Additive-Belief-Based Preferences*

David Dillenberger† Collin Raymond‡

July 13, 2020

Abstract

We introduce a new class of preferences — which we call additive-belief-based (ABB) — that captures a general and yet tractable approach to belief-based utility, and that encompasses many popular models in the behavioral literature. We show that the general class of ABB preferences and two prominent special cases, which allow utility to depend on each period's beliefs but not on changes in beliefs across periods, are fully characterized by suitable relaxations of the standard Independence axiom. We identify the intersection of ABB preferences with the class of recursive preferences and characterize attitudes towards the timing of resolution of uncertainty for ABB preferences. Our approach helps to better understand, in terms of testable predictions, existing models and leads to new models that can accommodate previously uncaptured behavioral patterns.

JEL codes: D80, D81, D83, D91.

Key words: Anticipatory utility, Compound lotteries, Preferences over beliefs, Recursive preferences, Resolution of uncertainty.

1 Introduction

1.1 Motivation

It is both intuitive and well documented that beliefs about future consumption or life events directly affect well-being. For example, an individual may enjoy looking forward to an upcoming vacation and particularly so if the risks of severe weather conditions are deemed very low; on the other hand, the same individual may worry about a future medical procedure. Using survey techniques, Loewenstein (1987) provides early evidence on the effects of these anticipatory motives in economics.

*We thank Andrew Caplin, Eddie Dekel, Kfir Eliaz, Daniel Gottlieb, Faruk Gul, Botond Kőszegi, Muriel Niederle, Kirby Nielsen, Pietro Ortoleva, Wolfgang Pesendorfer, Andrew Postlewaite, Hendrik Rommeswinkel, Philipp Sadowski, Todd Sarver, Ran Spiegler, and Tomasz Strzalecki for useful comments and suggestions. We gratefully acknowledge the financial support of NSF under awards SES 1758653 and SES 1758775.

†Department of Economics, University of Pennsylvania (ddill@sas.upenn.edu)

‡Department of Economics, Purdue University (collinbraymond@gmail.com)

There is also widespread evidence from other fields discussing how anticipation of pain produces psychological-stress reactions: a notable example is Lazarus (1966), while Berns et al. (2006) provide evidence from fMRI studies. Moreover, a desire to avoid anxiety or other negative emotions has been suggested as the primary motivation for why individuals avoid information even when it is rational to acquire it (e.g. Oster et al., 2013).

As a result, decision-making models in which individuals derive utility not only from material outcomes but also from their current and future beliefs have become increasingly prominent. Existing models usually take one of two forms. The first, as in Caplin and Leahy (2001), allows utility to depend on the (absolute) level of beliefs; i.e., on how likely it is that certain states/payoffs occur. In line with previous literature, we refer to this set of models as *anticipatory utility* models. In the second, as in Kőszegi and Rabin (2009), utility depends not on the level of beliefs, but on changes in beliefs across periods. We refer to this class as *changing beliefs* models. Both classes have proved useful in explaining intuitive behavioral phenomena that are hard to reconcile with the standard "consequentialist" model, such as asset and portfolio puzzles (Pagel, 2018), patterns of over-consumption in the face of income shocks (Kőszegi and Rabin, 2009), and selective avoidance of medical information (Caplin and Eliaz, 2003; Kőszegi, 2003).

Existing research tends to make specific functional form assumptions about the way utility depends on beliefs. However, despite their potential explanatory power, little work has been done on trying to understand what types of behavior belief-based utility, as a general class, can potentially accommodate, or must rule out. There are several potential downsides to the current approach. For example, it may falsely reject the hypothesis that beliefs matter merely because the data does not support specific functional form assumptions. It may also obscure the main features of belief-based utility that drive particular behavioral patterns (mistakenly attributing them instead to a specific model). Moreover, although both anticipatory and changing beliefs models are motivated by many of the same stylized facts, there is little understanding of what distinctive behavioral predictions they may make. It is conceivable that some of the existing models generate the same behavior despite having distinct functional forms.

In this paper we address these issues. We suggest a unified framework for the type of models described above and provide their testable implications. In particular, we introduce and analyze a new class of utility functions — which we call *additive-belief-based (ABB)* utility — that captures a general approach to belief-based utility and that encompasses both the anticipatory and changing beliefs classes. We also point out a useful partition of the class of anticipatory utility models into (i) *prior-anticipatory utility* models, where utility depends on beliefs at the beginning of the time period, before information has been received; and (ii) *posterior-anticipatory utility* models, where utility depends on beliefs at the end of the time period, after information has been received. A comprehensive review of models that fall into these categories is in Section 5. Additionally, Table

1 in that section summarizes the taxonomy of popular models that are nested by ABB preferences. We show that both the general class and the special subsets we consider are characterized by new, intuitive relaxations of the familiar Independence axiom of expected utility, applied to our setting. We also demonstrate how functional forms restrictions translate to non-standard behaviors, such as intrinsic (i.e., non-instrumental) attitudes towards information, and, conversely, how such behaviors allow us to distinguish between different cases of ABB utility. For example, prior-anticipatory utility models, previously unused in the literature, are able to accommodate a richer set of behavioral phenomena compared to the oft-used posterior-anticipatory utility models.

Our results can thus guide applied theorists, by providing a toolkit to understanding the behavioral implications of particular modeling choices, as well as empirical researchers, in testing — and distinguishing between — potential belief-based mechanisms.

1.2 Preview of Results

As will be described below, individuals in our framework are assumed to gain utility from their beliefs, but they cannot directly choose them.[1] Rather, they hold prior beliefs, receive information, and form posterior beliefs by applying Bayes' rule. Therefore, individuals can control their beliefs only by choosing particular information structures. To link this to observable behavior, we look at preferences over the combination of information structures and prior beliefs. These can naturally be elicited in experimental and field settings. Formally, taking advantage of the mapping between information structures and compound lotteries, we take as our domain of preferences the set of two-stage compound lotteries, that is, lotteries whose prizes are different lotteries over final outcomes.[2] The restriction to two, rather than to arbitrary T, stages is mostly to ease exposition and simplify notation. In Appendix A.4 we comment on how our analysis can be extended to that richer domain.

Let P be a typical two-stage lottery. In period 0 it induces prior beliefs $\phi(P)$ over final outcomes; the individual knows that the overall probability to receive x_j in period 2 is $\phi(P)(x_j)$. In period 1, P generates a signal i with probability $P(p_i)$. Signal i generates posterior beliefs over outcomes; the individual now knows that in period 2 he will receive x_j with probability $p_i(x_j)$. In period 2, all uncertainty resolves and the individual receives x_j and has degenerate beliefs on this outcome

[1] A well known example in which individuals optimally choose their beliefs is Brunnermeier and Parker (2005). Bénabou and Tirole (2016) and Bénabou (2015) survey models where individuals can distort their beliefs.

[2] In particular, we assume that individuals cannot take intermediate actions that may affect their final payoffs. While many of the models we refer to do allow for such actions, we will omit them from our analysis for several reasons. First, for expository purposes, we focus on the simplest model that still incorporates all variables of interest into the utility function. Second, as we discuss in Section 4, individuals in our model have *intrinsic* preferences over information. It will be cleanest to characterize such preferences within a framework that rules out any *instrumental* value of information. Third, allowing for intermediate actions requires us to take a stand about the solution concept that governs how individuals determine which action to take (e.g., backward induction or personal equilibrium), rather than solely focus on the parameters that identify their preferences. The choice of solution concept also requires carefully modeling the timing of actions relative to when beliefs, and belief-based utility, are realized.

(denoted δ_{x_j}). From the ex ante point of view, the total utility of this scenario is:

$$
\begin{aligned}
V_{ABB}(P) \quad = & \quad \underbrace{\sum_j \phi(P)(x_j)u(x_j)}_{\text{expected utility from material payoffs}} \\
+ & \quad \underbrace{\sum_i P(p_i)\nu_1(\phi(P), p_i)}_{\text{expected utility from beliefs in period 1}} \\
+ & \quad \underbrace{\sum_i P(p_i) \sum_j p_i(x_j)\nu_2(p_i, \delta_{x_j})}_{\text{expected utility from beliefs in period 2}}
\end{aligned}
$$

The first term represents the expected consumption utility — the expected utility that the individual receives from material outcomes (in period 2). The second term represents period 1's belief-based utility — the expected utility from having interim beliefs p_i in period 1, conditional on having prior beliefs $\phi(P)$. The last term represents period 2's belief-based utility — the individual's expected utility from x_j being realized in period 2, conditional on having interim beliefs p_i.

As a concrete example, suppose there are two outcomes, H (high) and L (low), so that beliefs are summarized by the probability of H. Suppose that ex-ante the two outcomes are equally likely. Then the expected utility over material outcomes is $\frac{1}{2}u(H) + \frac{1}{2}u(L)$. In period 1, the individual receives a binary signal: half the time it is good, and beliefs move to $\frac{3}{4}$ with corresponding utility $\nu_1(\frac{1}{2}, \frac{3}{4})$; half the time it is bad, and beliefs fall to $\frac{1}{4}$ with corresponding utility $\nu_1(\frac{1}{2}, \frac{1}{4})$. Expected belief-based utility in period 1 is thus $\frac{1}{2}\nu_1(\frac{1}{2}, \frac{3}{4}) + \frac{1}{2}\nu_1(\frac{1}{2}, \frac{1}{4})$. In period 2, the individual learns whether he got H or L. After the good signal utility is either $\nu_2(\frac{3}{4}, 1)$ or $\nu_2(\frac{3}{4}, 0)$; and after the bad signal it is either $\nu_2(\frac{1}{4}, 1)$ or $\nu_2(\frac{1}{4}, 0)$. Expected belief-based utility in period 2 is $\frac{1}{2}\left(\frac{3}{4}\nu_2(\frac{3}{4}, 1) + \frac{1}{4}\nu_2(\frac{3}{4}, 0)\right) + \frac{1}{2}\left(\frac{1}{4}\nu_2(\frac{1}{4}, 1) + \frac{3}{4}\nu_2(\frac{1}{4}, 0)\right)$. ABB utility is the sum of these three components.

ABB utility is general enough to nest many important specifications, yet specific enough to highlight the direct effect of beliefs on flow utility. For example, in Kőszegi and Rabin (2009) ν_1 and ν_2 (whose exact forms are given in Section 6.1) are the gain-loss utilities from the belief change from $\phi(P)$ to p_i and from p_i to δ_{xj}, respectively. In Caplin and Leahy's (2001) model of anticipatory anxiety, both ν_1 and ν_2 depend only on their second argument.

In Section 2, we initially show how the prominent sub-classes of ABB functionals are related to each other, demonstrating that models that allow for changing beliefs nest those of prior anticipatory beliefs, which in turn nest those of posterior anticipatory beliefs. We then provide necessary and sufficient conditions for continuous preferences to be represented with an ABB functional. First, Prior Conditional Two-Stage Independence (PTI) requires standard Independence (in mixing compound lotteries) to hold only if *all* compound lotteries involved in the mixing induce the same prior distribution over final outcomes. That is, if $\phi(P) = \phi(Q) = \phi(R)$, then P is preferred to Q if and only if the mixture of P and R is preferred to the (same-proportion) mixture of Q and R. Second, Cross Sectional Two-Stage Independence (CTI) requires consistency/uniformity across priors in the scale used to measure preferences.

We next turn to showing how, in addition to CTI, strengthening PTI allows us to characterize models of prior-anticipatory beliefs. The key behavior that distinguishes utility from changes in beliefs and utility from the level of beliefs is how broadly Independence (again over compound lotteries) holds. If individuals only care about the levels of their beliefs, then Independence should hold whenever the two lotteries involved in the initial comparison, but not necessarily the one they are both mixed with, induce the same prior distribution over outcomes. That is, Strong Prior Conditional Two-Stage Independence (SPTI) drops from PTI the requirement that $\phi(R)$ agrees with $\phi(P) = \phi(Q)$. We last show that imposing Independence on all mixtures in the first stage (a property that subsumes all requirements above) characterizes posterior-anticipatory beliefs. Thus, our results demonstrate how a simple set of easily tested behavioral postulates allows distinguishing between different types of belief-dependent utility.

ABB preferences are not the only class of preferences that have been developed to explain informational preferences, even in the absence of the ability to condition actions on that information. A different vein of the literature, initially developed by Kreps and Porteus (1978) and extended by Segal (1990), focuses on recursive preferences over compound lotteries (and information). These models rely on a relaxation of the Independence axiom that is orthogonal to both PTI and CTI. In Section 3 we show that, in the context of two-stage compound lotteries, the intersection of the two models is precisely the class of preferences that admit a posterior-anticipatory beliefs representation.

Since individuals gain utility from their beliefs, they may exhibit non-degenerate preferences over information structures even if they cannot react to new information by altering their behavior, that is, even if they do not have actions to take in the interim stage. This feature, which is supported by a large body of experimental evidence (discussed in Section 5), is what distinguishes these models from the standard model, in which individuals would be indifferent to information when no actions are available. In Section 4 we investigate what restrictions on the functional forms are equivalent to well-known types of intrinsic informational preferences, such as preferences for early resolution of uncertainty (Kreps and Porteus, 1978) or preferences for one-shot resolution of uncertainty (Dillenberger, 2010). In doing so, we provide characterizations that generalize some earlier results which were made in the context of specific models. Our results allow us to determine how different classes of models can, or cannot, accommodate different intrinsic attitudes towards information. For example, a preference for early resolution of uncertainty is consistent with all the sub classes we consider. In contrast, we show that a preference for one-shot resolution of uncertainty cannot be exhibited by individuals who have only anticipatory motives, but rather requires a concern for changing beliefs.

In Section 5 we discuss how our three categories of utility functionals relate to specific models used in the literature. Finally, while this paper focuses on deriving results that are not tied to specific functional forms, in Section 6 we demonstrate how our results can enrich our understating of known models and extend the explanatory power of belief-based preferences. For example, we both extend and point out the limits of the relationship between loss aversion and information preferences within the model of Kőszegi and Rabin (2009). We also show how prior-anticipatory utility models, previously unused in the literature, are able to accommodate a richer set of behavioral phenomena, such as the joint presence of well documented violations of expected utility and preferences for more or less information, compared to the oft-used posterior-anticipatory utility models.

2 The Model

2.1 Preliminaries

Consider a set of prizes X, which is assumed to be a closed subset of some metric space. A simple lottery p on X is a probability distribution over X with a finite support. Let $\Delta(X)$ (or simply Δ) be the set of all simple lotteries on X. For any $p, q \in \Delta$ and $\alpha \in (0, 1)$, let $\alpha p + (1 - \alpha)q$ be the lottery that yields prize x with probability $\alpha p(x) + (1 - \alpha)q(x)$. Denote by δ_x the degenerate lottery that yields x with certainty and let $\bar{X} = \{\delta_x : x \in X\}$; we will often abuse notation and refer to δ_x simply as x. Similarly, denote by $\Delta(\Delta(X))$ (or simply Δ^2) the set of simple lotteries over Δ, that is, compound lotteries. For $P, Q \in \Delta^2$ and $\alpha \in (0, 1)$, denote by $R = \alpha P + (1 - \alpha)Q$ the lottery that yields simple (one-stage) lottery p with probability $\alpha P(p) + (1 - \alpha)Q(p)$. Denote by D_p the degenerate, in the first stage, compound lottery that yields p with certainty. Define a *reduction operator* $\phi : \Delta^2 \to \Delta$ that maps compound lotteries to reduced one-stage lotteries by $\phi(Q) = \sum_{p \in \Delta} Q(p)p$.[3] We refer to $\phi(Q)$ as prior beliefs (or simply a prior). When there is no risk of confusion, we sometimes refer to ϕ as the prior itself, without specifying the compound lottery that induced it. Our primitive is a binary relation $\succsim$ over Δ^2.

2.2 Functional Forms

We first formally define additive-belief-based utility.

Definition 1. *An additive-belief-based (ABB) representation is a tuple (u, ν_1, ν_2) that consists of continuous functions $u : X \to \mathbb{R}$, $\nu_1 : \Delta \times \Delta \to \mathbb{R}$, and $\nu_2 : \Delta \times \bar{X} \to \mathbb{R}$, such that $V_{ABB} : \Delta^2 \to \mathbb{R}$*

[3] Compound lotteries are isomorphic to the set of prior beliefs over outcomes plus a potential information structure. We can associate an information structure with the set of posterior beliefs it induces — the set of second-stage lotteries.

defined as

$$V_{ABB}(P) = \sum_j \phi(P)(x_j)u(x_j) + \sum_i P(p_i)\nu_1(\phi(P), p_i) + \sum_i P(p_i) \sum_j p_i(x_j)\nu_2(p_i, \delta_{x_j})$$

represents $\succsim$.

The general ABB representation allows utility to depend on changes in beliefs in period 1 and period 2. If utility depends on changes in beliefs, then ν_1 and ν_2 are functions of both their arguments. Alternatively, individuals may not care about changes, but rather about the levels of their beliefs in any given period. This can happen in one of two ways. The first case supposes that utility depends on beliefs at the beginning of any period, that is, ν_1 is solely a function of $\phi(P)$ and ν_2 is solely a function of p_i. We call this functional form prior-anticipatory utility.

Definition 2. *A prior-anticipatory representation is an ABB representation with the restrictions that* $\nu_1(\phi(P), p_i) = \hat{\nu_1}(\phi(P))$ *and* $\nu_2(p_i, \delta_{x_j}) = \hat{\nu_2}(p_i)$.

In the second case, utility is derived from beliefs at the end of any period (that period's posterior beliefs, after receiving information), that is, ν_1 is solely a function of p_i and ν_2 is solely a function of δ_{x_j}. We call this posterior-anticipatory utility.

Definition 3. *A posterior-anticipatory representation is an ABB representation with the restrictions that* $\nu_1(\phi(P), p_i) = \bar{\nu_1}(p_i)$ *and* $\nu_2(p_i, \delta_{x_j}) = \bar{\nu_2}(\delta_{x_j})$.

Clearly, both anticipatory representations above are subsets of V_{ABB}. More surprisingly, prior-anticipatory representation nests posterior-anticipatory representation.

Lemma 1. *If* $\succsim$ *has a posterior-anticipatory representation, then it has a prior-anticipatory representation.*

Proof of Lemma 1: By Definition 2 and setting $\tilde{\nu} = \sum_j \phi(P)(x_j)u(x_j) + \hat{\nu_1}(\phi(P))$, a prior-anticipatory representation is given by $\tilde{\nu_1}(\phi(P)) + \sum_i P(p_i)\hat{\nu_2}(p_i)$. Similarly, by Definition 3 and setting $\tilde{u}(x_j) = u(x_j) + \bar{\nu_2}(\delta_{x_j})$, a posterior-anticipatory representation is given by $\sum_i \phi(P)(x_j)\tilde{u}(x_j) + \sum_i P(p_i)\bar{\nu_1}(p_i)$. Clearly the latter is a subset of the former. $\square$

2.3 Characterization

We now characterize the functionals we have described using the relation $\succsim$. As will become apparent, our approach to restrict preferences is to impose Independence-type conditions on particular subsets of Δ^2. The first two axioms are standard.

Weak Order (WO) *The relation* $\succsim$ *is complete and transitive.*

Continuity (C) *The relation* $\succsim$ *is continuous.*

Our key axiom is Prior Conditional Two-Stage Independence (PTI). PTI requires the Independence axiom to hold within the set of compound lotteries which share the same prior beliefs. Observe that the set $\mathcal{P}(p) := \{Q \in \Delta^2 | \phi(Q) = p\}$ is convex for any $p \in \Delta$. Thus, PTI states that Independence holds along "slices" of the compound lottery space, where all elements of the slice have the same reduced form probabilities.

Prior Conditional Two-Stage Independence (PTI): *For any $P, P', Q \in \Delta^2$ and $\alpha \in [0,1]$, if $\phi(P) = \phi(P') = \phi(Q)$, then $P \succsim P'$ if and only if $\alpha P + (1-\alpha)Q \succsim \alpha P' + (1-\alpha)Q$.*

Recall that we identify preferences over compound lotteries with preferences over the combination of information structures and prior beliefs. PTI then requires that within a set of information structures that correspond to the same prior beliefs, the individual is an expected utility maximizer over their posterior beliefs; "non-standard" behavior may arise only when comparing across underlying prior beliefs.[4]

In addition to PTI, we need to link the evaluations made across different prior beliefs. This is the content of the following axiom, Cross Sectional Two-Stage Independence (CTI), which states that these comparisons are performed using the same "measurement rod".[5] In particular, CTI ensures that relative preferences of one set of posterior beliefs corresponding to prior ϕ, compared to a second set of posterior beliefs corresponding to prior ϕ', are not altered by mixing, so long as the mixing preserves the prior associated with each posterior belief.

Cross Sectional Two-Stage Independence (CTI): *For any $P, Q, R, S \in \Delta^2$ and $\alpha \in [0,1]$, if $\phi(P) = \phi(Q) \neq \phi(R) = \phi(S)$, $P \succsim R$, and $Q \succsim S$, then $\alpha P + (1-\alpha)Q \succsim \alpha R + (1-\alpha)S$.*

Our first main result shows that PTI and CTI, along with the standard two axioms above, characterize preferences that admit an ABB representation.

Proposition 1. *The relation $\succsim$ satisfies WO, C, PTI, and CTI, if and only if it has an ABB representation.*

All proofs are in Appendix A.1. To prove Proposition 1, we define a prior-conditional representation $V_{PC}(P) = \sum_i P(p_i)\nu(\phi(P), p_i)$, which is an expected utility functional (over the second-stage lotteries p_is) for a fixed prior $\phi(P)$. We then show that the relation $\succsim$ has a prior-conditional

[4] In the language of information preferences, PTI implies that preferences over statistical experiments — each of which chooses a signal with a probability that depends on the realized state of nature — are preserved by mixing (fixing a prior). Thus, if a sick individual prefers medical test A to B, he must prefer a λ chance of test A and $1-\lambda$ chance of test C to a randomization with the same proportion between B and C. Notice that the ranking between a pair of statistical experiments at one prior does not restrict the ranking of those experiments for any other prior — knowing that test A is preferred to B when sick conveys no information about the ranking of A and B (or any mixture involving them) when healthy.

[5] We could also state PTI and CTI as a single axiom, by not requiring $\phi(Q) \neq \phi(S)$ in the statement of CTI. Note that under this modification, taking $Q = S$ implies PTI. We have decided to state PTI and CTI as two separate requirements since they are conceptually different and play different roles in deriving the result representation. Indeed, we will see below that fixing CTI, both anticipatory representations are obtained by strengthening PTI.

representation if and only if it has an ABB representation. By PTI, an immediate application of the Mixture Space Theorem yields that fixing ϕ we have a prior-conditional representation. But while fixing ϕ the rankings within $\mathcal{P}(\phi)$ will not be affected by any monotone transformation of the prior-conditional utility, the rankings across different slices might. Axiom CTI rules this out: it guarantees that all such transformations are ϕ-independent, and can be taken without loss of generality to be the identity. We will revisit the observation that the general ABB representation can be compactly written as a simple expected utility functional in Section 2.4, when discussing its uniqueness properties.

PTI is not very restrictive, as it requires mixing not to reverse rankings only when all lotteries involved in the mixing have the same reduced form probabilities. A natural way to strengthen it is to suppose that only the compound lotteries involved in the *original* preference comparison need to have the same reduced form probabilities — the common compound lottery that they are mixed with need not. This means that the pair of lotteries which are compared after the mixing will have the same reduced form probabilities as each other, but need not have the same reduced form probabilities as the original pair. The next axiom formalizes this intuition.

Strong Prior Conditional Two-Stage Independence (SPTI): *For any $P, P', Q \in \Delta^2$ and $\alpha \in [0,1]$, if $\phi(P) = \phi(P')$, then $P \succsim P'$ if and only if $\alpha P + (1-\alpha)Q \succsim \alpha P' + (1-\alpha)Q$.*

SPTI rules out complementarity between prior beliefs and the corresponding information systems. That is, irrespectively of a fixed underlying prior, the individual consistently chooses among information systems based on the expected utility criterion over posterior beliefs; the relative value of posterior beliefs under prior ϕ does not change if we move to prior ϕ'. Thus, violations of expected utility may occur only when the individual compares two compound lotteries that do not refine the same prior beliefs. This means that SPTI still allows the trade-off between informational concerns (attitude towards the way uncertainty is resolved) and payoff concerns (prior beliefs) to vary across the entire domain. For example, suppose there are two outcomes $H > L$, and let $p = p(H)$. Then it may be that if p is low, the individual will prefer to slightly reduce the overall chance of receiving H in exchange for having all uncertainty resolves immediately; yet will care less about the resolution process where p is high and will not accept such compromise.

SPTI clearly implies PTI, but it is logically independent of CTI. Replacing PTI with SPTI yields our second characterization result:

Proposition 2. *The relation $\succsim$ satisfies WO, C, CTI, and SPTI, if and only if it has a prior-anticipatory representation.*

Similar to the previous proposition, the main step in proving Proposition 2 is showing that $\succsim$ has a prior-anticipatory representation if and only if it has a (prior-separable) representation of the form $V_{ps} = \nu_{ps1}(\phi(P)) + \sum_i P(p_i)\nu_{ps2}(p_i)$.

In order to characterize posterior-anticipatory representations, we further strengthen when Independence applies. The next axiom implies both CTI and SPTI, as it requires Independence to hold when mixing *any* two compound lotteries in the first stage.

Two-Stage Independence (TI): *For any* $P, P', Q \in \Delta^2$ *and* $\alpha \in [0,1]$, $P \succsim P'$ *if and only if* $\alpha P + (1-\alpha)Q \succsim \alpha P' + (1-\alpha)Q$.

TI implies that prior beliefs do not matter when considering preferences over information structures.[6] Our next result shows that it is equivalent to a posterior anticipatory representation.

Proposition 3. *The relation* $\succsim$ *satisfies WO, C, and TI, if and only if it has a posterior-anticipatory representation.*

2.4 Special Cases

The functional forms we have previously derived above are quite general. In many cases, we may want to impose further restrictions on the set of functionals we consider.

One typical assumption within the literature is that, in either stage, the individual receives the same utility (often normalized to 0) from beliefs that do not change. We describe these preferences as belief stationarity invariant (BSI).

Definition 4. *An ABB representation is belief stationarity invariant (BSI) if* $\nu_1(p,p) = \nu_2(\delta_x, \delta_x) = 0$ *for all* $x \in X$ *and* $p \in \Delta$.

A second type of assumption is that the utility derived from beliefs does not depend on which period those beliefs are realized. We call such preferences belief time invariant (BTI).[7]

Definition 5. *An ABB representation is belief time invariant (BTI) if* $\nu_1 = \nu_2$ *over their relevant shared domain.*

To relate BSI and BTI to behavior, we use the next axiom, which is due to Segal (1990).

Time Neutrality (TN): *For any* $p \in \Delta$, *if* $P = D_p$ *and* $Q = \sum_i p(x_i)\delta_{x_i}$, *then* $P \sim Q$.

Time Neutrality supposes that the individual is indifferent between a compound lottery that fully resolves in the first period and one that resolves only in the second period (so that the

[6] Seo (2009) also features mixing of compound lotteries. However, he has three stages, where the first stage is objective, second is subjective, and the third is objective. He provides an axiom called First-Stage Independence, which is analogous to TI. We use slightly different nomenclature to highlight the fact that the axioms operate on different domains.

[7] BTI rules out situations where individuals may get a stronger or weaker "kick" from beliefs if they occur sooner (for example, via discounting). We can weaken BTI to allow for such considerations, and say that ABB representation is pseudo-belief time invariant (PBTI) if for some scalar $\kappa > 0$, $\nu_1 = \kappa\nu_2$ over their relevant shared domain. However, in the end of the proof of Proposition 4 we show that PBTI does not restrict preferences alone and, furthermore, that BSI and PBTI in conjunction have no observable implications as well, as long as κ can be chosen arbitrarily.

information structure reveals no information in period 1), provided that they induce the same distribution over final outcomes, that is, they are both in $\mathcal{P}(p)$.

Although BSI and BTI do not restrict preferences alone, in conjunction they do.

Proposition 4. *Suppose $\succsim$ has an ABB representation. The following statements are true:*

1. *The relation $\succsim$ has a representation which is belief stationarity invariant.*
2. *The relation $\succsim$ has a representation which is belief time invariant.*
3. *The relation $\succsim$ has a representation which is both belief stationarity invariant and belief time invariant, if and only if it satisfies Time Neutrality.*

2.5 Uniqueness

The fact that we can obtain either a BSI or a BTI representation without loss of generality raises the question to what extent are ABB preferences uniquely identified.

The uniqueness property can be broken up into two parts. For both parts (as well as the discussion on uniqueness in the Appendix), we will make the assumption that, conditional on any prior, preferences are non-trivial. In other words, for all $\bar{\phi} \in \Delta$, there exists $P, P^{'} \in \Delta^2$ such that $\phi(P) = \phi(P^{'}) = \bar{\phi}$ and $P \succ P^{'}$.[8]

First, an immediate application of the mixture space theorem implies that if V and V' are both ABB representations of the same preference relation, then they differ by a positive affine transformation.

Proposition 5. *Suppose $\succsim$ has an ABB representation V. The ABB representation V' also represents $\succsim$ if and only if there exist scalars $\alpha > 0$ and β such that $V' = \alpha V + \beta$.*

Second, there are individual terms that can be subtracted from one component and absorbed in another, leaving the numerical value intact. In Appendix A.2 we show that the uniqueness results of the sub-components u, ν_1, and ν_2 are more subtle because any outcome that generates material utility must also appear in the support of the beliefs entering ν_1 and ν_2. Thus, one should expect that, without any further restrictions, there is some freedom to assign utility that is generated by any x appearing in the support of the lottery to either material utility or belief-based utility.

If, instead, we focus on the standard normalization applied in the literature (i.e., the one imposed by BSI), then u is unique up to an affine transformation, while ν_1 and ν_2 are unique up to common

[8]To see why without such assumption standard uniqueness results fail, note that if preferences conditional on a prior are trivial then they can be represented by arbitrary non-expected utility functionals. However, if conditional on any prior preferences are non-trivial, then continuity implies that for any prior $\bar{\phi}$, and for any other prior $\phi^{'}$ close enough to $\bar{\phi}$, there must be $P, P^{'}$ such that $\phi(P) = \bar{\phi}$, $\phi(P^{'}) = \phi^{'}$, and $P \sim P^{'}$. Then, applying the techniques in Claim 2 of the proof of Proposition 1, normalizations applied to one prior must be consistent with normalizations applied to any other prior, allowing us to to utilize standard techniques to establish uniqueness.

scaling. Since any ABB preferences have a BSI representation, this uniqueness result is entirely general.[9] Thus, we achieve a complete separation between preferences over final outcomes, which embed risk attitudes and are captured by u, and preferences over beliefs, which embed preferences for information and are captured by ν_1 and ν_2.

Proposition 6. *Suppose $\succsim$ has an ABB representation (u, ν_1, ν_2) that satisfies BSI. The ABB representation (u', ν_1', ν_2') also represents $\succsim$ and satisfies BSI if and only if there exist scalars $\alpha > 0$ and β_u such that $u'(x) = \alpha u + \beta_u$; $\nu_1'(\rho, p) = \alpha \nu_1(\rho, p)$; and $\nu_2'(p, \delta_x) = \alpha \nu_2(p, \delta_x)$.*

3 ABB Utility and Recursive Preferences

ABB preferences are not the only preferences used to model decisions over compound risk; an alternative specification is of preferences that are recursive. Recursive preferences have played an extensive role in a variety of models attempting to capture, among other things, choices over compound lotteries and information (see Kreps and Porteus, 1978; Segal, 1990; Grant, Kajii, and Polak, 1998; Dillenberger, 2010; Dillenberger and Segal, 2017; and Sarver, 2018).

In the definition below, $\mathbb{CE}_W(p)$ denotes the certainty equivalent of $p \in \Delta$ corresponding to the real function W on Δ, that is, $W(p) = W(\delta_{\mathbb{CE}_W(p)})$.[10]

Definition 6. *Suppose $\succsim$ over Δ^2 can be represented by the functional V. Then it has a recursive representation (V_1, V_2), where $V_i : \Delta \to \mathbb{R}$, if and only if for all $P = \sum_i P(p_i) D_{p_i}$, we have $V(P) = V_1(\sum_i P(p_i) \delta_{\mathbb{CE}_{V_2}(p_i)})$.*

Segal (1990) provides a behavioral equivalence for these functional forms using a substitution axiom he called Compound Independence, which we refer to as Recursivity.

Recursivity (R): *For any $p, q \in \Delta$, $Q \in \Delta^2$, and $\alpha \in [0, 1]$, $D_p \succsim D_q$ if and only if $\alpha D_p + (1-\alpha) Q \succsim \alpha D_q + (1-\alpha) Q$.*

Similarly to our previous main assumptions, Recursivity applies Independence to a particular "slice" of compound lotteries: the original pair of lotteries being compared must be degenerate in the first stage. This slice is orthogonal to that considered by CTI and PTI (or SPTI). Segal (1990) shows that the relation $\succsim$ satisfies WO, C, and R, if and only if it admits a recursive representation.

One immediate question is to what extent these two classes of utility, ABB and recursive, are related. Our next result shows that their intersection is exactly those preferences which admit a posterior-anticipatory representation.

[9] We only consider transformations that generate different actual values for all sub-functions involved in the transformation, and do not consider transformations that add and subtract elements to a specific sub-functoinal that leave its own actual value unchanged. For example, $\sum_x p(x) \gamma_\nu(p, \delta_x) = \sum_x p(x)(\gamma_\nu(p, \delta_x) + \epsilon_p(x))$ whenever $\sum_x p(x) \epsilon_p(x) = 0$.

[10] For the certainty equivalent to be well-defined, we need to impose some order on the set X. It will be the case whenever we take the set of prizes to be an interval $X \subset \mathbb{R}$ and both functions V_i in Definition 6 are monotone with respect to first-order stochastic dominance.

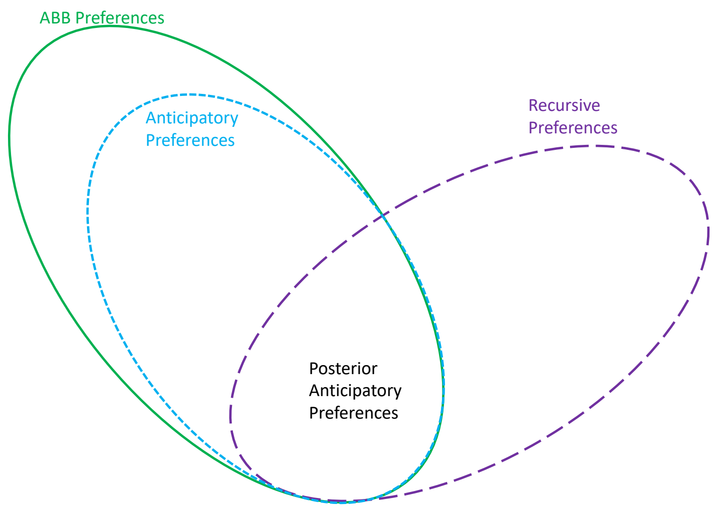

Figure 1: Relationships between models

Proposition 7. *The following statements are equivalent:*

- *The relation* $\succsim$ *satisfies WO, C, PTI, CTI, and R.*
- *The relation* $\succsim$ *has a posterior-anticipatory representation.*
- *The relation* $\succsim$ *has a recursive representation where* V_1 *is expected utility.*

Figure 1 depicts the relationships discussed here and the results of the last section. In general, ABB models can be directly tested (and falsified) in the natural domain of information preferences given a fixed prior — this is precisely the subdomain where axiom PTI bites. In contrast, recursive preferences, *as a general class*, have no observable restrictions when the prior is fixed; testing the assumption of Recursivity (Axiom R) requires observing preferences as the prior changes. Proposition 7 then implies that this latter property applies as well to the subclass of preferences that have posterior-anticipatory representation. The proposition further implies that any posterior-anticipatory representation captures individuals who have "emotions over emotions". This is because recursive models transform any two-stage compound lottery into a simple lottery over the second-stage certainty equivalents. Those certainty equivalents capture all second-period utility from beliefs changing in the second stage (e.g., disappointment/elation). Since first-period preferences take the certainty equivalents as the possible outcomes, those anticipated second period emotions are part and parcel of the "material" payoffs of the first-stage preferences. In all ABB models which do not have posterior-anticipatory representation, future changes in beliefs are independent of the effects of previous changes in beliefs. In other words, first period's beliefs-based utility, as captured by ν_1, is all about changes in material outcomes and does not depend on second period's belief-based utility, ν_2.

In addition to axiom R, Segal (1990) also introduces several other restrictions on preferences over compound lotteries (such as the Reduction of Compound Lotteries axiom and the requirement that Independence holds among both the set of full early resolving lotteries and fully late resolving lotteries). To complete our analysis, in Appendix A.3 we establish their relationship with CTI, PTI, SPTI, and TI, and further interpret these connections via the functional forms.

4 ABB Utility and the Timing of Resolution of Uncertainty

Individuals with ABB utility will have intrinsic preferences over information, that is, they may prefer one information structure to another even without the ability to condition actions on either of them. Many papers looking at specific examples of ABB preferences derive results regarding preferences over information, while focusing on two concepts: preferences for early versus late resolution of uncertainty and preferences for one-shot versus gradual resolution of uncertainty. In an analogous vein, characterizations of these informational attitudes for recursive preferences have been a major focus of the decision-theoretic literature. However, there do not exist equivalent characterizations for ABB preferences. Our results will allow to compare how different classes of models (recursive and ABB) can accommodate (or not) different non-instrumental attitudes towards information.

In many works, such as Kreps and Porteus (1978) and Grant, Kajii, and Polak (1998), individuals are allowed not only to prefer uncertainty to be fully resolved earlier (in period 1 rather than in period 2), but also to always prefer Blackwell-more-informative signals in period 1, that is, earlier resolution of uncertainty. Drawing on Grant, Kajii, and Polak (1998), we define a preference for early resolution of uncertainty as follows.

Definition 7. *The relation $\succsim$ displays a preference for early resolution of uncertainty if*

$$\beta\alpha D_q + (1-\beta)\alpha D_p + (1-\alpha)Q \succsim \alpha D_r + (1-\alpha)Q$$

for any $Q \in \Delta^2$, $p, q, r \in \Delta$, and $\alpha, \beta \in (0,1)$, such that $r = \beta p + (1-\beta)q$.

That is, preference for early resolution of uncertainty implies affinity towards splitting any branch that leads to some posterior beliefs r into several branches, whenever the split consists of a mean-preserving spread of r. As Grant, Kajii, and Polak (1998) show, this notion indeed characterizes intrinsic information loving (preference for more informative signals). Preference for late resolution of uncertainty is analogously defined, by requiring the reverse ranking for such lotteries.

We now characterize preferences that exhibit preferences for either early or late resolved lotteries. Similarly to known results about recursive preferences, attitude towards the resolution of uncertainty is characterized in our model by the curvature of the appropriate components. In the next result we refer to preferences which have a prior-anticipatory representation, but do not have a posterior-anticipatory representation, as having a prior*- anticipatory representation.[11]

Proposition 8. *The following statements are true:*

1. *Suppose $\succsim$ has an ABB representation. Then $\succsim$ exhibits a preference for early (resp., late) resolution of uncertainty if and only if $\nu_1(\rho, \cdot) + \sum_x \nu_2(\cdot, x)$ is convex (resp., concave).*

2. *Suppose $\succsim$ has a prior*-anticipatory representation. Then $\succsim$ exhibits a preference for early (resp., late) resolution of uncertainty if and only if $\hat{\nu_2}$ is convex (resp., concave).*

3. *Suppose $\succsim$ has a posterior-anticipatory representation. Then $\succsim$ exhibits a preference for early (resp., late) resolution of uncertainty if and only if $\bar{\nu_1}$ is convex (resp., concave).*

A distinct notion of preferences for resolution of uncertainty is introduced by Dillenberger (2010). He supposes that individuals satisfy Time Neutrality (axiom TN described earlier) and that they prefer either compound lotteries in which *all* uncertainty is resolved in period 1 or in period 2 to any other compound lotteries which induce the same prior beliefs.[12] He defines this as a preference for one-shot resolution of uncertainty.

Definition 8. *The relation $\succsim$ exhibits a preference for one-shot resolution of uncertainty (PORU) if $P \sim Q \succsim R$, for all $P, Q, R \in \Delta^2$ such that $\phi(P) = \phi(Q) = \phi(R)$, $P = D_p$, and $Q = \sum_i p(x_i)\delta_{x_i}$.*

Our next result characterizes ABB preferences that exhibit PORU; it further shows that the class of anticipatory preferences is inconsistent with this property.

Proposition 9. *The following statements are true:*

1. *Suppose $\succsim$ has an ABB representation. Then $\succsim$ exhibits PORU if and only if*

$$\sum_i P(p_i) \sum_x p_i(x)\nu_1(\phi(P), \delta_x) \geq \sum_i P(p_i) \sum_x p_i(x)\nu_1(\phi(P), p_i) + \sum_i P(p_i) \sum_x p_i(x)\nu_1(p_i, \delta_x)$$

2. *If $\succsim$ has a prior-anticipatory representation, then it can never exhibit strict PORU.*[13]

For item (1), suppose $\succsim$ has an ABB representation. Observe first that PORU implies TN. Therefore, by Proposition 4, if $\succsim$ exhibits PORU then it must have an ABB representation that satisfies both BSI and BTI, and in particular $\nu_1 = \nu_2$. And since the expected utility from material

[11] Recall from Definition 2 that for a prior*- anticipatory representation $\nu_2(p_i, \delta_{x_j}) = \hat{\nu_2}(p_i)$ and from Definition 3 that for posterior-anticipatory preferences $\nu_1(\phi(P), p_i) = \bar{\nu_1}(p_i)$.

[12] This is tightly linked to the notion of a preference for clumped information, introduced by Kőszegi and Rabin (2009), which we explore in Section 6.1.

[13] That is, there are no triples as in Definition 8 for which $P \sim Q \succ R$.

payoffs is the same in all lotteries compared, the result follows.[14] The intuition behind item (2) derives from Corollary 1 of Appendix A.3, where we show that if $\succsim$ has a prior-anticipatory representation, then TN implies that ν_2 is an expected utility functional, and thus does not generate any intrinsic preferences towards information. The rest of the utility functional depends only on ϕ, the reduced form probability of the lottery, independently of the pattern of resolution of uncertainty. The result suggests PORU as a sufficient condition to rule out anticipatory preferences.[15] In contrast, Proposition 8 implies that displaying preference for early resolution of uncertainty cannot help ruling out a specific form of belief-based utility.

All our results thus far are general and are not tied to specific functional form assumptions. In Section 6 we demonstrate how they can be applied to enrich our understanding of known models, focusing on the content of Propositions 8 and 9.

5 The Scope of ABB Preferences

In this section we discuss how our three categories of utility functionals relate to particular models used in the literature. Table 1 provides a list of papers which use functional forms nested by ABB. Many of these models also allow individuals to take intermediate actions, so their domain and representation may appear different than that presented in this paper.[16]

Some models in the literature fit into the framework of changing beliefs models — they have an ABB representation, but do not have any anticipatory representation. These include models that were meant to captures utility derived from changing beliefs, such as Kőszegi and Rabin (2009) and Pagel (2018). In other models, utility may not be derived from changes in beliefs per se. Rather, utility is derived from levels of beliefs, but the function that determines the levels depends on the prior beliefs. These include the models of Mullainathan and Shleifer (2005), where individuals seek signals that confirm their priors, and Caplin and Eliaz (2003), which has the form

[14] Note that the inequality in item (1) has the form of (the opposite of) the triangle inequality, which should hold 'on average'. Indeed, one example of a function that satisfies it is $\nu_1(\rho, p) = -d(\rho, p)$, where d is some standard distance measure on the unit simplex. In this case, the utility loss from beliefs moving equals the total expected distance traveled by beliefs. Indeed, for such function the inequality holds term by term (this is the negative of the triangle inequality) and thus also in expectation. One interpretation for such functional form would be that the agent is simply averse to any changes in beliefs, presumably due to some hidden costs of adjustments (think of an agent who pre-committed to some belief-contingent action, or someone who placed bets on his intermediate beliefs) that are 'large enough', overwhelming any effect of good news.

[15] In the context of recursive non-expected utility preferences with TN, Dillenberger (2010) shows equivalence between PORU and a property of static preferences called Negative Certainty Independence (NCI). (The class of preferences that satisfy NCI was then characterized by Cerreia Vioglio et al., 2015.) This equivalence ceases to hold for the general ABB models. The reason is that in that setting dynamic preferences are not fully determined by (recursive application of) static preferences. In particular, the first term on the right-hand side of the inequality in item 1 of Proposition 9 can always be manipulated so that PORU is violated even though the restriction of preferences to fully early resolved lotteries does satisfy NCI.

[16] As in the previous section, we denote models that have prior-anticipatory but not posterior-anticipatory representation as having prior*-anticipatory utility.

Changing Beliefs	Prior*-Anticipatory Utility	Posterior-Anticipatory Utility
Caplin and Eliaz (2003)	See text	Kreps and Porteus (1979)
Mullainathan and Shleifer (2005)		Epstein and Zin (1989)
Kőszegi and Rabin (2009)		Caplin and Leahy (2001)
Kőszegi (2010)		Kőszegi (2003)
Pagel (2018)		Caplin and Leahy (2004)
		Kőszegi (2006)
		Eliaz and Spiegler (2006)
		Eliaz and Schotter (2010)
		Schweitzer and Szech (2018)

Table 1: Some models nested by ABB

of prior-dependent preferences over the time of resolution of uncertainty. Within our domain, we also capture the model of Kőszegi (2010), who models expectations (i.e., beliefs) that interact with material payoffs (although his domain also allows for actions and material payoffs in period 1).

Other models in the literature adhere to the anticipatory-utility model. In particular, our posterior-anticipatory functional form nests all the models in Caplin and Leahy's (2001) framework, including Kőszegi (2003), Caplin and Leahy (2004), Kőszegi (2006), Eliaz and Spiegler (2006), Eliaz and Schotter (2010), and Schweitzer and Szech (20186).

We know of no existing models that explicitly capture pure prior-anticipatory motivations, that is, models that admit a prior-anticipatory — but not posterior-anticipatory — representation. To better understand the gap between the two anticipatory representations, in Proposition 15 of Appendix A.3 we provide an alternative characterization of posterior-anticipatory representation, which amounts to adding to the axioms underlying the prior-anticipatory representation an additional Independence requirement on the subset of early resolving lotteries (compound lotteries in which all uncertainty is resolved in the first stage). Descriptively, prior-anticipatory motives are important as they can accommodate a behavior that violates expected utility over early resolving lotteries — in accordance with frequently observed experimental results, such as the Allais paradox — and *cannot* be accommodated by posterior-anticipatory preferences. To see that this sub-class is nonempty, let f and g be two arbitrary non-expected utility functionals. In Claim 3 of Appendix A.1 we establish that the function $V(P) = f(\phi(P)) + \sum_i P(p_i)g(p_i)$ is a prior*-anticipatory utility.

Caplin and Leahy (2001) is one of the first papers to explicitly study belief-based utility. They provide an axiomatization of their functional form, but take as their domain the set of "psychological lotteries", which includes lotteries not just over material outcomes, but also over psychological states (i.e., beliefs). Thus, their domain includes objects (psychological states) which are not observable, and are not directly chooseable. Our approach, which confines attention to preferences over compound lotteries, ensures that all our restrictions are stated solely in terms of preferences over observable objects. In fact, given this domain, Caplin and Leahy's preferences are exactly equivalent to posterior-anticipatory preferences.

Observation 1. *The relation $\succsim$ can be represented by a Caplin and Leahy (2001) utility function if and only if it has a posterior-anticipatory representation.*[17]

While, as Caplin and Leahy allude to, posterior-anticipatory preferences nest those recursive preferences which are expected utility over early resolving lotteries, unmentioned by them—but following immediately from Proposition 7 and Observation 1— is that they are also nested by the more general class of recursive preferences. This immediately implies that we can carry over known tools and intuitions from recursive models to analyze posterior-anticipatory models.

Ely, Frankel, and Kamenica (2015) also model individuals who care about changes in their beliefs. Their model of surprise has the same structure as our ABB representation, but is formally not within the class of models we study because it is discontinuous. Their model of suspense is similar in spirit, but formally different from ABB models not only due to its lack of continuity, but also because of a non-linear transformation that is applied to the expected utility from changes in beliefs.

Gul, Natenzon, and Pesendorfer (2016) study a model that shares some key features of ours. Although their domain and objects of choice are different, both papers consider utility functions where individuals gain utility from beliefs and from material payoffs in a way that is additively separable. However, while we suppose individuals calculate the expectation of a belief-based utility using objective probabilities, they allow for non-additive measures.

The ABB representation is reminiscent of the functional forms used in the rational inattention literature (e.g. Sims, 2003). For example, if we assume $\nu_2 = 0$, then we can interpret ν_1 as the cost of beliefs shifting from the prior to the posterior. There are two caveats to this analogy. First, while in our current framework individuals do not take intermediate actions, in rational inattention models individuals only pay the cost of learning in order to match actions to states. The second difference is more subtle. The rational inattention literature typically supposes that individuals choose information in a way that is unobserved by the analyst — in an extreme case, information is obtained solely by self-reflection. Our domain, on the other hand, specifically supposes we observe individuals choosing between known, objectively given information structures. Therefore, one can interpret the cost of information in our setting as the amount that must be paid conditional on choosing a particular information structure. This is distinct from a rational inattention story, where even in the face of objective information, an individual could avoid paying costs by simply

[17]In particular, Caplin and Leahy (2001) representation is Equation (1) in their paper. The mapping is that our $u + \nu_2$ is equivalent to their u_2, and our ν_1 is equivalent to their $u_1(\phi)$ (they use ϕ as a functional, rather than to denote reduce form probabilities).

not processing that information.

While a vast literature, spanning both economics and psychology, demonstrated non-instrumental preferences for information, it has primarily focused on testing attitudes towards the resolution of uncertainty.[18] There has been little work directly testing how we should model such preferences; for example, whether recursivity is a reasonable assumption. Our results can guide future experimental tests to understand the underlying structure of preferences for information and belief-based utility.

6 Applications

We now discuss how our results can be applied to some models encompassed by the ABB class. We focus on showing the implications for informational preferences (as in Section 4), which is one of the most commonly discussed applications of belief-based preferences. Since posterior-anticipatory preferences are subset of recursive preferences for which these implications are known, we confine our attention to the two other categories of interest.

6.1 Changing Beliefs Preferences

Kőszegi and Rabin (2009) is our leading example of changing beliefs preferences. Their specification is as follows: given lottery p, let $c_p(\psi)$ be the payoff at percentile ψ of the distribution induced by p. Then $\nu_1 = \kappa_1 \int \mu(u(c_{p_i}(\psi)) - u(c_{\phi(P)}(\psi)))d\psi$, $\nu_2 = \kappa_2 \int \mu(u(c_{\delta_x}(\psi)) - u(c_{p_i}(\psi)))d\psi$, where μ is a gain-loss utility function such that (i) μ is continuous, strictly increasing, twice differentiable for $x \neq 0$, and satisfying $\mu(0) = 0$; (ii) $\mu(y) + \mu(-y) < \mu(x) + \mu(-x)$ whenever $y > x \geq 0$; (iii) $\mu''(x) \leq 0$ for $x > 0$ and $\mu''(x) \geq 0$ for $x < 0$; and (iv) $\frac{\lim_{x\to 0} \mu'(|x|)}{\lim_{x\to 0} \mu'(-|x|)} = \lambda > 1$.

We will focus on understanding under what restrictions Kőszegi and Rabin's model satisfies PORU. This is of interest both theoretically and experimentally (e.g., Zimmerman, 2014; Falk and Zimmerman, 2016; Nielsen, 2017; Masatlioglu, Orhun and Raymond, 2017). Accordingly, assume that $\kappa_1 = \kappa_2$. Kőszegi and Rabin confined their attention to only two final outcomes and show that if the gain-loss function is piecewise linear (item (iii) with $\mu''(x) = 0$ for all $x \neq 0$), then their preferences satisfy a preference for clumped information; a phenomenon that is tightly linked to PORU. They further conjectured, but did not prove, that their result generalizes to any number of outcomes.

Our next result first establishes that Kőszegi and Rabin's conjecture is correct; piecewise linear gain-loss utility implies PORU. If we consider only compound lotteries that have two final outcomes ({P $|\phi(P)$ is binary}), this is also true so long as there is no too much diminishing sensitivity in

[18]See, for example, Chew and Ho (1994); Ahlbrecht and Weber (1997); Arai (199)7; Lovallo and Kahneman (2000); Eliaz and Schotter (2010); Von Gaudecker et al. (2011); Brown and Kim (2014); Kocher, Krawczyk, and Van Winden (2014); Ganguly and Tassoff (2017); Falk and Zimmerman (2016); Nielsen (2017); Masatlioglu, Raymond, and Orhun (2017).

the loss domain, that is, if $\mu'(x)$ is large enough for all $x < 0$. We then demonstrate that this prediction is not global. In particular, if losses are not experienced much more strongly than gains, then PORU doesn't hold. In other words, loss aversion is not sufficient for PORU. To simplify exposition, in stating the following result we define the gain function as the restriction of μ to the domain of the weakly positive numbers (and the loss function analogously).

Proposition 10. *Suppose $\succsim$ is represented by Kőszegi and Rabin utility function as specified above, with $\kappa_1 = \kappa_2$. The following statements are true:*

1. *If $\mu''(x) = 0$ for all $x \neq 0$, then $\succsim$ exhibits PORU.*
2. *Consider $\{P \mid \phi(P)$ is binary$\}$. For any gain function, if $\mu'(x)$ is large enough for all $x < 0$, then $\succsim$ exhibits PORU.*
3. *For any loss function, if gains are sufficiently similar to losses (i.e. $\mu(x) + \mu(-x)$ goes to 0 for all x) then $\succsim$ violates PORU.*[19]

Kőszegi and Rabin's functional form (along with closely related variants) is the most prominent changing-beliefs model in the literature. However, little work has been done trying to understand to what extent the intuitions developed for that model extend. An immediate implication of their asymmetry between gains and losses is that gain-loss utility must be concave in a small neighborhood of the reference point. Our next result connects this property to our general result about PORU (Proposition 9). It shows that "local reference dependence", in that losses must be experienced more strongly than gains relative to the prior belief, is implied by PORU. The form of loss aversion imposed by Kőszegi and Rabin is simply one way of generating such local reference dependence. The assumption that ν_1 is smooth in the first argument is often innocuous, e.g., if with only two outcomes, so long as ν_1 is monotone in the first argument then smoothness is satisfied almost everywhere.

Observation 2. *Suppose $\succsim$ has an ABB representation that exhibits PORU. If $\nu_1(\phi, p)$ is smooth in the first argument, then it must be concave in the second argument at ϕ.*

6.2 Prior-Anticipatory Preferences

As previously noted, we are unaware of models that fit into the class of prior*-anticipatory models. We have already mentioned the ability of these preferences to capture non-standard risk attitudes, in contrast to posterior-anticipatory models (which must satisfy expected utility over early resolving lotteries). As we demonstrate below, this feature is *not* inconsistent with also displaying nontrivial

[19] Duraj and He (2019) show that if there are more than two final outcomes, then PORU is not satisfied whenever the gain function exhibits diminishing sensitivity.

informational preferences, a requirement that poses difficulties to many recursive preferences (Definition 6). For example, Grant, Kajii, and Polak (2000) show that assuming preference for early resolution of uncertainty implies that the recursive forms of two prominent classes of non-expected utility models — Rank-Dependent Utility (Quiggin, 1982) and Betweenness (Chew, 1983; Dekel, 1986) — each almost collapses back to recursive expected utility.

Our next result complements the discussion above by showing that a well known behavioral phenomenon — first-order risk aversion, which captures risk aversion over small stakes — is inconsistent with global preference for early or late resolution of uncertainty within standard recursive models, but is consistent with these properties in the prior-anticipatory framework.[20]

Proposition 11. *The following statements are true:*

1. *Suppose $\succsim$ admits a recursive representation where both V_1 and V_2 are monotone. Suppose its restriction to early resolving lotteries, represented by V_1, (to late resolving lotteries, represented by V_2, respectively) exhibits first-order risk aversion. Then $\succsim$ cannot exhibit preference for early resolution of information (late resolution of information, respectively).*

2. *There exist a prior*-anticipatory preference relation that exhibits first-order risk aversion over both early and late resolving lotteries, as well as a global preference for either early resolution of uncertainty or late resolution of uncertainty.*

To conclude, Part 2 of Proposition 11 suggests that prior-anticipatory preferences may increase the explanatory power of belief-based models by simultaneously generating commonly observed behavior in both the domains of risk and information. We have seen that posterior-anticipatory models cannot accommodate violations of expected utility and that the commonly used changing-beliefs specification of Kőszegi and Rabin (2009) as in Section 6.1, fails to capture preferences for earlier or later resolution of uncertainty. Part 1 of Proposition 11 shows that Recursive models can match either violations of expected utility that are related to first-order risk aversion — with a prominent example being loss aversion — or can allow for preferences for earlier or later resolution of uncertainty, but not both at the same time. On the other hand, as shown in Proposition 9, neither class of anticipatory preferences can exhibit strict PORU, while changing beliefs models can. This suggests a trade-off between the changing beliefs and anticipatory approaches; knowing which avenue is empirically most relevant is a topic left for future research.

[20] This result implies the converse of Proposition 1 (ii) of Grant et al. (2000). This is because their result is for Rank-Dependent Utility. In that model, so long as the probability weighting function for late resolving lotteries is differentiable, V_2 is smooth away from degenerate lotteries. And since the weighting function is strictly increasing, it must be differentiable almost everywhere. However, our result is distinct from Proposition 1 (i) of Grant et al (2000); that result applies to Betweenness preferences, which do not necessarily exhibit first-order risk aversion.

A Appendix

A.1 Proofs

Proof of Proposition 1. Let $V_{PC}(P) := \sum_i P(p_i)\nu_{PC}(\phi(P), p_i)$. We refer to V_{PC} as a prior-conditional representation.

Claim 1. *The relation $\succsim$ has a prior-conditional representation if and only if it has an ABB representation.*

Proof of Claim 1. Consider the first two terms in the ABB representation. Observe that the term $\sum_i P(p_i) \sum_j p_i(x_j)u(x_j)+\sum_i P(p_i)\nu_1(\phi(P), p_i)$ can be rewritten as $\sum_i P(p_i)\hat{\nu}_1(\phi(P), p_i)$. Similarly, any $\sum_i P(p_i)\hat{\nu}_1(\phi(P), p_i)$ can be rewritten as $\sum_i P(p_i)\sum_j p_i(x_j)u(x_j) + \sum_i P(p_i)\nu_1(\phi(P), p_i)$.

Consider now the third term in the ABB representation. Note that any $\sum_i P(p_i)\sum_j p_i(x_j)\nu_2(p_i, \delta_{x_j})$ can be rewritten as $\sum_i P(p_i)\hat{\nu}_2(p_i)$, since p_i embeds all the $x_j s$ in it's support. Moreover, given any $\sum_i P(p_i)\hat{\nu}_2(p_i)$, we can rewrite it as $\sum_i P(p_i)\sum_j p_i(x_j)\nu_2(p_i, \delta_{x_j})$.

Thus preferences have an ABB representation if and only if they can be represented by

$$\sum_i P(p_i)\hat{\nu}_1(\phi(P), p_i) + \sum_i P(p_i)\hat{\nu}_2(p_i)$$

Simplifying further, observe that any $\sum_i P(p_i)\hat{\nu}_1(\phi(P), p_i) + \sum_i P(p_i)\hat{\nu}_2(p_i)$ can be rewritten as $\sum_i P(p_i)\widetilde{\nu}(\phi(P), p_i)$; and any $\sum_i P(p_i)\widetilde{\nu}(\phi(P), p_i)$ can be rewritten as $\sum_i P(p_i)\hat{\nu}_1(\phi(P), p_i) + \sum_i P(p_i)\hat{\nu}_2(p_i)$. We have just proved that $\succsim$ has a representation of the form V_{ABB} if and only if it has a representation of the form $\sum_i P(p_i)\widetilde{\nu}(\phi(P), p_i)$. □

We now use our new representation for Claim 2.

Claim 2. *The relation $\succsim$ has a prior-conditional representation if and only if it satisfies WO, C, PTI, and CTI.*

Proof of Claim 2. Observe that the prior-conditional representation holds if and only if for any fixed ϕ preferences are expected utility, which is known to be equivalent to WO, C, and PTI. Each of these conditional expected utility functionals are ordinally unique up to a monotone transformation f_ϕ, which a priori may depend on the reduced form probabilities. We now show that this is impossible, so that f is independent of ϕ and thus can be taken without loss of generality to be affine.

We partition the set of compound lotteries into two types of orthogonal equivalence classes. The first is the set $\mathcal{I}$ of equivalence classes induced by $\succsim$; denote an arbitrary class by I (with $P, Q \in I$ if and only if $P \sim Q$), and for any $P \in \Delta$ let $I(P) = \{Q : P \sim Q\}$. The second set includes the equivalence classes induced by ϕ, with a generic element $\mathcal{P}(\phi) = \{P \in \Delta : \phi(P) = \phi\}$.[21] Note that if there are disjoint $I, I' \in \mathcal{I}$ that have nonempty intersection with $\mathcal{P}(\phi)$, then all equivalence classes of $\succsim$ that contain an element of $\mathcal{P}(\phi)$ must form a convex set. To see this, let $I(\phi) = \{I(P) : P \in \mathcal{P}(\phi)\}$ and consider any two elements $I, I' \in I(\phi)$ with corresponding $P \in I \cap \mathcal{P}(\phi)$ and $P' \in I' \cap \mathcal{P}(\phi)$. Since by PTI standard Independence holds on $\mathcal{P}(\phi)$, preferences must be continuous over mixtures of P and P' and so for any $\alpha \in [0,1]$, $\alpha P + (1-\alpha)P'$ must be ranked in terms of $\succsim$ between P and P' (and conversely, for any Q ranked in terms of $\succsim$ between P and P', we can find an $\alpha \in [0,1]$ such that $\alpha P + (1 - \alpha P') \sim Q$.)

We next show that for any $I \in \mathcal{I}$, f_ϕ is independent of ϕ on I. We do it by dividing the set $\mathcal{I}$ into two classes and show that f is independent of ϕ within each class. The classes have non-intersecting interiors, but may overlap at their boundary points.

1. This class includes all $I \in \mathcal{I}$ such that $I = \bigcup_{\phi \in \Phi} \mathcal{P}(\phi)$ for some set of priors Φ. This means that if a given $\mathcal{P}(\phi)$ has an element in one such I, then all elements of $\mathcal{P}(\phi)$ must be in I. We can then normalize the utility of all elements of $\{\mathcal{P}(\phi) : \phi \in \Phi\}$ to some constant k. Then f is independent of any $\phi \in \Phi$ on I (since all f_ϕs take in k as an argument and must output the same number).

2. This class includes all $I \in \mathcal{I}$ such that there exists a ϕ where both $I \cap \mathcal{P}(\phi) \neq \emptyset$ and $(\mathcal{I} \setminus I) \ni J \cap \mathcal{P}(\phi) \neq \emptyset$ hold. Thus, if a given $\mathcal{P}(\phi)$ has an element in one such I, then another element of $\mathcal{P}(\phi)$ must be in a different I' (which by construction is also part of this class). Denote the set of these Is as $\mathbb{I}$. Denote the closure of any set Z as $\mathrm{cl}(Z)$.

   Pick a ϕ such that $I(\phi) \subseteq \mathrm{cl}(\mathbb{I})$ and either (i) there exists a worst $\succsim$-equivalence class that a member of $\mathcal{P}(\phi)$ is in, denoted $I_{\min}(\phi)$, and this is is the lowest element of $\mathrm{cl}(\mathbb{I})$: $P \in I_{\min}(\phi) \implies Q \succsim P, \forall Q \in I \in \mathbb{I}$; or (ii) there exists a sequence of $P_n \in \mathcal{P}(\phi)$ such that, for each n, $P_n \in I_n \in \mathbb{I}$ and the sequence I_n converges to the lowest element of $\mathrm{cl}(\mathbb{I})$. In either case we denote by $\underline{I(\phi)}$ an element of the greatest lower bound of the set of indifference classes each containing some $P \in \mathcal{P}(\phi)$. Thus, we pick out a ϕ such that $I(\phi)$ is not singleton, and there exists $I \in \mathrm{cl}(I(\phi))$ that is included in the worst indifference class of $\mathbb{I}$.

   Set $I^1 = \mathrm{cl}(I(\phi))$. Inductively, set $I^n = I(\phi)$ for a ϕ such that $I_{\min}(\phi)$ is the lowest element of $\mathrm{cl}(\mathbb{I} \setminus \bigcup_{i=1\ldots n-1} I^i)$. We refer to each I^n as a sub-class. For any n, we denote by ϕ^n the ϕ that generates I^n. Note that this construction spans the entire $\mathbb{I}$; that each I^i is a convex set and is closed (by construction); that given any two I^k and I^j, their intersection can only consist of a single equivalence class I; and that there must be a countable number of these sub-classes.

[21] If $\mathcal{P}(\phi) \subseteq I$ for some I, then, for this ϕ, the decision maker is simply indifferent to the resolution of uncertainty.

Start with I^1. There are two cases to consider. The first case is where $I(\phi^1) = I^1$. In this case, the maximal and minimal $\succsim$-equivalence classes of I^1 contain members of $\mathcal{P}(\phi^1)$. For any I^i, let $\overline{I^i}$ be a lottery in the $\succsim$-maximal $I \in I^i$ and $\underline{I^i}$ a lottery in the $\succsim$-minimal $I \in I^i$. Note that by the uniqueness result of the vNM functional on each slice, we can normalize $\sum_j \overline{I^1}(p_j)\widetilde{\nu}(\phi^1, p_j) = 1$ and $\sum_j \underline{I^1}(p_j)\widetilde{\nu}(\phi^1, p_j) = 0$. Using the fact that within a slice Independence is satisfied, for any $P \in \mathcal{P}(\phi^1)$ we can assign a number $\lambda(P) \in [0, 1]$ such that $\sum_j P(p_j)\widetilde{\nu}(\phi^1, p_j) = \lambda(P)$.

In the second case, $I(\phi^1) \subset I^1$. In this case, the maximal and minimal $\succsim$-equivalence classes of I^1 are reached in the limit by a sequence $\{I_k\}$ each contains a member of $\mathcal{P}(\phi^1)$. Suppose it is the minimal $\succsim$-equivalence class (the other case is analogous). Then, as described above, we denote by $\underline{I^1}$ an element of the greatest lower bound of $\{I_k\}$ and repeat the normalization process described in the previous case, but where $\lambda(P)$ is identified as the limit of our mixing operation using $\overline{I^i}$ and a sequence of elements in $I_k \cap \mathcal{P}(\phi^1)$.

We have now completed our normalization for I^1. We extend the normalization process inductively for any I^i, where $i > 1$, considering three different cases. The proof below supposes that $I(\phi^i) = I^i$. When $I(\phi^i) \subset I^i$ the mixture operation involving $\overline{I^i}$ and $\underline{I^i}$ can be interpreted as mixing with a sequence of lotteries and taking the limit.

Remark 1. *In what follows, we abuse notation and denote by $I^i \cap \mathcal{P}(\phi)$ a subset of $\mathcal{P}(\phi)$, where each of its elements belongs to some $I \in I^i$. We also write $P \in I^i$ rather than "$P \in I \in I^i$".*

(a) The first case is where $I^{i-1} \cap I^i = \emptyset$. In this case, there is a "gap" in the indifference classes between I^{i-1} and I^i. Suppose that $\sum_j \overline{I^{i-1}}(p_j)\widetilde{\nu}(\phi^{i-1}, p_j) = \kappa$. We fix a $\delta > 0$ and set $\sum_j \underline{I^i}(p_j)\widetilde{\nu}(\phi^i, p_j) = \kappa + \delta$ and $\sum_j \overline{I^i}(p_j)\widetilde{\nu}(\phi^i, p_j) = \kappa + 1 + \delta$. For any $P \in \mathcal{P}(\phi^i)$ assign a number $\lambda(P) \in [\kappa + \delta, \kappa + \delta + 1]$ such that $\sum_j P(p_j)\widetilde{\nu}(\phi^i, p_j) = \lambda(P)$.

    Now consider any $\phi' \neq \phi^i$ such that $I^i \cap \mathcal{P}(\phi') \neq \emptyset$. Let $\underline{R^i(\phi')}$ be a $\succsim$ $-$minimal element in $I^i \cap \mathcal{P}(\phi')$. Similarly, let $\overline{R^i(\phi')}$ be a $\succsim$ $-$maximal element in $I^i \cap \mathcal{P}(\phi')$. Find $P \in I^i \cap \mathcal{P}(\phi^i)$ such that $P \sim \underline{R^i(\phi')}$ and set $\sum_j \underline{R^i(\phi')}(p_j)\widetilde{\nu}(\phi', p_j) = \lambda(P)$. Similarly, find $P' \in I^i \cap \mathcal{P}(\phi^i)$ such that $P' \sim \overline{R^i(\phi')}$ and set $\sum_j \overline{R^i(\phi')}(p_j)\widetilde{\nu}(\phi', p_j) = \lambda(P')$.

(b) The second case is where $I^{i-1} \cap I^i \neq \emptyset$ (recall that by construction these two sets can only overlap at a single indifference class) and there is no $\tilde{\phi}$ such that both $\mathcal{P}(\tilde{\phi}) \cap I^{i-1}$ and $\mathcal{P}(\tilde{\phi}) \cap I^i$ consist of more than a singleton. In this case, there is no "gap" in the indifference classes between I^{i-1} and I^i but there is no prior ϕ such that $I(\phi)$ has non-trivial intersection with both I^{i-1} and I^i. Suppose that $\sum_j \overline{I^{i-1}}(p_j)\widetilde{\nu}(\phi^{i-1}, p_j) = \kappa$. We set $\sum_j \underline{I^i}(p_j)\widetilde{\nu}(\phi^i, p_j) = \kappa$ and $\sum_j \overline{I^i}(p_j)\widetilde{\nu}(\phi^i, p_j) = \kappa + 1$. For any $P \in \mathcal{P}(\phi^i)$ assign a number $\lambda(P) \in [\kappa, \kappa + 1]$ such that $\sum_j P(p_j)\widetilde{\nu}(\phi^i, p_j) = \lambda(P)$.

As in the previous case, now consider any $\phi' \neq \phi^i$ such that $I^i \cap \mathcal{P}(\phi') \neq \emptyset$. Let $\underline{R^i(\phi')}$ be a $\succsim$ −minimal element in $I^i \cap \mathcal{P}(\phi')$. Similarly, let $\overline{R^i(\phi')}$ be a $\succsim$ −maximal element in $I^i \cap \mathcal{P}(\phi')$. Find $P \in I^i \cap \mathcal{P}(\phi^i)$ such that $P \sim \underline{R^i(\phi')}$ and set $\sum_j \underline{R^i(\phi')}(p_j)\widetilde{\nu}(\phi', p_j) = \lambda(P)$. Similarly, find $P' \in I^i \cap \mathcal{P}(\phi^i)$ such that $P' \sim \overline{R^i(\phi')}$ and set $\sum_j \overline{R^i(\phi')}(p_j)\widetilde{\nu}(\phi', p_j) = \lambda(P')$.

(c) The third case is where $I^{i-1} \cap I^i \neq \emptyset$ (recall that by construction these two sets can only overlap at a single indifference class) and there exists a $\tilde{\phi}$ such that both $\mathcal{P}(\tilde{\phi}) \cap I^{i-1}$ and $\mathcal{P}(\tilde{\phi}) \cap I^i$ consist of more than a singleton. In this case, there is no "gap" in the indifference classes between I^{i-1} and I^i and there is a prior, denoted $\tilde{\phi}$ below, such that $I(\tilde{\phi})$ has non-trivial intersection with both I^{i-1} and I^i.

Suppose that $\sum_j \overline{I^{i-1}}(p_j)\widetilde{\nu}(\phi^{i-1}, p_j) = \kappa$. This implies that there are $P, P' \in \mathcal{P}(\tilde{\phi})$ such that $\sum_j P(p_j)\widetilde{\nu}(\tilde{\phi}, p_j) = \kappa$ and $\sum_j P'(p_j)\widetilde{\nu}(\tilde{\phi}, p_j) = \kappa'$ for some $\kappa' < \kappa$. Observe that given its cardinal uniqueness, this fully pins down $\widetilde{\nu}(\tilde{\phi}, \cdot)$. Moreover, observe that there exists a P'' such that $\sum_j P''(p_j)\widetilde{\nu}(\tilde{\phi}, p_j) = \kappa'' > \kappa$.

By construction there exists some $Q \in \mathcal{P}(\phi^i)$ such that $Q \sim P''$. We set $\sum_j \underline{I^i}(p_j)\widetilde{\nu}(\phi^i, p_j) = \kappa$. Moreover, we can find a $\lambda(Q)$ such that $\lambda(Q)\underline{I^i} + (1 - \lambda(Q))\overline{I^i} \sim Q$. Let ι solves $\lambda(Q)\kappa + (1 - \lambda(Q))\iota = \kappa''$ and set $\sum_j \overline{I^i}(p_j)\widetilde{\nu}(\phi^i, p_j) = \iota$.

As in the previous cases, now consider any $\phi' \neq \phi$ such that $I^i \cap \mathcal{P}(\phi') \neq \emptyset$. Let $\underline{R^i(\phi')}$ be a $\succsim$ −minimal element in $I^i \cap \mathcal{P}(\phi')$. Similarly, let $\overline{R^i(\phi')}$ be a $\succsim$ −maximal element in $I^i \cap \mathcal{P}(\phi')$. Find $P \in I^i \cap \mathcal{P}(\phi^i)$ such that $P \sim \underline{R^i(\phi')}$ and set $\sum_j \underline{R^i(\phi')}(p_j)\widetilde{\nu}(\phi', p_j) = \lambda(P)$. Similarly, find $P' \in I^i \cap \mathcal{P}(\phi^i)$ such that $P' \sim \overline{R^i(\phi')}$ and set $\sum_j \overline{R^i(\phi')}(p_j)\widetilde{\nu}(\phi', p_j) = \lambda(P')$.

Note that, by construction, we have $f_\phi(\lambda(P)) = f_{\phi'}(\lambda(P))$ and $f_\phi(\lambda(P')) = f_{\phi'}(\lambda(P'))$. By CTI, for any $\alpha \in [0, 1]$, $\alpha P + (1 - \alpha)P' \sim \alpha\underline{R^i(\phi')} + (1 - \alpha)\overline{R^i(\phi')}$ and thus $f_\phi(\alpha\lambda(P) + (1 - \alpha)\lambda(P')) = f_{\phi'}(\alpha\lambda(P) + (1 - \alpha)\lambda(P'))$. This implies that within I^i we can take, without loss of generality, $f_{\phi'} = f_\phi$. By repeating this same process over all ϕ's that have an element in I^i we can show that for all relevant priors ϕ' we can set $f_{\phi'} = f_\phi$ on I^i. In other words, on each I^i we can take f to be independent of the prior.[22]

We now turn to piecing together the entire utility function. First, recall that we have a countable set $I^1, I^2, ...$ (ordered in terms of increasing preferences). Take the set $\mathbb{G}^1 = \mathcal{I} \backslash \mathbb{I}$. These are the "leftover" equivalence classes that compose the first class we considered. Note that we can also group these together into convex sets. We do so inductively. Pick an I such that I is the worst

[22] Observe that if we mix two lotteries in I^i that induce different priors, then their mixed lottery, which induce some prior ϕ'' either lives in I^i and the same calibration exercise can be performed for ϕ'', or outside of I^i, in which case it should be relegated to the step where we deal with the relevant $I^j \neq I^i$.

equivalence class in $\mathrm{cl}(\mathbb{G})$. Then find $\hat{I} \in \mathrm{cl}(\mathbb{G})$ with the property that for any $P \in \hat{I}$ there is no $I' \in \mathbb{I}$ with $Q \prec P$ for some $Q \in I'$. Then denote G^1 as the set of equivalence classes from I to $\hat{I}$ (inclusive); $G^1 = \{I''|P \succsim Q \succsim R$ for all $P \in \hat{I}$, $Q \in I''$, and $R \in I\}$. Now set $\mathbb{G}^2 = \mathbb{G}^1 \backslash G^1$, and find G^2 using the same process, and continue in this fashion to get a collection $\{G^l\}$.

We have already shown that f is independent of ϕ on any G^l or I^i. On any one of these ranges denote the relevant f as f_j where $j \in \{G^l, I^i\}$. Suppose there is an indifference class I such that I is in two elements of $\{G^l, I^i\}$, call them H and H'. Suppose without loss of generality that for some $P \in I$ with $\sum_j P(p_j)\widetilde{\nu}(\phi, p_j) = a$, we have $f_H(a) \geq f_{H'}(a)$. Then, since both f_H and $f_{H'}$ are strictly increasing, we can simply set $f^*_H(\cdot) = f_{H'}(\cdot) - [f_{H'}(a) - f_H(a)]$, generating continuity at a. Denote the adjusted collection of transformation functions by $\{h_H\}$. We can now simply take a single function f defined by $f = h_H$ on H and performing a monotone transformation to recover our utility function $V(P) = f^{-1}(f(\sum_j P(p_j)\tilde{v}(\phi(P), p_j)))$. □

This proves the equivalence in the proposition. ■

Proof of Proposition 2. First we define a prior-separable representation as $V_{ps} = \nu_{ps1}(\phi(P)) + \sum_i P(p_i)\nu_{ps2}(p_i)$

Claim 3. *The relation $\succsim$ has a prior-separable representation if and only if it has a prior-anticipatory representation.*

Proof of Claim 3. Recall from the proof of Lemma 1 that $\succsim$ has a prior-anticipatory representation if and only if it has a representation $\hat{\nu}_1(\phi(P)) + \sum_i P(p_i)\nu_2(p_i)$. Note that this is simply the sum of a utility function defined over the reduced lottery and a recursive utility that is expected utility in the first stage. □

We now use our new representation for Claim 4.

Claim 4. *The relation $\succsim$ has a prior-separable representation if and only if it satisfies WO, C, CTI, and SPTI.*

Proof of Claim 4. It is easy to check that the axioms are necessary for the representation. For sufficiency, observe that SPTI implies PTI, which, in turns, implies that there exists a representation of the form $\sum_i P(p_i)\widetilde{\nu}(\phi(P), p_i)$. Moreover, by SPTI, if $\sum_i P(p_i)\widetilde{\nu}(\phi(P), p_i) = \sum_i Q(p_i)\widetilde{\nu}(\phi(P), p_i)$, then $\sum_i(\alpha P + (1-\alpha)R)(p_i)\widetilde{\nu}(\phi((\alpha P + (1-\alpha)R)), p_i) = \sum_i(\alpha Q + (1-\alpha)R)(p_i)\widetilde{\nu}(\phi((\alpha P + (1-\alpha)R)), p_i)$ for any R, which is true if and only if $\widetilde{\nu}$ is additively separable in it's first argument: $\sum_i P(p_i)\widetilde{\nu}(\phi(P), p_i) = \hat{\nu}_1(\phi(P)) + \sum_i P(p_i)\nu_2(p_i)$. To see this, observe that with n sub-lotteries, the utility function V_{ps} can be thought of as a function of $n+1$ arguments — the n sub-lotteries and the prior beliefs. Since the representation is additively separable,conditional on the prior, preferences must satisfy separability (i.e, preferential independence in Debreu, 1960) across the sub-lotteries (and all subsets of the sub-lotteries). Further observe that SCTI implies that all subsets of the sub-lotteries and the prior also satisfy separability (preferential independence). Thus, by Debreu (1960) (see also Wakker, 1993) the representation must be additively separable in all components. □

This proves the equivalence in the proposition. ■

Proof of Proposition 3. First we define a prior-separable expected utility representation as $V_{pseu} = \sum_x \phi(P)(x)\nu_{pseu1}(x) + \sum_i P(p_i)\nu_{pseu2}(p_i)$.

Claim 5. *The relation $\succsim$ has a prior-separable expected utility representation if and only if it has a posterior-anticipatory representation.*

Proof of Claim 5. From Lemma 1, $\succsim$ has a posterior-anticipatory representation if and only if it has a representation $\sum_i P(p_i) \sum_j p_i(x_j)\hat{u}(x_j) + \sum_i P(p_i)\nu_1(p_i)$. This is simply the sum of a an expected utility functional defined over the reduced lottery and a recursive utility that is expected utility in the first stage. □

We now use the new representation for Claim 6.

Claim 6. *The relation $\succsim$ has a prior-separable expected utility representation if and only if it satisfies WO, C and TI.*

Observe that by the mixture space theorem, $\succsim$ satisfies WO, C, and TI if and only if it can be represented by the functional $\sum_i P(p_i)\grave{\nu}(p_i)$. Moreover, if preferences can be represented by $\sum_i P(p_i)\grave{\nu}(p_i)$ then clearly they have a prior-separable expected utility representation (where $\nu_{pseu1}(x) = 0$). Similarly, any prior anticipatory representation can be written as $\sum_i P(p_i)\grave{\nu}(p_i)$ where $\grave{\nu}(p_i) = \nu_{pseu2}(p_i) + \sum_x p_i(x)\nu_{pseu1}(x)$.

This proves the equivalence in the proposition. □

Proof of Proposition 4. We prove each of the statements in order.

1. We first show that if $\succsim$ has an ABB representation then it has a BSI representation in a series of two claims.

   Claim 7. *There exists an equivalent representation $(u, \hat{\nu}_1, \hat{\nu}_2)$ which satisfies the condition $\hat{\nu}_1(\rho, \rho) = 0$ for all ρ.*

Proof of Claim 7. Denote as $N(p)$ the number of elements in the support of p and sum up below only amongst those elements with positive probability. Define: $\hat{\nu}_1(\rho, p) = \nu_1(\rho, p) - \nu_1(p, p)$ and $\hat{\nu}_2(p, \delta_x) = \nu_2(p, \delta_x) + \frac{\nu_1(p,p)}{N(p)p(x)}$. By construction, $\hat{\nu}_1(\rho, \rho) = 0$. Moreover, preferences did not change as the new representation gives utility:

$$
\begin{aligned}
& \sum_x \rho(x)u(x) + \sum_p P(p)\hat{\nu}_1(\rho, p) + \sum_p \sum_x P(p)p(x)\hat{\nu}_2(p, \delta_x) \\
= \; & \sum_x \rho(x)u(x) + \sum_p P(p)[\nu_1(\rho, p) - \nu_1(p, p)] + \sum_p \sum_x P(p)p(x)[\nu_2(p, \delta_x) + \frac{\nu_1(p,p)}{N(p)p(x)}] \\
= \; & \sum_x \rho(x)u(x) + \sum_p P(p)\nu_1(\rho, p) - \sum_p P(p)\nu_1(p, p) + \sum_p \sum_x P(p)p(x)\nu_2(p, \delta_x) \\
+ \; & \sum_p P(p)\nu_1(p, p) \sum_x \frac{1}{N(p)} \\
= \; & \sum_x \rho(x)u(x) + \sum_p P(p)\nu_1(\rho, p) - \sum_p P(p)\nu_1(p, p) + \sum_p P(p)\nu_1(p, p) + \sum_p \sum_x P(p)p(x)\nu_2(p, \delta_x)
\end{aligned}
$$

which is the original utility function. $\square$

Claim 8. *There exists an equivalent representation $(\widetilde{u}, \hat{\nu}_1, \widetilde{\nu}_2)$, which satisfies the condition $\widetilde{\nu}_2(\delta_x, \delta_x) = 0$ for all $x \in X$.*

Proof of Claim 8. Define $\widetilde{\nu}_2(p, \delta_x) = \hat{\nu}_2(p, \delta_x) - \hat{\nu}_2(\delta_x, \delta_x)$ and $\widetilde{u}(x) = u(x) + \hat{\nu}_2(\delta_x, \delta_x)$. Note that $\widetilde{\nu}_2(\delta_x, \delta_x) = 0$ for all x. Observe that this does not change preferences since utility under this representation is:

$$
\begin{aligned}
& \sum_x \widetilde{u}(x)\rho(x) + \sum_p P(p)\hat{\nu}_1(\rho, p) + \sum_p \sum_x P(p)p(x)\widetilde{\nu}_2(p, \delta_x) \\
= \; & \sum_x \rho(x)[u(x) + \hat{\nu}_2(\delta_x, \delta_x)] + \sum_p P(p)\hat{\nu}_1(\rho, p) + \sum_p P(p)p(x)[\hat{\nu}_2(p, \delta_x) - \hat{\nu}_2(\delta_x, \delta_x)] \\
= \; & \sum_x \rho(x)u(x) + \sum_p P(p)\hat{\nu}_1(\rho, p) + \sum_p \sum_x P(p)p(x)\hat{\nu}_2(p, \delta_x) \\
+ \; & \sum_x \rho(x)\hat{\nu}_2(\delta_x, \delta_x) - \sum_p \sum_x P(p)p(x)\hat{\nu}_2(\delta_x, \delta_x)
\end{aligned}
$$

which is simply the original utility function.$\square$

Thus, we have a utility representation $(\widetilde{u}, \hat{\nu}_1, \widetilde{\nu}_2)$ which satisfies BSI. $\square$

2. We next show that $\succsim$ always has a representation which is BTI. Take the representation $(\widetilde{u}, \hat{\nu}_1, \widetilde{\nu}_2)$ defined in the previous part. Define

$$\widetilde{\nu}_2'(p, \delta_x) = \widetilde{\nu}_2(p, \delta_x) + [\hat{\nu}_1(p, \delta_x) - \widetilde{\nu}_2(p, \delta_x)]$$

and

$$\hat{\nu}_1'(\rho, p) = \hat{\nu}_1(\rho, p) - \sum_x p(x)[\hat{\nu}_1(p, \delta_x) - \widetilde{\nu}_2(p, \delta_x)]$$

Observe $(\widetilde{u}, \hat{\nu}_1', \widetilde{\nu}_2')$ represents the same preferences. Utility under the second representation is:

$$\begin{aligned}
& \sum_x \widetilde{u}(x)\rho(x) + \sum_p P(p)\hat{\nu}_1'(\rho, p) + \sum_p \sum_x P(p)p(x)\widetilde{\nu}_2'(p, \delta_x) \\
= \; & \sum_x \widetilde{u}(x)\rho(x) + \sum_p P(p)[\hat{\nu}_1(\rho, p) - \sum_x p(x)[\hat{\nu}_1(p, \delta_x) - \widetilde{\nu}_2(p, \delta_x)]] \\
+ \; & \sum_p \sum_x P(p)p(x)[\widetilde{\nu}_2(p, \delta_x) + [\hat{\nu}_1(p, \delta_x) - \widetilde{\nu}_2(p, \delta_x)]] \\
= \; & \sum_x \widetilde{u}(x)\rho(x) + \sum_p P(p)\hat{\nu}_1(\rho, p) - \sum_p \sum_x P(p)p(x)[\hat{\nu}_1(p, \delta_x) - \widetilde{\nu}_2(p, \delta_x)] \\
+ \; & \sum_p \sum_x P(p)p(x)\widetilde{\nu}_2(p, \delta_x) + \sum_p \sum_x P(p)p(x)[\hat{\nu}_1(p, \delta_x) - \widetilde{\nu}_2(p, \delta_x)] \\
= \; & \sum_x \widetilde{u}(x)\rho(x) + \sum_p P(p)\hat{\nu}_1(\rho, p) + \sum_p \sum_x P(p)p(x)\widetilde{\nu}_2(p, \delta_x)
\end{aligned}$$

which are the original preferences.

Moreover, observe that by construction

$$\hat{\nu}_1'(p, \delta_x) = \hat{\nu}_1(p, \delta_x) - [\hat{\nu}_1(\delta_x, \delta_x) - \widetilde{\nu}_2(\delta_x, \delta_x)] = \hat{\nu}_1(p, \delta_x) - [0 - 0]$$

Also

$$\widetilde{\nu}_2'(p, \delta_x) = \widetilde{\nu}_2(p, \delta_x) + [\hat{\nu}_1(p, \delta_x) - \widetilde{\nu}_2(p, \delta_x)] = \hat{\nu}_1(p, \delta_x)$$

Thus we satisfy BTI. However, we no longer satisfy BSI. This is because

$$\hat{\nu}_1'(\rho, \rho) = \hat{\nu}_1(\rho, \rho) - \sum_x \rho(x)[\hat{\nu}_1(\rho, \delta_x) - \widetilde{\nu}_2(\rho, \delta_x)]$$

no longer necessarily equals 0. □

3. We now show that $\succsim$ has a representation which is both BSI and BTI if and only if it satisfies TN.

For the only if part, observe that for $P = D_p$

$$\begin{aligned} V_{ABB}(P) &= E_p(u) + \nu_1(p,p) + \sum_j p(x_j)\nu_2(p,\delta_{x_j}) \\ &= E_p(u) + \sum_j p(x_j)\nu_2(p,\delta_{x_j}) \end{aligned}$$

where the second equality is by BSI.

For $Q = \sum_i p(x_j)\delta_{x_j}$ we have

$$\begin{aligned} V_{ABB}(Q) &= E_p(u) + \sum_j p(x_j)\nu_1(p,\delta_{x_j}) + \sum_j p(x_j)\nu_2(\delta_{x_j},\delta_{x_j}) \\ &= E_p(u) + \sum_j p(x_j)\nu_1(p,\delta_{x_j}) \end{aligned}$$

where the second equality is again by BSI.

By BTI, $\sum_j p(x_j)\nu_2(p,\delta_{x_j}) = \sum_j p(x_j)\nu_1(p,\delta_{x_j})$, which implies $V_{ABB}(P) = V_{ABB}(Q)$, that is, TN is satisfied.

To prove the other direction, we can simply assume preferences satisfy BSI. Observe that time neutrality implies that

$$\sum_x \widetilde{u}(x)\rho(x) + \hat{\nu}_1(\rho,\rho) + \sum_x \rho(x)\widetilde{\nu_2}(\rho,\delta_x) = \sum_x \widetilde{u}(x)\rho(x) + \sum_x \rho(x)\hat{\nu}_1(\rho,\delta_x) + \sum_x \rho(x)\widetilde{\nu_2}(\delta_x,\delta_x)$$

or, taking the fact that BSI holds,

$$\sum_x \rho(x)\widetilde{\nu_2}(\rho,\delta_x) = \sum_x \rho(x)\hat{\nu}_1(\rho,\delta_x)$$

Observe that $\hat{\nu}_1(\rho,\delta_x)$ only appears as a term as part of the sum $\sum_x \rho(x)\hat{\nu}_1(\rho,\delta_x)$. Thus, we cannot separately identify the individual parts of $\sum_x \rho(x)\hat{\nu}_1(\rho,\delta_x)$. Since $\sum_x \rho(x)\widetilde{\nu_2}(\rho,\delta_x) = \sum_x \rho(x)\hat{\nu}_1(\rho,\delta_x)$, we can suppose without loss of generality that $\rho(x)\widetilde{\nu_2}(\rho,\delta_x) = \rho(x)\hat{\nu}_1(\rho,\delta_x)$ term by term. □

This completes the proof of Proposition 4. ■

Lastly, as we mention in Footnote 7, we show that if $\succsim$ has an ABB representation, then it has a representation which is both BSI and PBTI.[23] First, normalize the representation using claims 7 and 8 so that it satisfies BSI. We then normalize the representation so that BTI holds as in the second part of the proof of this proposition. As we have mentioned there, $\hat{\nu}'_1(\rho,\rho)$ no

[23]Since $\succsim$ always has a representation which is BTI, it also has a PBTI representation where $\kappa = 1$.

longer necessarily equals 0. But, since we started with a BSI representation, we already had that $\hat{\nu}_1'(\delta_x, \delta_x) = \widetilde{\nu}_2'(\delta_x, \delta_x) = 0$ so those values do not change.

In order to simplify notation, call the functionals after these two steps u, ν_1, and ν_2 respectively. Thus, $\nu_1 = \nu_2$ over their shared domain, and $\nu_2(\delta_x, \delta_x) = 0 = \nu_1(\delta_x, \delta_x)$.

Now we will define a representation that satisfies both BSI and PBTI. We do this in a way that mirrors Claim 7. Denote as $N(p)$ the number of elements with positive probability in p and sum up below only amongst those elements. Define: $\hat{\nu}_1(\rho, p) = \nu_1(\rho, p) - \nu_1(p, p)$. Importantly, this redefinition implies $\hat{\nu}_1(\rho, \delta_x) = \nu_1(\rho, \delta_x) - \nu_1(\delta_x, \delta_x) = \nu_1(\rho, \delta_x)$.

We then turn to solve for $\hat{\nu_2}$. Denote $z(p) = \nu_1(p, p)$. For our representation to satisfy PBTI we need that $\hat{\nu_2}(p, \delta_x) = \kappa\hat{\nu}_1(p, \delta_x) = \kappa\nu_1(p, \delta_x) = \kappa\nu_2(p, \delta_x)$ for some κ. If p has $N(p)$ outcomes in its support, then these are $N(p)$ equations and $N(p) + 1$ unknowns. We also need it to be the case that $\sum p(x)\hat{\nu_2}(p, \delta_x) = z(p)$. Substituting in we get $\kappa \sum p(x)\nu_2(p, \delta_x) = z(p)$ or $\kappa = \frac{z(p)}{\sum p(x)\nu_2(p,\delta_x)}$. Observe that this uniquely pins down κ and so uniquely pins down $\hat{\nu_2}$ for each p. Thus, PBTI is satisfied. Moreover, observe that by construction $\hat{\nu_2}(\delta_x, \delta_x) = 0$ still and $\hat{\nu}_1(\rho, \rho) = 0$, and so BSI is satisfied as well. □

Proof of Proposition 5. From Claim 1 we know that we can confine attention to a prior-conditional representation of $\succsim$. Observe that fixing $\phi(P)$, $\sum_i P(p_i)\nu_{PC}(\phi(P), p_i)$ is an expected utility functional, and so possesses the same uniqueness results; i.e., it is unique up to affine transformations of scalars $\alpha_P > 0$ and β_P. But, since $\sum_i P(p_i)\nu_{PC}(\phi(P), p_i) \geq \sum_i Q(q_i)\nu_{PC}(\phi(Q), q_i)$ if and only if $\beta_P + \alpha_P \sum_i P(p_i)\nu_{PC}(\phi(P), p_i) \geq \beta_Q + \alpha_Q \sum_i Q(q_i)\nu_{PC}(\phi(Q), q_i)$, it must be the case that $\beta_P = \beta_Q$ and $\alpha_P = \alpha_Q$. ■

Proof of Proposition 6. To see the result for a BSI representation, first take the uniqueness result for general ABB preferences (Proposition 12 in Appendix A.2). Suppose that (u, ν_1, ν_2) is a BSI representation. We first show that any non-constant transformation γ_1, with $\gamma_1(x) \neq 0$ for some x, cannot generate a BSI representation. Suppose that there is some x_i such that $\gamma_1(x_i) \neq 0$. Consider the two-stage lottery $D_{\delta_{x_i}}$. Then $\nu_1'(\delta_{x_i}, \delta_{x_i}) = 0 - \gamma_1(x_i) \neq 0$ so this cannot be a BSI representation. Next we show that any non-constant transformation γ_u, with $\gamma_u(x) \neq 0$ for some x, cannot generate a BSI representation. Suppose that there is some x_i such that $\gamma_u(x_i) \neq 0$. Consider the two-stage lottery $D_{\delta_{x_i}}$. Then $\nu_2'(\delta_{x_i}, \delta_{x_i}) = 0 - \gamma_u(x_i) \neq 0$ so this cannot be a BSI representation. For similar resasons $\beta_1 = \beta_2 = 0$. Lastly, we show that any transformation where $\sum_x p(x)\gamma_\nu(p, \delta_x) \neq 0$ for some p cannot generate a BSI representation. If there were p_i such that this were true, then $v_1'(p, p) = 0 + \sum_x p(x)\gamma_\nu(p, \delta_x) \neq 0$ violating the BSI representation. Moreover, if $\sum_x p(x)\gamma_\nu(p, \delta_x) = 0$ then $v_1' = v_1$ and so we exclude it from consideration. ■

Proof of Proposition 7. We first show that $\succsim$ has a posterior-separable expected utility representation (i.e., it satisfies WO, C, and TI) if and only if it satisfies WO, C, PTI, CTI, and R.

Necessity is immediate. To show sufficiency, note that $\succsim$ has a representation of the form $V_{PC} = \sum_i P(p_i)\nu_{PC}(\phi(P), p_i)$ if it satisfies WO, C, PTI and CTI. If R is satisfied, then it must be the case that ν_{PC} is independent of the first argument. Thus we have a representation of the form $\sum_i P(p_i)\hat{\nu}_{PC}(p_i)$, which is equivalent to the following representation: $\sum_i P(p_i)\sum_j p_i(x_j)\hat{u}(x_j) + \sum_i P(p_i)\nu_1(p_i)$.

Recall that $\succsim$ has a posterior-anticipatory representation if and only if it has a representation $\sum_i P(p_i)\grave{\nu}(p_i)$. By Segal (1990), this is a recursive representation where V_1 is expected utility. This representation clearly satisfies WO, C, and TI. ■

Proof of Proposition 8. For item (1), observe that we can ignore the first term of the ABB representation, as it is the same under any two compound lotteries with the same reduced form probabilities. Suppose then that $\nu_1(\rho, \cdot) + \sum_x \nu_2(\cdot, x)$ is convex. Then, by Grant, Kajii and Polak (1998) the individual must exhibit a preference for early resolution of uncertainty. Conversely, if the term above is not convex, then it must be concave in a local neighborhood of some p_i. We can replicate the argument in Grant, Kajii and Polak (1998). Take some compound lottery that delivers as one sub-lottery p_i, and take a linear bifurcation of p_i so that the new sub-lotteries are arbitrarily close to p_i. Then by Grant, Kajii and Polak (1998) the individual must be worse off (since locally the utility function is concave).

For item (2), take any P and Q as specified in the statement of the proposition. Observe that $\phi(P) = \phi(Q)$. Direct calculations then show that $P \succsim Q$ if and only if $\beta\acute{\nu_2}(p_1) + (1-\beta)\acute{\nu_2}(p_2) \geq \acute{\nu_2}(\beta p_1 + (1-\beta)p_2)$. And since the triple p_1, p_2, and β were arbitrary, the inequality holds if and only if $\acute{\nu_2}$ is convex. Similarly, the inequality is reversed if and only if $\acute{\nu_2}$ is concave.

For item (3), simply replace in the entire paragraph above $\acute{\nu_2}$ with $\bar{\nu_1}$. ■

Proof of Proposition 9. For item (i), note that by Definition 7, PORU implies TN. By Proposition 4, the representation satisfies both BSI and BTI. We now use these functional form restrictions when calculating the values of the lotteries in question. We have $V(D_p) = V(\sum_i p(x_i)\delta_{x_i}) = \sum_j \phi(P)(x_j)u(x_j) + \sum_x \phi(P)(x)\nu_1(\phi(P), \delta_x)$, which is the left hand side of the inequality in (i) in addition to the expected utility from material payoffs. PORU implies that these two compound lotteries are better than any other P with the same reduced probabilities $\phi(P)$. Indeed, the value of any such P is $V(p) = \sum_j \phi(P)(x_j)u(x_j) + \sum_i P(p_i)\nu_1(\phi(P), p_i) + \sum_{p_i} P(p_i)\sum_x p_i(x)\nu_1(p_i, \delta_x)$, which is the right hand side of the inequality in addition to the same expected utility from material payoffs.

For item (ii), first recall that if $\succsim$ has a prior-anticipatory representation, then it can also be represented by the functional $\hat{\nu}_1(\phi(P)) + \sum_i P(p_i)\nu_2(p_i)$. Corollary 1 below shows that in this case TN implies that ν_2 is an expected utility functional, and thus the second term does not generate any anomalous preferences towards information. Clearly the first term $\hat{\nu}_1(\phi(P))$ depends only on the prior beliefs, independently of the pattern of resolution of uncertainty. The individual is thus indifferent among all lotteries that induce the same prior beliefs, and in particular cannot display strict PORU. ■

Proof of Proposition 10. We prove each item in turn. Our technique for proving the first two parts is to show that "path by path" (described more below) the inequality that Proposition 9 points out is sufficient (as well as necessary) for PORU. Therefore, it must also hold after averaging across all paths.

1. We show that Kőszegi and Rabin (2009) preferences with linear gain-loss utility satisfy PORU. The gain loss utility experienced at percentile ψ by moving from belief p to q is $\mu(u(c_q(\psi) - u(c_p(\psi)))$. Consider the two-stage lotteries P and Q, where P is one-shot (fully resolves in the first stage), Q gradually resolves, and both induce same prior beliefs over final outcomes. Because expected consumption utility is the same in both P and Q, we only need to consider gain-loss utility. Conditional on x_i being realized, the gain-loss utility at percentile ψ in P is given by: $\mu(u(x_i) - u(c_{\phi(P)}(\psi)))$. Suppose $u(x_i) - u(c_{\phi(P)}(\psi))) > 0$ (the other case is analogous).[24] Observe that every time x_i is realized for Q, it happens via some single stage lottery p_k. Then moving from $\phi(P)$ to p_k and from there to δ_{x_i} in Q gives gain-loss utility $A := \mu(u(x_i)) - u(c_{p_k}(\psi))) + \mu(u(c_{p_k}(\psi)) - u(c_{\phi(P)}(\psi)))$. There are four possibilities to consider.

    - $u(c_{p_k}(\psi)) = u(x_i)$. Then $A = \mu(u(x_i) - u(c_{\phi(P)}(\psi)))$, so the utility is the same as in P.
    - $u(c_{p_k}(\psi)) > u(x_i)$. Then $A = \mu(u(x_i)) - u(c_{p_k}(\psi))) + \mu(u(c_{p_k}(\psi)) - u(x_i)) + \mu(u(x_i) - u(c_{\phi(P)}(\psi))) < \mu(u(x_i) - u(c_{\phi(P)}(\psi)))$. Thus, at ψ, the utility of Q is less than that of P.
    - $u(c_{\phi(P)}(\psi)) \leq u(c_{p_k}(\psi)) < u(x_i)$. Then $A = \mu(u(x_i) - u(c_{\phi(P)}(\psi)))$, so the utility is the same as in P.
    - $u(c_{\phi(P)}(\psi)) > u(c_{p_k}(\psi))$. Then $A = \mu(u(c_{p_k}(\psi)) - u(c_{\phi(P)}(\psi))) + \mu(u(c_{\phi(P)}(\psi)) - u(c_{p_k}(\psi))) + \mu(u(x_i) - u(c_{\phi(P)}(\psi))) < \mu(u(x_i) - u(c_{\phi(P)}(\psi)))$. Thus, at ψ, the utility of Q is less than that of P.

    Combining the results above, we found that, at percentile ψ, the utility of Q is less than that of P. Observe that the total probability attached to receiving x_i, $\phi(P)(x_i)$ is the same across the two compound lotteries. Therefore, aggregating across different beliefs paths from the prior to p_k and then to δ_{x_i} in Q results in the same probability as of moving directly from

[24] If instead $u(x_i) = u(c_{\phi(P)}(\psi))$ then only the first, second, and fourth bullet points below can be relevant, but otherwise the same results hold.

the prior to δ_{x_i} in P. We conclude that the gain-loss utility from Q is lower than that of Q. Since ψ was arbitrary, the result follows. $\square$

2. To prove this part, we keep looking at the environment of part (1) but confine attention to the case of only two final prizes. We then go through each of the four cases again.

   - $u(c_{p_k}(\psi)) = u(x_i)$. Then $A = \mu(u(x_i) - u(c_{\phi(P)}(\psi)))$, so the utility is the same as in P.
   - $u(c_{p_k}(\psi)) > u(x_i)$. Note that if μ' is large enough in the loss domain, then $\mu(u(x_i)) - u(c_{p_k}(\psi)))$ is arbitrarily negative and so does A. Hence the utility of Q is less than that of P.
   - $u(c_{\phi(P)}(\psi)) \leq u(c_{p_k}(\psi)) < u(x_i)$. Then, if we let $\mu(u(x_i) - u(c_{\phi(P)}(\psi))) = G$, by concavity of the gain function $A \leq 2G$.
   - $u(c_{\phi(P)}(\psi)) > u(c_{p_k}(\psi))$. If μ' is large enough in the loss domain, then $\mu(u(c_{p_k}(\psi)) - u(c_{\phi(P)}(\psi)))$ is arbitrarily negative and so does A. Hence the utility of Q is less than that of P.

   Observe that the only time that a gain occurs from gradual resolution is in the third case. However, anytime the prior belief ϕ moves to a non-degenerate p_k such that at some percentile the third case is effective, with strictly positive probability either the second case or the fourth case must also occur (for a different x_i). In other words, whenever a gain occurs, a loss must occur as well. By taking μ' to be large enough in the loss domain, we can always make the losses arbitrarily large to overwhelm any of the gains.

   We now turn to considering the case where $u(x_i) - u(c_{\phi(P)}(\psi))) < 0$ (unlike in the piecewise linear case, here it less obvious that the proof which worked for gains will also work for losses). There are again four cases to consider.

   - $u(c_{p_k}(\psi)) = u(x_i)$. Then the utility is the same as in P.
   - $u(c_{p_k}(\psi)) < u(x_i)$. Since losses are larger than gain (μ' is sufficiently large in the loss domain), $A < \mu(u(x_i) - u(c_{\phi(P)}(\psi)))$.
   - $u(c_{\phi(P)}(\psi)) \geq u(c_{p_k}(\psi)) > u(x_i)$. Then $A = \mu(u(x_i) - u(c_{p_k}(\psi))) + \mu(u(c_{p_k}(\psi)) - u(c_{\phi(P)}(\psi)))$, which by the convexity of μ in the loss domain must be more negative than $\mu(u(x_i) - u(c_{\phi(P)}(\psi)))$.
   - $u(c_{\phi(P)}(\psi)) < u(c_{p_k}(\psi))$. Since losses are larger than gain, $A < \mu(u(x_i) - u(c_{\phi(P)}(\psi)))$.

   Observe that there are no gains from gradual resolution for any of these four cases.

   Thus the gain-loss utility, and hence the total utility, of Q is lower than that of P. $\square$

3. Given a loss function, we first consider a gain function that is exactly symmetric ($-\mu(-x) = \mu(x)$). Suppose prior beliefs are uniform over some three prizes with corresponding utility values $z_0, z_1, z_2 \in \Delta$ such that $z_2 - z_1 = z_1 - z_0 > 0$. Slightly abuse notation, we will identify below an outcome with its utility value (or assume, without loss of generality, that u is linear). Take $p_1 = \delta_{z_0}$, $p_2 = \frac{1}{2}\delta_{z_1} + \frac{1}{2}\delta_{z_2}$, and let $Q = \frac{1}{3}D_{p_1} + \frac{2}{3}D_{p_2}$.

   Consider an early resolved compound lottery $P = \frac{1}{3}D_{\delta_{z_0}} + \frac{1}{3}D_{\delta_{z_1}} + \frac{1}{3}D_{\delta_{z_2}}$. That is, beliefs according to P in Period 1 are either degenerate on z_0, z_1, or z_2. Gain-loss utility if z_0 is realized is then

$$\int_0^1 \mu(0-c(\phi))dp = \int_0^{\frac{1}{3}} \mu(0)dp + \int_{\frac{1}{3}}^{\frac{2}{3}} \mu(z_0-z_1)dp + \int_{\frac{2}{3}}^1 \mu(z_0-z_2)dp = \frac{1}{3}\mu(z_0-z_1) + \frac{1}{3}\mu(z_0-z_2);$$

   if z_1 is realized, gain-loss utility is

$$\int_0^1 \mu(1-c(\phi))dp = \int_0^{\frac{1}{3}} \mu(z_1-z_0)dp + \int_{\frac{1}{3}}^{\frac{2}{3}} \mu(0)dp + \int_{\frac{2}{3}}^1 \mu(z_1-z_2)dp = \frac{1}{3}\mu(z_1-z_0) + \frac{1}{3}\mu(z_1-z_2);$$

   and if z_2 is realized, gain-loss utility is

$$\int_0^1 \mu(2-c(\phi))dp = \int_0^{\frac{1}{3}} \mu(z_2-z_0)dp + \int_{\frac{1}{3}}^{\frac{2}{3}} \mu(z_2-z_1)dp + \int_{\frac{2}{3}}^1 \mu(0)dp = \frac{1}{3}\mu(z_2-z_0) + \frac{1}{3}\mu(z_2-z_1).$$

   Aggregating, and taking into account the probability of each realization and that the z values are equally spaced, we get the gain-loss utility for P is

$$\begin{aligned}\frac{1}{9}\mu(z_0-z_1) + \frac{1}{9}\mu(z_0-z_2) + \frac{1}{9}\mu(z_1-z_0) + \frac{1}{9}\mu(z_1-z_2) + \frac{1}{9}\mu(z_2-z_0) + \frac{1}{9}\mu(z_2-z_1) \\ = \frac{2}{9}\mu(z_0-z_1) + \frac{2}{9}\mu(z_1-z_0) + \frac{1}{9}\mu(z_0-z_2) + \frac{1}{9}\mu(z_2-z_0).\end{aligned}$$

   The gain-loss utility for Q in period 1 is

$$\begin{aligned} & \frac{1}{3}[\int_0^{\frac{1}{3}} \mu(0)d\psi + \int_{\frac{1}{3}}^{\frac{2}{3}} \mu(z_0-z_1)d\psi + \int_{\frac{2}{3}}^1 \mu(z_0-z_2)d\psi] \\ + & \frac{2}{3}[\int_0^{\frac{1}{3}} \mu(z_1-z_0)d\psi + \int_{\frac{1}{3}}^{\frac{1}{2}} \mu(0)d\psi + \int_{\frac{1}{2}}^{\frac{2}{3}} \mu(z_2-z_1)d\psi + \int_{\frac{2}{3}}^1 \mu(0)d\psi] \\ = & \frac{1}{3}[\frac{1}{3}\mu(z_0-z_1) + \frac{1}{3}\mu(z_0-z_2)] + \frac{2}{3}[\frac{1}{3}\mu(z_1-z_0) + \frac{1}{6}\mu(z_2-z_1)]\end{aligned}$$

In period 2 the gain loss utility from p_1 is 0. The gain loss utility from p_2 is

$$\frac{1}{2}[\int_0^{\frac{1}{2}}\mu(z_2-z_1)d\psi+\int_{\frac{1}{2}}^1\mu(0)d\psi]+\frac{1}{2}[\int_0^{\frac{1}{2}}\mu(0)d\psi+\int_{\frac{1}{2}}^1\mu(z_1-z_2)d\psi]=\frac{1}{4}\mu(z_2-z_1)+\frac{1}{4}\mu(z_1-z_2)$$

Total gain-loss utility from Q is then

$$\begin{aligned}\frac{1}{9}\mu(z_0-z_1)+\frac{1}{9}\mu(z_0-z_2)+\frac{2}{9}\mu(z_1-z_0)+\frac{1}{9}\mu(z_2-z_1)+\frac{1}{6}\mu(z_2-z_1)+\frac{1}{6}\mu(z_1-z_2)\\=\frac{5}{18}\mu(z_0-z_1)+\frac{1}{9}\mu(z_0-z_2)+\frac{1}{2}\mu(z_1-z_0)\end{aligned}$$

Subtracting the total utility of P from the total utility of Q we are left with

$$\frac{1}{18}\mu(z_0-z_1)+\frac{5}{18}\mu(z_1-z_0)-\frac{1}{9}\mu(z_2-z_0)$$

By symmetry this is arbitrarily close to ,

$$\frac{2}{9}\mu(z_1-z_0)-\frac{1}{9}\mu(z_2-z_0)=\frac{1}{9}\mu(z_1-z_0)+\frac{1}{9}\mu(z_2-z_1)-\frac{1}{9}\mu(z_2-z_0)\geq 0$$

Where the last inequality follows by the concavity of μ in the gains domain and that the z values are equally spaced. □

This completes the proof of the proposition. ■

Proof of Observation 2.

We first prove the claim for the case of two possible final outcomes, so that beliefs can be summarized by a number. Let Q be a lottery that fully resolves in the second stage. We denote the prior belief by ϕ. Let $W(\phi,p)=\nu_1(\phi,p)+\sum_x p(x)\nu_2(p,\delta_x)$ denote the value of getting sublottery p, given prior ϕ, and then from p resolving in period 2. Since we assume that $\succsim$ exhibits PORU, TN must hold and thus $\nu_1=\nu_2$, which we simply denote by ν. Using item (3) of Proposition 4, the ABB representation is also belief stationarity invariant, and we have $W(\phi,1)=\nu(\phi,1)$ and $W(\phi,0)=\nu(\phi,0)$.

Observe that the value of Q is equal to $\phi W(\phi,1)+(1-\phi)W(\phi,0)=\phi\nu(\phi,1)+(1-\phi)\nu(\phi,0)=W(\phi,\phi)$. A necessary condition for $\succsim$ to exhibit PORU is that for P with support p,p' such that $P(p)W(\phi,p)+P(p)W(\phi,p)\leq W(\phi,\phi)$. This immediately implies that for any p in a small neighbourhood around ϕ, $W(\phi,p)$ must lie below the line segment connecting $W(\phi,1)$ and $W(\phi,0)$, which we denote by L.

In particular, the slope of $W(\phi,\bullet)$ must be both weakly larger than the slope of L at ϕ from below, and weakly lower than the slope of L at ϕ from above.

In the derivations below we take the derivative of functions with respect to the second element, which we call p; and then evaluate the derivative at r. We are interesting in the case where r converges to ϕ. We have already established that $\lim_{r\to\phi^+} \frac{\partial W(\phi,p)}{\partial p}|_r \leq \frac{W(\phi,1)-W(\phi,0)}{1}$. The LHS of this is equal to $\lim_{r\to\phi^+} \frac{\partial \nu(\phi,p)}{\partial p} + \frac{\partial p\nu(p,1)}{\partial p} + \frac{\partial(1-p)\nu(p,0)}{\partial p}|_r$. Using the chain rule, the condition becomes

$$\lim_{r\to\phi^+} \frac{\partial \nu(\phi,p)}{\partial p} + \nu(p,1) + p\frac{\partial \nu(p,1)}{\partial p} - \nu(p,0) + (1-p)\frac{\partial \nu(p,0)}{\partial p}|_r \leq \nu(\phi,1) - \nu(\phi,0),$$

or

$$\lim_{r\to\phi^+} \frac{\partial \nu(\phi,p)}{\partial p} + \phi\frac{\partial \nu(p,1)}{\partial p} + (1-\phi)\frac{\partial \nu(p,0)}{\partial p}|_r \leq 0$$

Similarly, we need

$$\lim_{r\to\phi^-} \frac{\partial \nu(\phi,p)}{\partial p} + \phi\frac{\partial \nu(p,1)}{\partial p} + (1-\phi)\frac{\partial \nu(p,0)}{\partial p}|_r \geq 0$$

Suppose ν is smooth in it's first argument; then

$$\lim_{r\to\phi^-} \frac{\partial \nu(\phi,p)}{\partial p}|_r \geq \lim_{r\to\phi^+} \frac{\partial \nu(\phi,p)}{\partial p}|_r$$

This implies concavity of ν at the reference point.

If there are more than two outcomes, we replicate the previous steps, but now for each directional derivative with respect to the second argument of $\nu(\phi,p)$, and then evaluating at ϕ. Thus, each directional derivative must be decreasing, (i.e., the function is concave in that direction). The result follows since a function is concave if and only if all its directional derivatives are falling. ■

Proof of Proposition 11. We prove each item in turn.

(a) We prove the main statement; the respective one is proved analogously. Fix some natural number $n \geq 2$ and corresponding n outcomes. Fix the probabilities of any $n-2$ of the outcomes chosen, and confine attention to lotteries over the remaining two, x_1 and x_2 (fixing the mass needed to assign for them). Because preferences are strictly monotone, they must be differentiable almost everywhere along the one-dimensional line representing all lotteries over x_1 and x_2. Pick some p along this line such that V_2 is smooth at p.

Consider a two-stage lottery D_p. Now, take a linear bifurcation of p, which generates a compound lottery Q_ϵ, so that Q_ϵ yields with probability $\frac{1}{2}$ q^1_ϵ and with probability $\frac{1}{2}$ q^2_ϵ.

We suppose (as part of our construction) that (i) as $\epsilon \to 0$, q^1_ϵ and q^2_ϵ both converge to p; and (ii) q^1_ϵ is always preferred to q^2_ϵ (that is, $D_{q^1_\epsilon} \succ D_{q^2_\epsilon}$).

Since there are only two outcomes and preferences are monotone, there is a bijection between any one-stage lottery q and its expected value, $e(q)$, and between q and its certainty equivalent (calculated using either V_i), $CE_{V_i}(q)$. Thus, there exists a bijection between $e(q)$ and $CE_{V_i}(q)$, which we denote by $f_{V_i} = CE_{V_i}(e^{-1})$. Since the mapping from lotteries to expected values is smooth (and so it's inverse is smooth) and, by assumption, the mapping from lotteries to certainty equivalents (and so it's inverse) is smooth, we have that f is smooth.

Therefore, we can represent preferences over Q_ϵ with $V(Q_\epsilon) = V_1(\frac{1}{2}CE_{V_2}(e^{-1}(e(q^1_\epsilon))) + \frac{1}{2}CE_{V_2}(e^{-1}(e(q^2_\epsilon)))) = V_1(\frac{1}{2}f_{V_2}(e(q^1_\epsilon) + \frac{1}{2}f_{V_2}(e(q^2_\epsilon))$. Observe that this is exactly the situation considered in Section 5 of Segal and Spivak (1990). They show that if V_1 displays first-order risk aversion, then for small enough ϵ, the individual must strictly prefer D_p to Q_ϵ and, in particular, cannot exhibit preference for early resolution of uncertainty. $\square$

(b) By item (ii) of Proposition 8, we need to have convex $\hat{\nu_2}$. Consider RDU preferences. By Proposition 4 of Segal and Spivak (1990), RDU exhibits first-order risk aversion whenever the probability weighting function is strictly convex. By Wakker (1994) the RDU functional is convex if and only if the probability weighting function is convex. So taking any $\hat{\nu_2}$ with this property, as mentioned in the main text, will do. $\square$

This completes the proof of Proposition 11. $\blacksquare$

A.2 Uniqueness of ABB representations

In this section we present more detailed uniqueness results.

Proposition 12. *Suppose $\succsim$ has an ABB representation (u, ν_1, ν_2). The ABB representation (u', ν'_1, ν'_2) also represents $\succsim$ if and only if there exists scalars $\alpha > 0, \beta_u, \beta_1, \beta_2$, and continuous functions $\gamma_u : X \to \mathbb{R}$, $\gamma_1 : X \to \mathbb{R}$, and $\gamma_\nu : \Delta \times X \to \mathbb{R}$ such that*

- $u'(x) = \alpha u(x) + \beta_u + \gamma_u(x) + \gamma_1(x)$
- $\nu'_1(\rho, p) = \alpha\nu_1(\rho, p) + \beta_1 + \sum_x p(x)\gamma_\nu(p, \delta_x) - \sum_x \rho(x)\gamma_1(x)$
- $\nu'_2(p, \delta_x) = \alpha\nu_2(p, \delta_x) + \beta_2 - \gamma_\nu(p, \delta_x) - \gamma_u(x)$

Proof of Proposition 12. We first show that if

- $u'(x) = \alpha u(x) + \beta_u + \gamma_u(x) + \gamma_1(x)$

- $\nu_1'(\rho,p) = \alpha\nu_1(\rho,p) + \beta_1 + \sum_x p(x)\gamma_\nu(p,\delta_x) - \sum_x \rho(x)\gamma_1(x)$
- $\nu_2'(p,\delta_x) = \alpha\nu_2(p,\delta_x) + \beta_2 - \gamma_\nu(p,\delta_x) - \gamma_u(x)$

then (u',ν_1',ν_2') represents the same preferences as (u,ν_1,ν_2).

Consider the utility function generated by the former representation.

$$
\begin{aligned}
& \sum_x u'(x)\rho(x) + \sum_p P(p)\nu_1'(\rho,p) + \sum_p\sum_x P(p)p(x)\nu_2'(p,\delta_x) \\
= \; & \sum_x \rho(x)[\alpha u(x) + \beta_u + \gamma_u(x) + \gamma_1(x)] \\
+ \; & \sum_p P(p)[\alpha\nu_1(\rho,p) + \beta_1 + \sum_x p(x)\gamma_\nu(p,\delta_x) - \sum_x \rho(x)\gamma_1(x)] \\
+ \; & \sum_p\sum_x P(p)p(x)[\alpha\nu_2(p,\delta_x) + \beta_2 - \gamma_\nu(p,\delta_x) - \gamma_u(x)] \\
= \; & \alpha\sum_x \rho(x)u(x) + \beta_u + \sum_x \rho(x)\gamma_u(x) + \sum_x \rho(x)\gamma_1(x) \\
+ \; & \alpha\sum_p P(p)\nu_1(\rho,p) + \beta_1 + \sum_p\sum_x P(p)p(x)\gamma_\nu(p,\delta_x) - \sum_p P(p)\sum_x \rho(x)\gamma_1(x) \\
+ \; & \alpha\sum_p\sum_x P(p)p(x)\nu_2(p,\delta_x) + \beta_2 - \sum_p\sum_x P(p)p(x)\gamma_\nu(p,\delta_x) - \sum_p\sum_x P(p)p(x)\gamma_u(x)
\end{aligned}
$$

Denoting $\beta = \beta_u + \beta_1 + \beta_2$ and recalling that $\sum_p\sum_x P(p)p(x) = \sum_x \rho(x)$ we get

$$
\begin{aligned}
& \alpha[\sum_x \rho(x)u(x) + \sum_p P(p)\nu_1(\rho,p) + \sum_p\sum_x P(p)p(x)\nu_2(p,\delta_x)] + \beta \\
+ \; & \sum_x \rho(x)\gamma_u(x) + \sum_p\sum_x P(p)p(x)\gamma_\nu(p,\delta_x) + \sum_x \rho(x)\gamma_1(x) \\
- \; & \sum_p\sum_x P(p)p(x)\gamma_\nu(p,\delta_x) - \sum_x \rho(x)\gamma_u(x) - \sum_x \rho(x)\gamma_1(x) \\
= \; & \alpha[\sum_x \rho(x)u(x) + \sum_p P(p)\nu_1(\rho,p) + \sum_p\sum_x P(p)p(x)\nu_2(p,\delta_x)] + \beta
\end{aligned}
$$

which clearly are the same preferences as (u,ν_1,ν_2).

To prove the other direction, suppose (u,ν_1,ν_2) and (u',ν_1',ν_2') represent the same preferences.

Define $\hat{u}(x) = u(x) - u(x)$; $\hat{\nu}_2(p,\delta_x) = \nu_2(p,\delta_x) - \nu_2(p,\delta_x)$; and $\hat{\nu}_1(\rho,p) = \nu_1(\rho,p) + \sum_x \rho(x)u(x) + \sum_x p(x)\nu_2(p,\delta_x)$. These represent the same preferences as (u,ν_1,ν_2) but we can write $V(P) = \sum_p P(p)\hat{\nu}_1(\phi(P),p)$.

Now define $\hat{u}'(x) = u'(x) - u'(x)$; $\hat{\nu}_2'(p,\delta_x) = \nu_2'(p,\delta_x) - \nu_2'(p,\delta_x)$; and $\hat{\nu}_1'(\rho,p) = \nu_1'(\rho,p) + \sum_x \rho(x)u'(x) + \sum_x p(x)\nu_2'(p,\delta_x)$. These represent the same preferences as (u',ν_1',ν_2') but we can write $V'(P) = \sum_p P(p)\hat{\nu}_1'(\phi(P),p)$.

Since $V(P) = \sum_p P(p)\hat{\nu}_1(\phi(P),p)$ and $V'(P) = \sum_p P(p)\hat{\nu}_1'(\phi(P),p)$ we know that $\hat{\nu}_1'(\phi(P),p)$

must be an affine transformation of $\hat{\nu}_1(\phi(P),p)$; so that $\hat{\nu}_1'(\phi(P),p) = \alpha\hat{\nu}_1(\phi(P),p) + \beta$. Thus $V'(P) = \sum_p P(p)\alpha\hat{\nu}_1(\phi(P),p) + \beta$. Clearly, $\sum_p \alpha P(p)\hat{\nu}_1(\phi(P),p) + \beta$ has an ABB representation $(\alpha u + \beta_u, \alpha\nu_1 + \beta_1, \alpha\nu_2 + \beta_2)$, where $\beta_u + \beta_1 + \beta_2 = \beta$.

By construction $\alpha\hat{u} = \hat{u}' = 0$ and $\alpha\hat{\nu}_2 = \hat{\nu}_2' = 0$. Thus we can say $u'(x) = u'(x) - \alpha u(x) + \alpha u(x)$; $\nu_2'(p,\delta_x) = \nu_2'(p,\delta_x) - \alpha\nu_2(p,\delta_x) + \alpha\nu_2(p,\delta_x)$; and $\nu_1'(\phi(P),p) = \alpha\nu_1(\phi(P),p) - \sum_x \rho(x)[u'(x) - \alpha u(x)] - \sum_x p(x)[\nu_2'(p,\delta_x) - \alpha\nu_2(p,\delta_x)] + \beta$. Moreover, it is easy to verify that we can arbitrarily divide β among the terms.

Define $\gamma_u(x) = 0$; $\gamma_1(x) = u'(x) - \alpha u(x)$; and $\gamma_\nu(p,\delta_x) = -[\nu_2'(p,\delta_x) - \alpha\nu_2(p,\delta_x)]$. Then $u'(x) = \alpha u(x) + \gamma_u(x) + \gamma_1(x)\beta_u$; $\nu_2'(p,\delta_x) = \alpha\nu_2(p,\delta_x) - \gamma_\nu(p,\delta_x) - \gamma_u(x) + \beta_1$ and $\nu_1'(\phi(P),p) = \alpha\nu_1(\phi(P),p) + \sum_x p(x)\gamma_\nu(p,\delta_x) - \sum_x \rho(x)\gamma_1(x)$. Thus we have constructed the transformation. ■

For completeness, we now show that if we suppose utility depends only on the levels of beliefs, stronger uniqueness results also obtain. In this case, both belief-based functionals are unique up to expected utility preferences. Thus, the only part of the utility function not uniquely identified (up to standard transformations) are an individual's expected utility attitudes towards final outcomes.

Proposition 13. *Suppose $\succsim$ has an anticipatory representation (u,ν_1,ν_2). The ABB representation (u',ν_1',ν_2') also represents $\succsim$ if and only if there are scalars $\alpha > 0, \beta_u, \beta_1, \beta_2$ and continuous functions $\gamma_u : X \to \mathbb{R}$ and $\gamma_\nu : X \to \mathbb{R}$ such that*

- $u'(x) = \alpha u(x) + \beta_u + \gamma_u(x) + \gamma_\nu(x)$
- $\nu_1'(\rho) = \alpha\nu_1(\rho) + \beta_1 - \sum_x \rho(x)\gamma_\nu(x)$
- $\nu_2'(p) = \alpha\nu_2(p) + \beta_2 - \sum p(x)\gamma_u(x)$[25]

Proof of Proposition 13. The proof is analogous to the one of Proposition 12. We show necessity for prior anticipatory preferences as an example. Consider the utility function generated by the latter representation:

$$\begin{aligned}
& \sum_x u'(x)\rho(x) + \sum_p P(p)\nu_1'(\rho) + \sum_p\sum_x P(p)p(x)\nu_2'(p) \\
= & \sum_x \rho(x)[\alpha u(x) + \beta_u + \gamma_u(x) + \gamma_\nu(x)] + \sum_p P(p)[\alpha\nu_1(\rho) + \beta_1 - \sum_x \rho(x)\gamma_\nu(\delta_x)] \\
+ & \sum_p\sum_x P(p)p(x)[\alpha\nu_2(p) + \beta_2 - \gamma_u(x)] \\
= & \alpha\sum_x \rho(x)u(x) + \beta_u + \sum_x \rho(x)\gamma_u(x) + \sum_x \rho(x)\gamma_\nu(x) + \alpha\nu_1(\rho) + \beta_1 - \sum_x \rho(x)\gamma_\nu(\delta_x) \\
+ & \sum_p P(p)\alpha\nu_2(p) + \beta_2 - \sum_p\sum_x P(p)p(x)\gamma_u(x)
\end{aligned}$$

[25] Observe that in a posterior-anticipatory representation p is a degenerate lottery.

Denoting $\beta = \beta_u + \beta_1 + \beta_2$ and recalling that $\sum_p \sum_x P(p)p(x) = \sum_x \rho(x)$ we get

$$\alpha[\sum_x \rho(x)u(x) + \nu_1(\rho) + \sum_p P(p)\alpha\nu_2(p)] + \beta$$

which clearly are the same preferences as (u, ν_1, ν_2). ■

A.3 Other Properties

In addition to axiom R, Segal (1990) also introduced several other restrictions on preferences over compound lotteries. We first quickly review them. The strongest assumption is called Reduction of Compound Lotteries (ROCL), which supposes that individuals only care about the reduced form probabilities of any given compound lottery.

Reduction of Compound Lotteries (ROCL): *For all $P, Q \in \Delta^2$, if $\phi(P) = \phi(Q)$ then $P \sim Q$.*

To introduce the next assumption, we first define two special subsets of Δ^2.

- $\Gamma = \{D_p | p \in \Delta\}$, the set of degenerate lotteries in Δ^2. Γ is the set of late resolving lotteries.
- $\Lambda = \{Q \in \Delta^2 | Q(p) > 0 \Rightarrow p = \delta_x \text{ for some } x \in X\}$, the set of compound lotteries whose outcomes are degenerate in Δ. Λ is the set of early resolving lotteries.

We define the restriction of $\succsim$ to the subsets Γ and Λ as $\succsim_\Gamma$ and $\succsim_\Lambda$, respectively.[26] Assumption (I) imposes Independence on these two induced relations.

Independence (I): *The relations $\succsim_\Gamma$ and $\succsim_\Lambda$ satisfy Independence.*

Of course, one could also suppose Independence on either subset of preferences, for example, only $\succsim_\Lambda$ (the definition for preferences over late resolving lotteries is analogously defined).

Independence over Early Resolving Lotteries (I_Λ): *The relation $\succsim_\Lambda$ satisfies Independence.*

The last assumption is **Time Neutrality (TN)**, discussed previously.

Segal (1990), among other things, relates his proposed axioms to one another. In particular, he shows that if $\succsim$ satisfies WO and C, then (*i*) ROCL implies TN; (*ii*) ROCL and R imply I, and ROCL and I imply R; and (*iii*) R, I, and TN, imply ROCL. We can extend Segal's reasoning to include CTI, PTI, SPTI, and TI.[27]

Proposition 14. *Suppose $\succsim$ satisfies WO and C. The following statements are true.*[28]

[26] Both Γ and Λ are isomorphic to Δ, and therefore $\succsim_\Gamma$ and $\succsim_\Lambda$ can be interpreted as the the individual's preferences over simple lotteries in the appropriate period.

[27] We do not discuss the implications of SPTI and PTI individually, nor CTI individually, although our results can be extended. We do so because of the focus of our analysis is on CTI togerther with at least one of SPTI and PTI holding.

[28] Some items have already been established earlier; we add them here for completeness.

1. *(i) ROCL implies SPTI (ii) TI implies SPTI, R, and CTI; (iii) SPTI implies PTI.*

2. *(i) R, CTI, and PTI jointly imply TI (and so SPTI); (ii) CTI, SPTI and TN jointly imply ROCL*

3. *TN, R, CTI, and PTI jointly imply I (and so ROCL).*

Proof of Proposition 14. We show each part in turn

1. Observe that ROCL implies that all lotteries with the same reduced form probabilities are indifferent, which immediately implies SPTI. It is clear that TI implies SPTI, R and CTI since they are just TI applied to particular subsets of mixtures. We discussed previously that SPTI implies PTI. $\square$

2. Axioms R, PTI and CTI have been already shown to imply a posterior-anticipatory representation, which implies TI (and so SPTI). CTI and SPTI implies that we have a representation of the form $\hat{\nu}_1(\phi(P)) + \sum_i P(p_i)\nu_2(p_i)$. Over early resolving lotteries this takes the structure $\hat{\nu}_1(\phi(P)) + \sum_i P(\delta_{x_i})\nu_2(\delta_{x_i})$, which is simply a non-expected utility functional over the reduced form probabilities. TN implies this must be true also for any lottery with structure $\hat{\nu}_1(\phi(P)) + \nu_2(\phi(P))$, and so $\nu_2(\phi(P)) = \sum_i P(\delta_{x_i})\nu_2(\delta_{x_i})$, and so ν_2 satisfies reduction, and so ROCL is satisfied. $\square$

3. From item (2), R, CTI, and PTI imply SPTI, and we know that CTI, SPTI, and TN imply ROCL. $\square$

This concludes the proof of Proposition 14. $\blacksquare$

All relationships in Proposition 14 are interpreted via the lens of restrictions on preferences. In the context of our paper, it is perhaps more instructive to interpret them via the functional forms.

Corollary 1. *Suppose $\succsim$ has a prior-anticipatory representation. Then (i) TN or ROCL implies that ν_2 is an expected utility functional; and (ii) R or I_Λ implies that $\succsim$ has a posterior-anticipatory representation.*

If $\succsim$ has a posterior-anticipatory representation and satisfies TN, then it is expected utility.

Proof of Corollary 1: A prior-anticipatory representation implies that SPTI is satisfied. From the previous proof we know that SPTI and TN jointly imply ROCL, and that ROCL alone implies TN. Given the representation $\hat{\nu}_1(\phi(P)) + \sum_i P(p_i)\nu_2(p_i)$, TN implies that $\nu_2(p) = \sum_i p(\delta_{x_i})\nu_2(\delta_{x_i})$, and so ν_2 is expected utility. As shown previously, if R is satisfied then a posterior-anticipatory representation is implied. Moreover, Proposition 15 below shows that if I_Λ is satisfied, then a posterior-anticipatory representation is implied.

The representation of posterior-anticipatory preferences has the form $\sum_i P(p_i)\sum_j p_i(x_j)\hat{u}(x_j)+\sum_i P(p_i)\nu_1(p_i)$. Observe that over Λ these preferences have the structure $\sum_i P(p_i)\sum_j p_i(x_j)\hat{u}(x_j)+\sum_i P(\delta_{x_i})\nu_1(\delta_{x_i})$, which is expected utility. Thus, if TN is satisfied, preferences over Γ must also satisfy Independence. Given R, I and TN imply standard expected utility. □

These results allude to an alternative characterization of posterior-anticipatory preferences.

Proposition 15. *The following are equivalent:*

- *The relation $\succsim$ satisfies WO, C, CTI, SPTI, and I_Λ*
- *The relation $\succsim$ has a posterior-anticipatory representation*

Proof of Proposition 15: We use the fact that the relation $\succsim$ has a prior-separable expected utility representation if and only if it has a posterior-anticipatory representation. Given this, the result follows from the following claim. ■

Claim 9. *The relation $\succsim$ has a prior-separable expected utility representation if and only if it satisfies WO, C, CTI, SPTI, and I_Λ.*

Proof of Claim 9. It is easy to check that any V_{pseu} representation satisfies CTI, SPTI and I_Λ. For the other direction, notice that we have (given CTI and SPTI) a representation of the form $\hat{\nu}_1(\phi(P))+\sum_i P(p_i)\nu_2(p_i)$. Moreover, I_Λ implies that Independence is satisfied over lotteries in Λ. Observe that within Λ the representation has the form $\hat{\nu}_1(\phi(P))+\sum_i P(\delta_{x_i})\nu_2(\delta_{x_i})$. The second terms is simply an expected utility functional on Λ. Thus, the first term must be expected utility over the reduced form probabilities in order for Independence to be satisfied. □

A.4 Extension to T-stage lotteries

In this section we demonstrate how our characterization can be extended to a domain with arbitrary number of stages, T. For brevity, we will only discuss the extension of our most general ABB representation, stating the modified functional form and the underlying assumptions. Similar approaches extend our other two representation results.

Consider T-stage compound lotteries, and denote by C^t the set of lotteries that are degenerate for the first $(T-t)$ stages. Note that $t>s$ implies $C^t \subsetneq C^s$, and that our domain coincides with C^T. There is a natural isomorphism between elements of C^t and t-stage lotteries. To avoid confusion, we denote a typical element of C^t by P^t, and, for any such P^t, we denote the corresponding t-stage sublottery (that is, if it is in the support of some $(t+1)$-stage lottery) by $\tilde{P}^t$.[29] For any

[29] Note that we associate each t-stage lottery with lotteries that *start* with degenerates nodes, rather than with those that have (at least) the same amount of degenerate nodes across time, but in different order.

two lotteries $P,Q \in C^t$ and $\alpha \in [0,1]$ we denote by $[\alpha,t,P,Q]$ their α-mix at time t, that is, $[\alpha,t,P,Q](P^{t-1}) = \alpha P(P^{t-1}) + (1-\alpha)Q(P^{t-1})$. Observe that $[\alpha,t,P,Q] \in C^t$. For any t, we denote by $\phi(P^t)$ the reduced form probabilities over final outcomes induced by P^t.

The first two axioms extend our main axioms to the new domain.

A1. $t-PTI$: For any $\alpha \in [0,1]$, any $t \leq T$, and any $P,P^{'},Q \in C^t$ such that $\phi(P) = \phi(P^{'}) = \phi(Q)$; $P \succsim P^{'}$ if and only if $[\alpha,t,P,Q] \succsim [\alpha,t,P^{'},Q]$.

A2. $t-CTI$: For any $\alpha \in [0,1]$ and any $t \leq T$, if $\phi(P^t) = \phi({P^{'}}^t)$, $\phi(Q^t) = \phi({Q^{'}}^t)$, $P^t \succsim Q^t$, and ${P^{'}}^t \succsim {Q^{'}}^t$; then $[\alpha,t,P^t,{P^{'}}^t] \succsim [\alpha,t,Q^t,{Q^{'}}^t]$.

Before stating the conceptually new axiom, we introduce some notation. For any $P^{t-1},Q^{t-1} \in C^{t-1}$ and any $P^t \in C^t$, denote by $P^t_{\tilde{P}^{t-1}\to\tilde{Q}^{t-1}} \in C^t$ the lottery in which we replace the *sublottery* $\tilde{P}^{t-1}$ in the support of P^t with $\tilde{Q}^{t-1}$.

A3. Prior-Dependent Substitution: If $P^{t-1},Q^{t-1} \in C^{t-1}$, $\phi(P^{t-1}) = \phi(Q^{t-1})$, and $P^{t-1} \sim Q^{t-1}$, then $P^t \sim P^t_{\tilde{P}^{t-1}\to\tilde{Q}^{t-1}}$.

To define a T-period ABB representation we need just a bit further notation, identifying a sublottery with a specific node at the probability tree. Suppose in each time t there are J_t nodes according to P^T. Let $\tilde{P}^t_{j_t}$ be the sublottery that starts at node j_t in time t. Let $\rho(\tilde{P}^t_{j_t})$ be the probability of reaching $\tilde{P}^t_{j_t}$ according to P^T (that is, multiplying probabilities along the unique path leading to it, starting at the root of the tree).

Definition 9. *A T-period ABB representation is given by:*

$$V(P^T) = \sum_x u(x)\phi(P^T)(x) + \sum_{t=0}^{T-1}\sum_{j_t=1}^{J_t}\sum_{i_{t+1}=1}^{J_{t+1}} \rho(\tilde{P}^t_{j_t})\tilde{P}^t_{j_t}(\tilde{P}^{t+1}_{i_{t+1}})\nu_{t+1}(\phi(\tilde{P}^t_{j_t}),\phi(\tilde{P}^{t+1}_{i_{t+1}}))$$

Proposition 16. *The relation $\succsim$ satisfies A1-A3 if and only if it admits a T-period ABB representation.*

To prove this result, we will first show how to pin down ν_1, treating the sublotteries of length $T-1$ in the support of P^T as the final prizes. We then proceed forward, repeating the same arguments along the probability tree and sequentially trace all other ν_t, for $t \leq T$.

Proof of Proposition 16:

1. We start in Period 1 in order to first pin down ν_1.

   In Period 0, we apply $0-PTI$ and $0-CTI$ and using the same arguments as in the $T=2$ case (treating the sublotteries in the support of P^T as the final prizes) to obtain the representation

$$V(P^T) = \sum_{\tilde{P}^{T-1}} P^T(\tilde{P}^{T-1}) V_1(\phi(P^T), \tilde{P}^{T-1})$$

which, with some algebraic manipulations, can always be rewritten as

$$\sum_x u(x)\phi(P^T)(x) + \sum_{\tilde{P}^{T-1}} P^T(\tilde{P}^{T-1})\nu_1(\phi(P^T), \tilde{P}^{T-1}) + \sum_{\tilde{P}^{T-1}} P^T(\tilde{P}^{T-1}) V_2(\tilde{P}^{T-1}) \qquad (1)$$

We next argue that, without loss of generality, we can always normalize $\nu_1(\phi(P^{T-1}), \tilde{P}^{T-1}) = 0$. To see this, define $\hat{\nu_1}(\phi(P^T), \tilde{P}^{T-1}) = v_1(\phi(P^T), \tilde{P}^{T-1}) - v_1(\phi(P^{T-1}), \tilde{P}^{T-1})$; and $\hat{V}_2(\tilde{P}^{T-1}) = V_2(\tilde{P}^{T-1}) + v_1(\phi(P^{T-1}), \tilde{P}^{T-1})$. Observe that (1) is equivalent to

$$\sum_x u(x)\phi(P^T)(x) + \sum_{\tilde{P}^{T-1}} P^T(\tilde{P}^{T-1})\hat{\nu_1}(\phi(P^T), \tilde{P}^{T-1}) + \sum_{\tilde{P}^{T-1}} P^T(\tilde{P}^{T-1}) \hat{V}_2(\tilde{P}^{T-1})$$

but now we have $\hat{\nu_1}(\phi(P^{T-1}), \tilde{P}^{T-1}) = 0$. Therefore, we will now simply suppose that the representation in (1) with $\nu_1(\phi(P^{T-1}), \tilde{P}^{T-1}) = 0$. This normalization, in turns, implies that $P^{T-1} \succsim Q^{T-1}$ if and only if

$$\sum_x u(x)\phi(P^{T-1})(x) + V_2(\tilde{P}^{T-1}) \geq \sum_x u(x)\phi(Q^{T-1})(x) + V_2(\tilde{Q}^{T-1})$$

Moreover, if $\phi(P^{T-1}) = \phi(Q^{T-1})$, then $P^{T-1} \succsim Q^{T-1}$ if and only if $V_2(\tilde{P}^{T-1}) \geq V_2(\tilde{Q}^{T-1})$.

Now we apply the Prior-Dependent Substitution axiom. Take any $P^{T-1}, Q^{T-1} \in C^{T-1}$ such that $\phi(P^{T-1}) = \phi(Q^{T-1})$ and $P^{T-1} \sim Q^{T-1}$. The axiom requires that $V(P^T) = V(P^T_{\tilde{P}^{T-1} \to \tilde{Q}^{T-1}})$. Using what we established thus far, and applying the representation in (1), this implies that

$$V(P^T) - V(P^T_{\tilde{P}^{T-1} \to \tilde{Q}^{T-1}}) = \nu_1(\phi(P^T), \tilde{P}^{T-1}) - \nu_1(\phi(P^T), \tilde{Q}^{T-1}) = 0$$

which, in turns, implies that $\nu_1(\phi(P^T), \cdot)$ can only depend on the reduced form probabilities of its second argument.

2. Having pinned down ν_1, we move to Step 2. Here we focus on lotteries in C^{T-1}, tracing the form of V_2. Because we have pinned down ν_1, by A_3 it is without loss to use this subdomain to pin down V_2. We do this in a two-step manner.

   Recall that from the previous step, we obtain a representation

$$\sum_x u(x)\phi(P^T)(x) + \sum_{\tilde{P}^{T-1}} P^T(\tilde{P}^{T-1})\nu_1(\phi(P^T), \phi(\tilde{P}^{T-1})) + \sum_{\tilde{P}^{T-1}} P^T(\tilde{P}^{T-1})V_2(\tilde{P}^{T-1})$$

Define

$$\bar{V_2}(P^{T-1}) = \sum_x u(x)\phi(P^{T-1})(x) + V_2(\tilde{P}^{T-1})$$

and note that

$$\bar{V_2}(P^{T-1}) = V_1(P^{T-1})$$

Applying A1-A3 for $t = T-1$, we (analogously to Step 1) obtain the representation

$$\bar{V_2}(P^{T-1}) = \sum_x u(x)\phi(P^{T-1})(x) + \sum_{\tilde{P}^{T-2}} \tilde{P}^{T-1}(\tilde{P}^{T-2})\nu_2(\phi(\tilde{P}^{T-1}), \phi(\tilde{P}^{T-2})) + \sum_{\tilde{P}^{T-2}} \tilde{P}^{T-1}(\tilde{P}^{T-2})V_3(\tilde{P}^{T-2})$$

Thus

$$\begin{aligned} & \sum_x u(x)\phi(P^{T-1})(x) + \sum_{\tilde{P}^{T-2}} \tilde{P}^{T-1}(\tilde{P}^{T-2})\nu_2(\phi(\tilde{P}^{T-1}), \phi(\tilde{P}^{T-2})) + \sum_{\tilde{P}^{T-2}} \tilde{P}^{T-1}(\tilde{P}^{T-2})V_3(\tilde{P}^{T-2}) \\ = \; & \sum_x u(x)\phi(P^{T-1})(x) + V_2(\tilde{P}^{T-1}) \end{aligned}$$

which implies that

$$V_2(\tilde{P}^{T-1}) = \sum_{\tilde{P}^{T-2}} \tilde{P}^{T-1}(\tilde{P}^{T-2})\nu_2(\phi(\tilde{P}^{T-1}), \phi(\tilde{P}^{T-2})) + \sum_{\tilde{P}^{T-2}} \tilde{P}^{T-1}(\tilde{P}^{T-2})V_3(\tilde{P}^{T-2})$$

yielding a representation of V_2 as a function of ν_2 and the continuation value V_3.

3. We then continue iterating forward, using the exact same steps as in the previous stage, sequentially pinning down V_t as a function of ν_t and continuation value V_{t+1} for $t =, 3, .., T-1$. At time $T-1$ there is no continuation value, but in this stage we simply set $\nu_T = V_T$; note that V_T automatically depends only on the prior beliefs (and its second argument is always degenerate). ■

References

Ahlbrecht, Martin and Martin Weber, "Preference for gradual resolution of uncertainty," *Theory and Decision*, 1997, *43* (2), 167–185.

Arai, Dariush, "Temporal resolution of uncertainty in risky choices," *Acta psychologica*, 1997, *96* (1), 15–26.

Bénabou, Roland, "The economics of motivated beliefs," *Revue d'économie politique*, 2015, *125* (5), 665–685.

_ **and Jean Tirole**, "Mindful economics: The production, consumption, and value of beliefs," *Journal of Economic Perspectives*, 2016, *30* (3), 141–64.

Berns, Gregory S, Jonathan Chappelow, Milos Cekic, Caroline F Zink, Giuseppe Pagnoni, and Megan E Martin-Skurski, "Neurobiological substrates of dread," *Science*, 2006, *312* (5774), 754–758.

Brown, Alexander L and Hwagyun Kim, "Do individuals have preferences used in macro-finance models? An experimental investigation," *Management Science*, 2013, *60* (4), 939–958.

Brunnermeier, Marcus and Jonathan Parker, "Optimal Expectations," *American Economic Review*, 2005, *95*, 4.

Caplin, Andrew and John Leahy, "Psychological expected utility theory and anticipatory feelings," *The Quarterly Journal of Economics*, 2001, *116* (1), 55–79.

_ **and** _ , "The supply of information by a concerned expert," *The Economic Journal*, 2004, *114* (497), 487–505.

_ **and Kfir Eliaz**, "Aids policy and psychology: A mechanism-design approach," *RAND Journal of Economics*, 2003, pp. 631–646.

Cerreia-Vioglio, Simone, David Dillenberger, and Pietro Ortoleva, "Cautious Expected Utility and the Certainty Effect," *Econometrica*, March 2015, *83* (2), 693–728.

Chew, Soo Hong, "A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox," *Econometrica*, 1983, pp. 1065–1092.

_ **and Joanna L Ho**, "Hope: An empirical study of attitude toward the timing of uncertainty resolution," *Journal of Risk and Uncertainty*, 1994, *8* (3), 267–288.

Debreu, Gerard, "Topological methods in cardinal utility theory," *in Mathematical methods in the social sciences, Stanford University Press*, 1960, pp. 16–26.

Dekel, Eddie, "An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom," *Journal of Economic Theory*, 1986, *40* (2), 304–318.

Dillenberger, David, "Preferences for One-Shot Resolution of Uncertainty and Allais-Type Behavior," *Econometrica*, 2010, *78* (6), 1973–2004.

_ **and Uzi Segal**, "Skewed noise," *Journal of Economic Theory*, 2017, *169*, 344 – 364.

Duraj, Jetlir and Kevin He, "Dynamic Information Design with Diminishing Sensitivity Over News," *arXiv preprint arXiv:1908.00084*, 2019.

Eliaz, Kfir and Andrew Schotter, "Paying for confidence: An experimental study of the demand for non-instrumental information," *Games and Economic Behavior*, 2010, *70* (2), 304–324.

_ **and Ran Spiegler**, "Can anticipatory feelings explain anomalous choices of information sources?," *Games and Economic Behavior*, 2006, *56* (1), 87–104.

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica, "Suspense and surprise," *Journal of Political Economy*, 2015, *123* (1), 215–260.

Epstein, Larry G and Stanley E Zin, "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework," *Econometrica*, 1989, pp. 937–969.

Falk, Armin and Florian Zimmermann, "Beliefs and Utility: Experimental Evidence on Preferences for Information," CESifo Working paper, No. 6061 2016.

Ganguly, Ananda R and Joshua Tasoff, "Fantasy and dread: the demand for information and the consumption utility of the future," *Management Science*, 2017, *63* (12), 3999–4446.

Grant, Simon, Atsushi Kajii, and Ben Polak, "Intrinsic preference for information," *Journal of Economic Theory*, 1998, *83* (2), 233–259.

_ **,** _ **, and** _ , "Temporal Resolution of Uncertainty and Recursive Non-expected Utility Models," *Econometrica*, 2000, *68* (2), 425–434.

Gul, Faruk, Paulo Natenzon, and Wolfgang Pesendorfer, "Random Evolving Lotteries and Intrinsic Preference for Information," Technical Report 2016.

Kőszegi, Botond and Matthew Rabin, "Reference-dependent consumption plans," *The American Economic Review*, 2009, *99* (3), 909–936.

Kocher, Martin G, Michal Krawczyk, and Frans van Winden, "Let me dream on! Anticipatory emotions and preference for timing in lotteries," *Journal of Economic Behavior and Organization*, 2014, *98*, 29–40.

Kőszegi, Botond, "Health anxiety and patient behavior," *Journal of health economics*, 2003, *22* (6), 1073–1084.

_ , "Emotional agency," *The Quarterly Journal of Economics*, 2006, *121* (1), 121–155.

_ , "Utility from anticipation and personal equilibrium," *Economic Theory*, 2010, *44* (3), 415–444.

Kreps, David M and Evan L Porteus, "Temporal resolution of uncertainty and dynamic choice theory," *Econometrica*, 1978, pp. 185–200.

Lazarus, Richard S, "Psychological stress and the coping process," *New York, McGraw-Hill*, 1966.

Loewenstein, George, "Anticipation and the valuation of delayed consumption," *The Economic Journal*, 1987, *97* (387), 666–684.

Lovallo, Dan and Daniel Kahneman, "Living with uncertainty: attractiveness and resolution timing," *Journal of Behavioral Decision Making*, 2000, *13* (2), 179.

Masatlioglu, Yusufcan, A Yesim Orhun, and Collin Raymond, "Preferences for Non-Instrumental Information and Skewness," Technical Report, Mimeo 2017.

Mullainathan, Sendhil and Andrei Shleifer, "The market for news," *American Economic Review*, 2005, pp. 1031–1053.

Nielsen, Kirby, "Preferences for the resolution of uncertainty and the timing of information," Working paper 2017.

Oster, Emily, Ira Shoulson, and E. Ray Dorsey, "Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease," *American Economic Review*, 2013, *103* (2), 804–30.

Pagel, Michaela, "A news-utility theory for inattention and delegation in portfolio choice," *Econometrica*, 2018, *86* (2), 491–522.

Quiggin, John, "A theory of anticipated utility," *Journal of Economic Behavior & Organization*, 1982, *3* (4), 323–343.

Sarver, Todd, "Dynamic Mixture-Averse Preferences," *Econometrica*, 2018, *86* (4), 1347–1382.

Schweizer, Nikolaus and Nora Szech, "Optimal revelation of life-changing information," *Management Science*, 2018, *64* (11), 4967–5460.

Segal, Uzi, "Two-stage lotteries without the reduction axiom," *Econometrica*, 1990, pp. 349–377.

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

Sims, Christopher A., "Implications of rational inattention," *Journal of Monetary Economics*, 2003, *50* (3), 665–690.

von Gaudecker, Hans-Martin, Arthur van Soest, and Erik Wengström, "Heterogeneity in risky choice behavior in a broad population," *The American Economic Review*, 2011, pp. 664–694.

Wakker, Peter, "Additive representations on rank-ordered sets: II. The topological approach," *Journal of Mathematical Economics*, 1993, *22* (1), 1–26.

Zimmermann, Florian, "Clumped or piecewise? Evidence on preferences for information," *Management Science*, 2014, *61* (4), 740–753.