof.* By definition $V(\pi_0, f_0) = \max_{f\in F} V(\pi_0, f)$. If $V(\pi, f_0) \geq V(\pi_0, f_0)$, then the intrinsic value of information $\pi$ at menu $F$ is non-negative:

$$\max_{f\in F} V(\pi, f) - \max_{f\in F} V(\pi_0, f) \geq V(\pi, f_0) - V(\pi_0, f_0) \geq 0.$$

As the instrumental value of information is always non-negative, $V(\pi, F) - V(\pi_0, F) \geq 0$ and $\pi$ is valuable.

If there exists $f^* \in F \cap F^*$, then the instrumental value of $\pi$, $V(\pi, f^*) - \max_{f\in F} V(\pi, f) = 0$. In addition $\max_{f\in F} V(\pi, f) = V(\pi, f^*) \leq V(\pi_0, f^*) \leq \max_{f\in F} V(\pi_0, f)$, so the intrinsic value of $\pi$ is non-positive. $\square$

*Remark* 8. The first condition is helpful, as it requires only calculation of an optimal act in the uninformed case. This could simplify checking whether ambiguity aversion generates information aversion or not. In MEU models, this is equivalent to $V(\pi, f_0) = V(\pi_0, f_0)$, when the intrinsic value of information $\pi$ for menu $F$ vanishes.

*Proof of Proposition 11.* If there exists an uninformed optimal act $f_0$ that is $\pi$-measurable, then $V(\pi, f_0) = V(\pi^*, f_0) = V(\pi_0, f_0)$. By the above lemma, $\Delta V(\pi, F) \geq 0$ . $\square$

[52] $\mathcal{P}(s_2|\{s_1, s_2\}) = [\frac{1}{3}, \frac{3}{5}]$, so $(0.49, 0.49, 0.49) \succ_{\{s_1,s_2\}} (0, 1, 1)$.

*Proof of Corollary 4.* Let $x$ be the uninformed optimal act for DM 1. So $V^1(\pi_0, x) \geq V^1(\pi_0, f)$, for all $f$ in menu $F$. Since DM 2 is more ambiguity averse than DM 1, $u_2 = u_1$ and $c_2 \leq c_1$. So for all $f \in \mathcal{F}$,

$$V^2(\pi_0, f) = \min_{p \in \Delta(S)} \int_S u(f)dp + c_2(p) \leq \min_{p \in \Delta(S)} \int_S u(f)dp + c_1(p) = V^1(\pi_0, f)$$

and $V^1(\pi_0, x) = u(x) = V^2(\pi_0, x)$. Thus $V^2(\pi_0, x) \geq V^2(\pi_0, f)$ for all $f \in F$. Since $x$ is $\pi$-measurable, by Proposition 11 we have $\Delta V^2(\pi, F) \geq 0$. $\square$

**Marginal Value of Information**

For any menu $F$, consider two partitions $\pi_2 \geq \pi_1$. The marginal value of getting the finer information $\pi_2$ is:

$$V(\pi_2, F) - V(\pi_1, F) = [\max_{f \in F^{\pi_2}} V(\pi_2, f) - \max_{f \in F^{\pi_1}} V(\pi_2, f)] + [\max_{f \in F^{\pi_1}} V(\pi_2, f) - \max_{f \in F^{\pi_1}} V(\pi_1, f)]$$

The first term captures the instrumental value of getting finer information $\pi_2$ relative to $\pi_1$, and since $F^{\pi_1} \subseteq F^{\pi_2}$ this term is non-negative. The second part captures the intrinsic value of information $\pi_2$ relative to $\pi_1$.

Lemma 9 can be generalized as follows.

**Lemma 10.** *1. If there exists an optimal strategy $f^{*1}$ for decision problem $(\pi_1, F)$ such that $V(\pi_1, f^{*1}) \leq V(\pi_2, f^{*1})$, then $V(\pi_2, F) - V(\pi_1, F) \geq 0$.*

*2. If there exists an optimal strategy $f^{*2}$ for decision problem $(\pi_2, F)$ such that $f^{*2} \in F^{\pi_1}$ and $V(\pi_1, f^{*2}) \geq V(\pi_2, f^{*2})$, then $V(\pi_2, F) - V(\pi_1, F) \leq 0$.*

The proof is similar to the proof of Lemma 9 and hence omitted.

## B.16 Proposition 12

*Proof.* Fix $\pi = \{E_1, \cdots, E_n\}$. Suppose $\succcurlyeq_0$ has second-order belief representation $(u, \phi; \Theta, \mu)$ and $\succcurlyeq_0$ is ambiguity averse. Then by Klibanoff et al. (2005) Proposition 1, $\phi$ is concave. Let

$f$ be an act where $\succcurlyeq_0$ displays local ambiguity neutrality. Then

$$\begin{aligned}
V(\pi, f) &= \int_\Theta \phi[\sum_{i=1}^n p_{\theta'}(E_i)\phi^{-1}[\int_\Theta \phi(\int u(f)dp_{\theta_i}(\cdot|E_i))d\mu_{E_i}(\theta_i)]]d\mu(\theta') \\
&\leq \int_\Theta \phi[\sum_{i=1}^n p_{\theta'}(E_i)[\int_\Theta \int u(f)dp_{\theta_i}(\cdot|E_i)d\mu_{E_i}(\theta_i)]]d\mu(\theta') \\
&= \int_\Theta \phi[\sum_{i=1}^n p_{\theta'}(E_i)(\int_\Theta \int_{E_i} u(f)dp_{\theta_i}\frac{d\mu(\theta_i)}{\int p_{\theta''}(E_i)d\mu(\theta'')}]d\mu(\theta') \\
&= \int_\Theta \phi[\sum_{i=1}^n (\int_\Theta \int_{E_i} u(f)dp_{\theta_i}d\mu(\theta_i)\frac{p_{\theta'}(E_i)}{\int p_{\theta''}(E_i)d\mu(\theta'')}d\mu(\theta')] \\
&\leq \phi \int_\Theta [\sum_{i=1}^n (\int_\Theta \int_{E_i} u(f)dp_{\theta_i}d\mu(\theta_i)\frac{p_{\theta'}(E_i)}{\int p_{\theta''}(E_i)d\mu(\theta'')}d\mu(\theta')] \\
&= \phi[(\sum_{i=1}^n \int_\Theta \int_{E_i} u(f)dp_{\theta_i}d\mu(\theta_i)(\int_\Theta \frac{p_{\theta'}(E_i)}{\int p_{\theta''}(E_i)d\mu(\theta'')}d\mu(\theta'))] \\
&= \phi(\int_\Theta \int_S u(f)dp_\theta d\mu) = V(\pi_0, f)
\end{aligned}$$

The two inequalities follow from the concavity of $\phi$. The last equality holds because $\succcurlyeq_0$ displays local ambiguity neutrality at $f$.

The case for ambiguity loving $\succcurlyeq_0$ can be proved analogously. □

# References


Acciaio, B., H. Follmer, and I. Penner (2012). Risk assessment for uncertain cash flows: Model ambiguity, discounting ambiguity, and the role of bubbles. *Finance and Stochastics 16*(4), 669–709.

Amarante, M. and E. Filiz (2007). Ambiguous events and maxmin expected utility. *Journal of Economic Theory 134*(1), 1–33.

Anscombe, F. J. and R. J. Aumann (1963). A definition of subjective probability. *The Annals of Mathematical Statistics 34*(1), pp. 199–205.

Backus, D. K., B. R. Routledge, and S. E. Zin (2005). *Exotic Preferences for Macroeconomists*, Volume 19 of *NBER Macroeconomics Annual 2004*, Chapter 5, pp. 319–414. MIT Press.

Bellemare, C., M. Krause, S. Kroger, and C. Zhang (2005). Myopic loss aversion: Information feedback vs. investment flexibility. *Economics Letters 87*(3), 319–324.

Blackwell, D. (1951). The comparison of experiments. In *Proceedings, Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 93102. Univ. of California Press, Berkeley.

Blackwell, D. (1953). Equivalent comparisons of experiments. *The Annals of Mathematical Statistics 24*(2), 265–272.

Cao, H., T. Wang, and H. Zhang (2005). Model uncertainty, limited market participation, and asset prices. *Review of Financial Studies 18*(4), 1219–1251.

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2011). Uncertainty averse preferences. *Journal of Economic Theory 146*(4), 1275–1330.

Dillenberger, D. (2010). Preferences for one-shot resolution of uncertainty and Allais-type behavior. *Econometrica 78*, 1973–2004.

Dow, J. and S. R. d. C. Werlang (1992). Uncertainty aversion, risk aversion, and the optimal choice of portfolio. *Econometrica 60*(1), 197–204.

Eichberger, J., S. Grant, and D. Kelsey (2007). Updating choquet beliefs. *Journal of Mathematical Economics 43*(7–8), 888–899.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics 75*(4), 643–669.

Epstein, L. and M. Schneider (2003). Recursive multiple-priors. *Journal of Economic Theory 113*(1), 1–31.

Epstein, L. and T. Wang (1994). Intertemporal asset pricing under knightian uncertainty. *Econometrica 64*(2), 283–322.

Epstein, L. G. (1999). A definition of uncertainty aversion. *Review of Economic Studies 66*, 579–608.

Epstein, L. G. and M. LeBreton (1993). Dynamically consistent beliefs must be Bayesian. *Journal of Economic Theory 61*(1), 1–22.

Epstein, L. G. and M. Schneider (2010). Ambiguity and asset markets. *Annual Review of Financial Economics 2*(1), 315–346.

Epstein, L. G. and J. Zhang (2001). Subjective probabilities on subjectively unambiguous events. *Econometrica 69*(2), 265–306.

Epstein, L. G. and S. E. Zin (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica 57*(4), 937–969.

Epstein, L. G. and S. E. Zin (1991). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *The Journal of Political Economy 99*(2), 263–286.

Gajdos, T., T. Hayashi, J.-M. Tallon, and J.-C. Vergnaud (2008). Attitude toward imprecise information. *Journal of Economic Theory 140*(1), 27–65.

Ghirardato, P., F. Maccheroni, and M. Marinacci (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory 118*(2), 133–173.

Ghirardato, P. and M. Marinacci (2002). Ambiguity made precise: A comparative foundation. *Journal of Economic Theory 102*(2), 251–289.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*(2), 141–153.

Gneezy, U., A. Kapteyn, and J. Potters (2003). Evaluation periods and asset prices in a market experiment. *The Journal of Finance 52*(2), 821–837.

Grant, S., A. Kajii, and B. Polak (1998). Intrinsic preference for information. *Journal of Economic Theory 83*(2), 233–259.

Gumen, A. and A. Savochkin (2013). Dynamically stable preferences. *Journal of Economic Theory 148*(4), 1487–1508.

Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica 75*(2), 503–536.

Hanany, E. and P. Klibanoff (2007). Updating preferences with multiple priors. *Theoretical Economics 2*, 261–298.

Hanany, E. and P. Klibanoff (2009). Updating ambiguity averse preferences. *The B.E. Journal of Theoretical Economics 9*.

Hansen, L. P. and T. J. Sargent (2001). Robust control and model uncertainty. *The American Economic Review , Papers and Proceedings 91*(2), 60–66.

Hansen, L. P. and T. J. Sargent (2007). *Robustness*. Princeton University Press.

Herstein, I. and J. Milnor (1953). An axiomatic approach to measurable utility. *Econometrica 21*(2), 291–297.

Hoy, M., R. Peter, and A. Richter (2014). Take-up for genetic tests and ambiguity. *Journal of Risk and Uncertainty 48*(2), 111–133.

Jaffray, J.-Y. (1994). Dynamic decision making with belief functions. In R. R. Yager, J. Kacprzyk, and M. Fedrizzi (Eds.), *Advances in the Dempster-Shafer Theory of Evidence*, pp. 331–352. New York, NY, USA: John Wiley & Sons, Inc.

Ju, N. and J. Miao (2012). Ambiguity, learning, and asset returns. *Econometrica 80*(2), 559591.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005). A smooth model of decision making under ambiguity. *Econometrica 73*(6), 1849–1892.

Klibanoff, P., M. Marinacci, and S. Mukerji (2008). Recursive smooth ambiguity preferences. *Journal of Economic Theory 144*, 930–976.

Kopylov, I. (2007). Subjective probabilities on small domains. *Journal of Economic Theory 133*(1), 236 – 265.

Koszegi, B. (2003). Health anxiety and patient behavior. *Journal of Health Economics 22*(6), 1073–1084.

Koszegi, B. and M. Rabin (2009). Reference-dependent consumption plans. *American Economic Review 99*, 909–936.

Kreps, D. M. and E. L. Porteus (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica 46*(1), 185–200.

Lerman, C., C. Hughes, B. J. Trock, R. E. Myers, D. Main, A. Bonney, M. R. Abbaszadegan, A. E. Harty, B. A. Franklin, J. F. Lynch, et al. (1999). Genetic testing in families with hereditary nonpolyposis colon cancer. *Journal of the American Medical Association 281*(17), 1618–1622.

Lerman, C., J. Marshall, J. Audrain, and A. Gomez-Caminero (1996). Genetic testing for colon cancer susceptibility: Anticipated reactions of patients and challenges to providers. *International Journal of Cancer 69*(1), 58–61.

Lerman, C., S. Narod, K. Schulman, C. Hughes, A. Gomez-Caminero, G. Bonney, K. Gold, B. Trock, D. Main, and J. Lynch (1996). BRCA1 testing in families with hereditary breast-ovarian cancer: a prospective study of patient decision making and outcomes. *Journal of the American Medical Association 275*(24), 1885–1892.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006a). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica 74*(6), 1447–1498.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006b). Dynamic variational preferences. *Journal of Economic Theory 128*(1), 4–44.

Machina and Schmeidler (1992). A more robust definition of subjective probability. *Econometrica 60*(4), 745–780.

Machina, M. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature 27*(4), 1622–1668.

Machina, M. J. and D. Schmeidler (1995). Bayes without Bernoulli: Simple conditions for probabilistically sophisticated choice. *Journal of Economic Theory 67*(1), 106–128.

Nau, R. F. (2006). Uncertainty aversion with second-order utilities and probabilities. *Management Science 52*(1), 136–145.

Pires, C. P. (2002). A rule for updating ambiguous beliefs. *Theory and Decision 53*(2), 137–152.

Raiffa, H. (1961). Risk, ambiguity, and the savage axioms: Comment. *The Quarterly Journal of Economics 75*(4), 690–694.

Riedel, F. (2004). Dynamic coherent risk measures. *Stochastic Processes and their Applications 112*(2), 185–200.

Savage, L. J. (1954). *The Foundations of Statistics*. Wiley, New York.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica 57*(3), 571–587.

Segal, U. (1990). Two-stage lotteries without the reduction axiom. *Econometrica 58*(2), 349–377.

Seo, K. (2009). Ambiguity and second-order belief. *Econometrica 77*(5), 1575–1605.

Siniscalchi, M. (2011). Dynamic choice under ambiguity. *Theoretical Economics 6*(3), 379–421.

Skiadas, C. (1998). Recursive utility and preferences for information. *Economic Theory 12*(2), 293–312.

Strzalecki, T. (2011). Axiomatic foundations of multiplier preferences. *Econometrica 79*(1), 47–73.

Strzalecki, T. (2013). Temporal resolution of uncertainty and recursive models of ambiguity aversion. *Econometrica 81*(3), 1039–1074.

Zhang, J. (2002). Subjective ambiguity, expected utility and choquet expected utility. *Economic Theory 20*(1), 159–181.