Payoff-Belief-Separable Preferences*

David Dillenberger[†] Collin Raymond[‡]

September 2017

Preliminary and incomplete

Abstract

We introduce a new class of preferences — which we call payoff-belief-separable (PBS) utility — that captures a general additively separable approach to belief-based utility, and that encompasses many popular models in the behavioral literature. We axiomatize a general class of PBS preferences, as well as two prominent special cases that allow utility to depend on the level of each period's beliefs but not on changes in beliefs across periods. We also identify the intersection of PBS preferences with the class of recursive preferences and characterize attitudes towards the timing of resolution of uncertainty for such preferences.

JEL codes: D80, D81

Key words: Anticipatory Utility, Compound lotteries, Preferences over beliefs, Recursive Preferences.

1 Motivation

Models of decision-making in which individuals derive utility not only from material outcomes but also from their current and future beliefs have become increasingly prominent in economics (see, for example, Caplin and Leahy, 2001; Kőszegi and Rabin, 2009; and Ely, Frankel, and Kamenica, 2015). In most of these models, overall utility is additively separable between material payoffs and beliefs. Our aim in this paper is to introduce and analyze a new class of utility functions — which we call *payoff-belief-separable (PBS)* utility — that captures a general additively separable approach to belief-based utility.

Models within this class usually take one of two forms. The first, as in Caplin and Leahy (2001), allows individuals' 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 these set of models as

*Acknowledgments: to be added
[†]Department of Economics, University of Pennsylvania (ddill@sas.upenn.edu)
[‡]Department of Economics, Purdue University (collinbraymond@gmail.com)

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 in any given period. We refer to this class as *changing beliefs* models. Our PBS representation encompasses both these frameworks. Moreover, we 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.

While individuals in these models gain utility from their beliefs, they cannot directly choose them. Rather, individuals begin with a prior belief, receive information, and form interim beliefs by applying Bayes' rule. Therefore, individuals can control their beliefs only by choosing particular information structures. And since individuals gain utility from their beliefs, they may exhibit 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 is the one that distinguishes these models from the standard model, in which individuals would be indifferent between all possible information structures when no actions are available. To tightly link these models to observable behavior, we look at preferences over the combination of information structures and prior beliefs. These can be elicited in a natural way in experimental and field settings. Formally, taking advantage of the natural 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.

In order to provide some intuition regarding the class of PBS functions, let P be a typical two-stage compound lottery. In period 0 it induces a prior belief $\phi(P)$; 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 a posterior belief over outcomes; the individual now knows that in period 2 they 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 centering on this outcome (denoted δ_{x_j}). The total utility of this scenario, denoted $V_{PBS}(P)$, is given by:

$$
\begin{aligned}
V_{PBS}(P) &= \underbrace{\sum_i P(p_i) \sum_j p_i(x_j) u(x_j)}_{\text{expected utility from material payoffs}} \\
&+ \underbrace{\sum_i P(p_i) \nu_1(\phi(P), p_i)}_{\text{expected utility from beliefs in periods 0 and 1}} \\
&+ \underbrace{\sum_i P(p_i) \sum_j p_i(x_j) \nu_2(p_i, \delta_{x_j})}_{\text{expected utility from beliefs in periods 1 and 2}}
\end{aligned}
$$

The first term represents the expected consumption utility of the two-stage lottery — the expected utility that the individual receives from material outcomes (in period 2). The second term represents period 1's belief-based utility — the individual's 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.

We first show how the prominent sub-classes of PBS 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 a PBS functional. We show that PBS utility is characterized by a single relaxation of the standard vNM-Independence axiom (applied to preferences over compound lotteries): The Conditional Two-Stage Independence (CTI) axiom requires Independence (in mixing the 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.

We next turn to showing how strengthening CTI 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, the Strong Conditional Two-Stage Independence (SCTI) axiom drops from the previous axiom the requirement that $\phi(R)$ agrees with $\phi(P) = \phi(Q)$. We last show that adding a version of one-stage Independence — that is, Independence imposed on a specific subset of degenerate two-stage lotteries — to the previous axioms characterizes posterior-anticipatory beliefs. Thus, our results demonstrate how a simple set of familiar and easily tested conditions in terms of observed behavior allows distinguishing between different types of belief-dependent utility.

PBS preferences are not the only class of preference 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, primarily developed by Kreps and Porteus (1978) and extended by Segal (1990), focuses on recursive preferences over compound lotteries (and information). 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.

Lastly, we analyze what types of restrictions on the functional forms of PBS preferences are equivalent to well-known types of intrinsic (i.e., non-instrumental) 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, for example those of Kőszegi and Rabin (2009), which were made in the context of specific functional forms. Our results will allow us to compare how different classes of models (recursive and PBS) can, or cannot, accommodate different non-instrumental attitudes towards information.

2 Model, Characterization, and Special Cases

Preliminaries:

Although we motivate PBS preferences as capturing utility from beliefs as they evolve over time in response to information, we formally use as our domain the space of two-stage compound lotteries, as is typical in the literature. 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 lotteries $p, q \in \Delta$ and $\alpha \in (0,1)$, we 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 probability 1 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. We sometime write a lottery (over either Δ or Δ^2) explicitly as a list; for example, $P = (p_1, P(p_1); \ldots; p_n; P(p_n))$, or $((p_i, P(p_i))_{i=1}^n$, denotes the two-stage lottery that, for $i = 1, ..., n$, yields p_i with probability $P(p_i)$. Define a *reduction operator* $\phi : \Delta^2 \to \Delta$ that maps compound lotteries to reduced one-stage lotteries by $\phi(Q) = \sum_{p \in \Delta} Q(p)p$.

Two special subsets of compound lotteries are (i) $\Gamma = \{D_p | p \in \Delta\}$, the set of degenerate lotteries in Δ^2. Γ is the set of late resolving lotteries, where any $P \in \Gamma$ captures a situation in which the information structure reveals no information in period 1, so that the interim posterior and prior beliefs coincide; and (ii) $\Lambda = \{Q \in \Delta^2 | Q(p) > 0 \Rightarrow p \in \bar{X}\}$, the set of compound lotteries whose outcomes are degenerate in Δ. Λ is the set of early resolving lotteries, where any $P \in \Lambda$ captures a situation in which the information structure reveals all information in period 1, so that interim posteriors after observing any signal have one element in their support.

Our primitive is a binary relation $\succsim$ over Δ^2. We define the restriction of $\succsim$ to the subsets Γ and Λ as $\succsim_\Gamma$ and $\succsim_\Lambda$, respectively.[1]

Functional Forms:

We first formally define payoff-belief separable utility.

Definition 1. *A payoff-belief separable (PBS) representation is a tuple $(V_{PBS}, u, \nu_1, \nu_2)$ consisting of a continuous function $V_{PBS} : \Delta^2 \to \mathbb{R}$ that represents $\succsim$, and continuous $u : X \to \mathbb{R}$, $\nu_1 : \Delta \times \Delta \to \mathbb{R}$, and $\nu_2 : \Delta \times \bar{X} \to \mathbb{R}$, such that*

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

The general PBS functional form 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, many models in the literature assume that individuals do not care about the changes in their beliefs, but rather care about the levels of their beliefs. Individuals may care about their beliefs in any given period 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 and define it as follows.

Definition 2. *A prior-anticipatory representation is a PBS 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 and the functional form is given by:

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

[1] Note that 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.

Clearly, both anticipatory representations above are subsets of V_{PBS}. 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.*

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 Conditional Two-Stage Independence (CTI). CTI requires the Independence axiom to hold within the set of compound lotteries which share the same reduced form probabilities over outcomes (the same ϕ). Observe that the set $\{Q \in \Delta^2 | \phi(Q) = p\}$ is convex for any $p \in \Delta$. Thus, CTI says that Independence holds along "slices" of the compound lottery space, where all elements of the slice have the same reduced form probabilities.

Conditional Two-Stage Independence (CTI): *If $\phi(P) = \phi(P') = \phi(Q)$, then $P \succsim P'$ if and only if $\alpha P + (1-\alpha)Q \succsim \alpha P' + (1-\alpha)Q$.*

Axiom CTI can be further motivated through the lens of preferences for information. Recall that we identify preferences over compound lotteries with preferences over the combination of information structures and prior beliefs. CTI then requires that within a set of information structures that correspond to the same prior beliefs, the individual is an expected utility maximizer; "non-standard" behavior may arise only when we change the underlying prior beliefs.

Our first main result shows that CTI, along with the standard two axioms above, is all we need to characterize preferences that admit a PBS representation.

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

CTI is not so 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 CTI 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. We call the axiom which formalizes this intuition Strong Conditional Two-Stage Independence (SCTI).

Strong Conditional Two-Stage Independence (SCTI): *If $\phi(P) = \phi(P')$, then $P \succsim P'$ if and only if $\alpha P + (1-\alpha)Q \succsim \alpha P' + (1-\alpha)Q$.*

SCTI rules out complementarity between the prior distribution and the corresponding information systems. That is, irrespectively of the underlying prior beliefs, the individual consistently chooses among information systems based on the expected utility criterion; violations of expected utility may occur only when ranking compound lotteries that do not refine the same prior beliefs.

SCTI clearly implies CTI. It also gives us our second characterization result:

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

In order to characterize posterior-anticipatory representations, we further impose a different Independence-type condition that is logically independent from both CTI and SCTI. Axiom I_Λ imposes Independence over the set of lotteries that fully resolve in the first stage, Λ. (Since the set of lotteries which resolve fully in the first stage is isomorphic to the set of one-stage lotteries, Independence has a natural meaning on this sub-domain.)

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

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

Special Cases:

The functional forms we have derived above are quite general. In many cases, we may want to suppose 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 (typically normalized to 0) from beliefs that do not change. We describe these preferences as belief stationarity invariant (BSI).

Definition 4. *A PBS representation is belief-stationarity invariant (BSI) if $\nu_1(p_i, p_i) = \nu_1(q_i, q_i) = \nu_2(\delta_x, \delta_x) = \nu_2(\delta_y, \delta_y)$ for all x, y, p_i, q_i.*

A second type of assumption is that the utility derived from beliefs does not depend on the timing of when those beliefs occur. We call this belief time invariance (BTI).

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

In order to relate BSI and BTI to behavior, we discuss a certain restriction on preferences over compound lotteries. The next axiom is due to Segal (1990).

Time Neutrality (TN): *If $P \in \Gamma$, $Q \in \Lambda$, and $\phi(P) = \phi(Q)$, then $P \sim Q$.*

Time Neutrality supposes that a decision-maker is indifferent between a lottery that resolves early, or a lottery that resolves late, provided that they induce the same probability distribution over final outcomes. This implies that ordering of preferences is the same over fully early resolving lotteries and fully late resolving lotteries (mapping them to their reduces formed equivalents).

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

Proposition 4. *Suppose $\succsim$ has a PBS 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.*

3 PBS and Recursive Preferences

PBS 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 also 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; and Sarver, 2016).

Segal (1990) was the first to formally discuss recursive preferences on the domain of compound lotteries. 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, $W(p) = W(\delta_{\mathbb{CE}_W(p)})$.[2]

Definition 6. *Suppose preferences over Δ^2 can be represented by the functional V. We say that preferences have a recursive representation (V_1, V_2), where $V_i : \Delta \to \mathbb{R}$, if and only if for all $P = (p_1, P(p_1); \ldots; p_n; P(p_n))$, we have $V(P) = V_1(\mathbb{CE}_{V_2}(p_1), P(p_1); \ldots; \mathbb{CE}_{V_2}(p_n); P(p_n))$.*

Segal (1990) provided a behavioral equivalent for these functional forms using an 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 $\alpha D_p + (1-\alpha)Q \succsim \alpha D_q + (1-\alpha)Q$.*

[2] 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 are monotone with respect to first-order stochastic dominance.

Recursivity, like CTI, applies Independence to a particular subset of compound lotteries. In particular, the original pair of lotteries being compared must be degenerate in the first stage, that is, members of Γ. Like CTI, Independence is thus applied to a particular "slice" of the compound lottery space. However, it is an orthogonal slice to that considered by CTI (or SCTI). 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, PBS and recursive, are related. Although they apply Independence to different slices of the compound lottery space, it is unclear whether there exist preferences which have both representations. The next result answers this question in the affirmative, and moreover, shows that the intersection is exactly those preferences which admit a posterior-anticipatory representation.

Proposition 5. *The following are equivalent:*

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

Figure 1 depicts the relationships discussed in the last two sections.

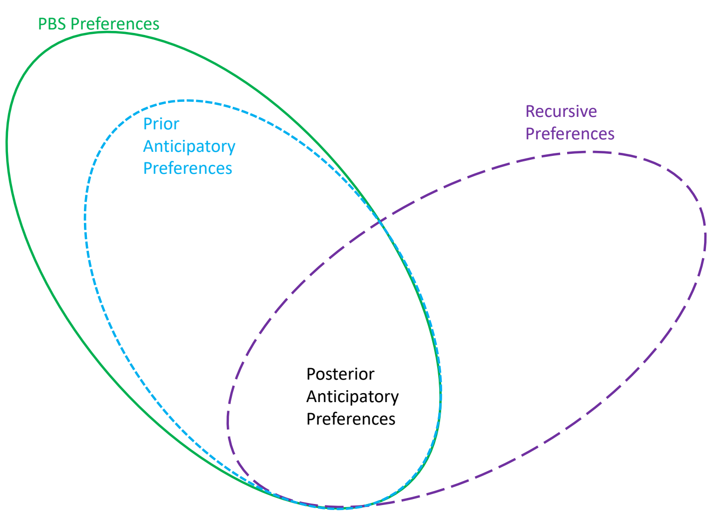

Figure 1: Relationships between Models

In general, PBS models can be directly tested in the most natural domain of information preferences given a fixed prior belief. In contrast, recursive preferences, *as a general class*, have no observable restrictions when the prior belief is fixed; testing the assumption of Recursivity (Axiom R) requires observing preferences as the prior changes. Proposition 5 then implies that this latter property applies as well to the subclass of preferences that have posterior-anticipatory representation (which is another manifestation of Axiom I_Λ). The proposition further implies that any posterior-anticipatory representation captures individuals who have "emotions over emotions". This is due to the fact that 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 PBS 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.

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 $P \sim Q$.*

Another assumption is to apply Independence not just to the set of full early resolving lotteries (Axiom I_Λ) but also to the set of fully late resolving lotteries. This restriction, which we denote by I_Γ, says that $\succsim_\Gamma$ satisfies Independence.

Independence (I): *Both I_Λ and I_Γ hold.*

A third assumption is **Time Neutrality (TN)**, discussed in the previous section.

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 and SCTI.

Proposition 6. *Suppose $\succsim$ satisfies WO and C. The following statements are true.*[3]

1. *(i) ROCL implies SCTI; and (ii) SCTI implies CTI.*
2. *(i) R and CTI jointly imply SCTI; (ii) I and SCTI jointly imply ROCL; and (iii) SCTI and TN jointly imply ROCL.*
3. *TN, R and CTI jointly imply I (and so ROCL).*

All relationships in Proposition 6 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.

[3] Items 1(ii) and 2(i) have already been established earlier; we add them here for completeness.

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

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

To give some intuitions, observe that a prior-anticipatory representation with TN implies that for any $Q \in \Lambda$ and $P \in \Gamma$ with the same reduced form probabilities: $\nu_1(\phi(Q)) + \sum_x Q(\delta_x)\nu_2(\delta_x) = \nu_1(\phi(P)) + \sum_i P(p_i)\nu_2(p_i)$. In other words, it must be that $\sum_x Q(\delta_x)\nu_2(\delta_x) = \sum_i P(p_i)\nu_2(p_i)$, which only occurs if ν_2 has an expected utility representation. Moreover, given a prior-anticipatory representation $\sum_i P(p_i)\sum_j p_i(x_j)u(x_j) + \nu_1(\phi(P)) + \sum_i P(p_i)\nu_2(p_i)$, I_Γ implies that the functional form $\sum_j \phi(P)(x_j)u(x_j) + \nu_1(\phi(P)) + \nu_2(\phi(P))$ must satisfy full Independence, which can only happen when both ν_1 and ν_2 have expected utility representations.

Of course, a posterior-anticipatory representation already imposes that ν_2 is expected utility. Thus, TN or I_Γ imposes that ν_1 also is expected utility, and so we obtain the standard model.

4 PBS and the Timing of Uncertainty Resolution

Individuals with PBS utility will have intrinsic preferences over information, that is, they may prefer one information structure to another even in the absence of the ability to condition actions on either of them. Many papers looking at specific examples of PBS functional forms 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 have been a major focus of the decision-theoretic literature. However, there do not exist equivalent characterizations for PBS preferences. Our results will allow us to compare how different classes of models (recursive and PBS) can accommodate (or not) different non-instrumental attitudes towards information.

Many authors, such as Kreps and Porteus (1978) and Grant, Kajii, and Polak (1998), conjecture that individuals not only prefer information to be fully resolved earlier (in period 1 rather than in period 2) but also that they always prefer Blackwell-more-informative signals in period 1, that is, earlier resolution of information. Drawing on Grant, Kajii, and Polak (1998), we can define a preference for early resolution of information.

Definition 7. $\succsim$ *displays a preference for early resolution of uncertainty (PERU) if for any* $Q, P \in \Delta^2$ *such that* $Q = (q_i, Q(q_i))_{i=1}^n$, $P = ((q_i, Q(q_i))_{i \neq j}; p_1, \beta Q(q_j); p_2, (1-\beta)Q(q_j))$ *for some* $\beta \in [0,1]$, *and* $q_j = \beta p_1 + (1-\beta)p_2$, *we have* $P \succsim Q$.

Preference for late resolution of information is analogously defined, by requiring that for the lotteries above $Q \succsim P$.

We now characterize anticipatory preferences that exhibit preferences for either earlier or later resolved lotteries.[4] 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.

Proposition 7. *The following statements are true:*

1. *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).*

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

Definition 8. *$\succsim$ exhibits a preference for one-shot resolution of uncertainty (PORU) if: (i) $\succsim$ satisfies TN, and (ii) For all $P, Q, R \in \Delta^2$ such that $\phi(P) = \phi(Q) = \phi(R)$, if $P \in \Lambda$ and $Q \in \Gamma$, then $P \sim Q \succsim R$.*

In contrast to Proposition 7, if preferences depend only on the level of beliefs, then they cannot generate strict PORU. In other words, there are no triple as in Definition 8 for which $P \sim Q \succ R$ hold.

Proposition 8. *If $\succsim$ has a prior-anticipatory representation, then it can never exhibit strict PORU.*

The intuition behind this proposition derives from Corollary 1, where we show that if $\succsim$ has a prior-anticipatory representation, then TN implies that ν_2 is expected utility. Recall that expected utility functionals do not generate any anomalous preferences towards information. The rest of the utility functional depends only on ϕ, the reduced form probability of the lottery. Moreover, since we identify anomalous attitudes towards information by using lotteries which all share the same reduced form probabilities, those other terms always take on the same value, regardless of the information structure. Thus, we can never observe strict PORU. This is in contrast to recursive preferences, where Dillenberger (2010) shows that strict PORU can occur and characterizes it in terms of a single condition on $\succsim_\Gamma = \succsim_\Lambda$.

[4] The most general form of PBS utility is too general to provide clean cut conditions.

Changing Beliefs	Prior*-Anticipatory Utility	Posterior-Anticipatory Utility
Caplin and Eliaz (2003)		Caplin and Leahy (2001)
Mullainathan and Shleifer (2005)		Kőszegi (2003)
Kőszegi and Rabin (2009)		Caplin and Leahy (2004)
Kőszegi (2010)		Eliaz and Schotter (2010)
Ely, Frankel and Kamenica (2015)		Szech and Schweitzer (2016)
Pagel (2016)		

Table 1: Some Models Nested by PBS

5 Discussion

We now discuss how our general framework applies to some specific functional forms used in the literature. In order to provide a sense of the breadth of models that our approach encompasses, Table 1 provides a list of papers which use functional forms nested by PBS. Many of these models also allow for actions, and so their domain and representation may appear different than that presented in this paper.[5]

Some models in the literature fit into the framework of changing beliefs models — they have a PBS representation, but do not have any anticipatory representation. These include some models that are explicitly meant to captures utility derived from changing beliefs, such as Kőszegi and Rabin (2009) and Pagel (2015). 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 PBS 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 PBS models not only in it's lack of continuity, but also in that a non-linear transformation is applied to the expected utility from changes in beliefs. 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 want to see signals that confirm their priors, as well as Caplin and Eliaz (2003), which takes on the form of a prior-dependent Kreps-Porteus representation. Within our domain, we also capture the model Kőszegi (2010), who explicitly 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 framework. In particular, all the models that adhere to the framework of Caplin and Leahy (2001) are capture by our model. More

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

precisely, our posterior-anticipatory utility delivers the Caplin-Leahy representation when applied to our domain (they originally allowed for intermediate actions). This includes models such as Kőszegi (2003), Caplin and Leahy (2004), Eliaz and Schotter (2010), and Szech and Schweitzer (2016).

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. As our characterization result shows, this amounts to an implicit assumption in the literature that anticipatory utility regarding the levels of beliefs is always accompanied by assuming the Independence axiom over early resolving lotteries, but not over late resolving lotteries. Descriptively, we believe prior-anticipatory motives are important, as they can accommodate behavior that violates expected utility over early resolving lotteries, in accordance with frequently observed experimental results, such as the Allais paradox.

We conclude by discussing the relationship between our model and several related axiomatizations which are also distinct from the literature on recursive functions. As previously discussed, our functional form nests that of Caplin and Leahy (2001). Caplin and Leahy provide an axiomatization of their functional form, but take as their domain the set of "psychological lotteries", which include lotteries not just over material outcomes, but also over psychological states (i.e., beliefs). Thus, their domain includes objects (psychological states) which are explicitly not observable, and not directly choosable. 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.

Yariv (2002) also provides a characterization of belief-dependent utility, although in a very different domain than ours, and for very different purposes. Her work relies on preferences being expressed over objects where beliefs are independently manipulable across periods. In contrast, our beliefs in period 2 are not independent of those in period 1.

Recently Gul, Natenzon and Pesendorfer (2016) have introduced a model that shares some key features of our model. Although the domain and objects of choice are quite different than ours, there are many key similarities in that 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, Gul, Nautenzon and Pesendorfer (2016) allow for non-additive measures.

6 Appendix

Proof of Lemma 1: By Definition 2, a prior-anticipatory representation is given by $\sum_i P(p_i)$ $\sum_j p_i(x_j)u(x_j) + \nu_1(\phi(P)) + \sum_i P(p_i)\nu_2(p_i)$. Note that preferences admit such a representation if and only if they can be represented by the functional $\hat{\nu}_1(\phi(P)) + \sum_i P(p_i)\nu_2(p_i)$, for some arbitrary non-expected utility functional $\hat{\nu}_1$. Similarly, by Definition 3, a posterior-anticipatory representation is given by $\sum_i P(p_i) \sum_j p_i(x_j)[u(x_j) + \nu_2(x_j)] + \sum_i P(p_i)\nu_1(p_i)$, which is equivalent to a representation of the form $\sum_i P(p_i) \sum_j p_i(x_j)\hat{u}(x_j) + \sum_i P(p_i)\nu_1(p_i)$. Clearly the second representation is a subset of the first.□

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

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

Proof of Claim 1. Consider the first two terms in the PBS 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 PBS 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 a PBS 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_{PBS} 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, 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 the three conditions in the statement of the claim.

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, and SCTI.*

Proof of Claim 4. It is easy to check that the axioms are necessary for the representation. For sufficiency, observe that SCTI implies CTI, 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 SCTI, 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: $\sum_i P(p_i)\widetilde{\nu}(\phi(P), p_i) = \hat{\nu}_1(\phi(P)) + \sum_i P(p_i)\nu_2(p_i)$.

This proves the equivalence in the proposition. $\square$

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 function 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, SCTI, and I_Λ.*

Proof of Claim 6. It is easy to check the necessity of the axioms to the representation. To show sufficiency, notice that we have (given SCTI) 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 this subset of lotteries 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.

This proves the equivalence in the proposition. □

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

To prove (1), we show that there is always a representation which is belief-stationary invariant 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 with positive support in p and sum up below only amongst those elements with positive support. 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:

$$\sum_x u(x)\rho(x) + \sum_p P(p)\hat{\nu}_1(\rho, p) + \sum_p \sum_x P(p)p(x)\hat{\nu}_2(p, \delta_x)$$

or

$$\sum_x \rho(x)u(x) + \sum_p P(p)[\nu_1(\rho, p) - \nu_1(p, p)] + \sum_p \sum_x P(p)p(x)[\nu_2(p, \delta_x) + \frac{\nu_1(p, p)}{N(p)p(x)}]$$

or

$$\sum_x \rho(x)u(x) + \sum_p P(p)\nu_1(\rho, p) - \sum_p P(p)\nu_1(p, p) + \sum_p \sum_x P(p)p(x)\nu_2(p, \delta_x) + \sum_p P(p)\nu_1(p, p) \sum_x \frac{1}{N(p)}$$

or

$$\sum_x \rho(x)u(x) + \sum_p P(p)\nu_1(\rho, p) - \sum_p P(p)\nu_1(p, p) + \sum_p P(p)\nu_1(p, p) + \sum_p \sum_x P(p)p(x)\nu_2(p, \delta_x)$$

which is the original utility function.□

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:

$$\sum_x \widetilde{u}(x)\rho(x) + \sum_p P(p)\hat{\nu}_1(\rho, p) + \sum_p \sum_x P(p)p(x)\widetilde{\nu}_2(p, \delta_x)$$

or

$$\sum_x \rho(x)[u(x) + \hat{\nu}_2(\delta_x, \delta_x)] + \sum_p P(p)\hat{\nu}_1(\rho, p) + \sum_p P(p)p(x)[\hat{\nu}_2(p, \delta_x) - \hat{\nu}_2(\delta_x, \delta_x)]$$

or

$$\sum_x \rho(x)u(x) + \sum_p P(p)\hat{\nu}_1(\rho, p) + \sum_p \sum_x P(p)p(x)\hat{\nu}_2(p, \delta_x) + \sum_x \rho(x)\hat{\nu}_2(\delta_x, \delta_x) - \sum_p \sum_x P(p)p(x)\hat{\nu}_2(\delta_x, \delta_x)$$

which is simply the original utility function.□

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

To prove part (2), we show that there is always a representation which is belief-time invariant. 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:

$$\sum_x \widetilde{u}(x)\rho(x) + \sum_p P(p)\hat{\nu}_1'(\rho, p) + \sum_p \sum_x P(p)p(x)\widetilde{\nu}_2'(p, \delta_x)$$

or

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

or

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

or

$$\sum_x \widetilde{u}(x)\rho(x) + \sum_p P(p)\hat{\nu}_1(\rho,p) + \sum_p \sum_x P(p)p(x)\widetilde{\nu}_2(p,\delta_x)$$

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

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

is no longer necessarily equals 0. □

Lastly we show part (3), that $\succsim$ has a representation which is both belief-stationary invariant and belief-time invariant if and only if it satisfies TN.

For the only if part, observe that for $P \in \Gamma$

$$\begin{aligned} V_{PBS}(P) &= E_{\phi(P)}(u) + \nu_1(\phi(P),\phi(P)) + \sum_j \phi(P)(x_j)\nu_2(\phi(P),\delta_{x_j}) \\ &= E_{\phi(P)}(u) + \sum_j \phi(P)(x_j)\nu_2(\phi(P),\delta_{x_j}) \end{aligned}$$

where the second equality is by BSI.

For $Q \in \Lambda$ we have

$$\begin{aligned} V_{PBS}(Q) &= E_{\phi(Q)}(u) + \sum_j \phi(Q)(x_j)\nu_1(\phi(Q),\delta_{x_j}) + \sum_j \phi(Q)(x_j)\nu_2(\delta_{x_j},\delta_{x_j}) \\ &= E_{\phi(Q)}(u) + \sum_j \phi(Q)(x_j)\nu_1(\phi(Q),\delta_{x_j}) \end{aligned}$$

where the second equality is again by BSI.

By BTI, $\sum_j \phi(P)(x_j)\nu_2(\phi(P),\delta_{x_j}) = \sum_j \phi(Q)(x_j)\nu_1(\phi(Q),\delta_{x_j})$. And if $\phi(P) = \phi(Q)$, then indeed $V_{PBS}(P) = V_{PBS}(Q)$, 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 simply 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. □

Proof of Proposition 5. We first show that $\succsim$ has a posterior-separable expected utility representation (i.e. it satisfies WO, C, SCTI, and I_Λ) if and only if it satisfies WO, C, CTI, and R.

Necessity is immediate. To show sufficiency, note that we have a representation of the form $V_{PC} = \sum_i P(p_i)\nu_{PC}(\phi(P), p_i)$. If recursivity 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: $\sum_i P(p_i)\sum_j p_i(x_j)\hat{u}(x_j) + \sum_i P(p_i)\nu_1(p_i)$.

Given a recursive representation and CTI, recall that Proposition 3 and the first part of the proof of this proposition imply that I_Λ is satisfied. This implies V_1 is expected utility. □

Proof of Proposition 6: Observe that ROCL implies that all lotteries with the same reduce form probabilities are indifferent, which immediately implies SCTI. It's clear that SCTI implies CTI.

R and CTI have been already shown to imply a posterior-anticipatory representation, which implies SCTI. We also know I_Λ and SCTI jointly imply R as well as posterior-anticipatory representation. Imposing I_Γ as well implies that we have a representation over fully late resolving lotteries of the form $\sum_x \phi(P)(x)\nu_{PSEU1}(x) + \nu_{PSEU2}(\phi(P))$, which must satisfy Independence. This implies $\nu_{PSEU2}(p_i)$ has an expected utility representation over final outcomes, which means that $\sum_x \phi(P)(x)\nu_{PSEU1}(x) + \sum_i P(p_i)\nu_{PSEU2}(p_i)$ is expected utility and so satisfies ROCL.

SCTI 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. Time neutrality implies this must true 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.

Last, observe that R and CTI imply that SCTI must be satisfied, and we know that SCTI and TN imply ROCL. □

Proof of Corollary 1: A prior-anticipatory representation implies that SCTI is satisfied. We know that SCTI plus TN imply reduction (from the previous proof), and that reduction alone implies TN. Given our 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. If I_Λ or R is satisfied then previous proofs imply that a posterior-anticipatory representation is implied. Moreover, I_Γ implies that Independence holds when preferences have given the structure: $\hat{\nu}_1(\phi(P)) + \nu_2(\phi(P))$, which implies both functionals must be expected utility with reduction.

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 these preferences, over fully early resolving lotteries, 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 late resolving lotteries must also satisfy Independence. So, given I_Γ, we know Independence is satisfied when the functional has the form $\sum_i P(p_i)\sum_j p_i(x_j)\hat{u}(x_j) + \nu_1(\phi(P))$, which implies ν_1 is expected utility; and so the entire functional is also expected utility which satisfies reduction. □

Proof of Proposition 7: For the first part, 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 $\beta\hat{\nu_2}(p_1) + (1-\beta)\hat{\nu_2}(p_2) \geq \hat{\nu_2}(\beta p_1 + (1-\beta)p_2)$. And since the triple p_1, p_2, and β were arbitrary, the inequality holds if and only if $\hat{\nu_2}$ is convex. Similarly, the inequality is reversed if and only if $\hat{\nu_2}$ is concave.

For the second part, simply replace in the entire paragraph above $\hat{\nu_2}$ with $\bar{\nu_1}$. □

References

[1] Andrew Caplin and Kfir Eliaz. Aids policy and psychology: A mechanism-design approach. *RAND Journal of Economics*, pages 631–646, 2003.

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

[3] Andrew Caplin and John Leahy. The supply of information by a concerned expert. *The Economic Journal*, 114(497):487–505, 2004.

[4] David Dillenberger. Preferences for one-shot resolution of uncertainty and Allais-type behavior. *Econometrica*, 78(6):1973–2004, 2010.

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

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

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

[8] Faruk Gul, Paulo Natenzon, and Wolfgang Pesendorfer. Random evolving lotteries and intrinsic preference for information. Technical report, 2016.

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

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

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

[12] Sendhil Mullainathan and Andrei Shleifer. The market for news. *American Economic Review*, pages 1031–1053, 2005.

[13] Michaela Pagel. A news-utility theory for inattention and delegation in portfolio choice. *Working Paper*, 2016.

[14] Todd Sarver. Optimal reference points and anticipation. Technical report, Discussion Paper, Center for Mathematical Studies in Economics and Management Science, 2012.

[15] Nikolaus Schweizer and Nora Szech. Optimal revelation of life-changing information. Working paper, 2016.

[16] Uzi Segal. Two-stage lotteries without the reduction axiom. *Econometrica*, pages 349–377, 1990.

[17] Leeat Yariv. Believe and let believe:axiomatic foundations for belief dependent utility functionals. Technical report, 2001.