Preferences for One-Shot Resolution of Uncertainty and Allais-Type Behavior

Dillenberger, David

9 April 2008

Preferences for One-Shot Resolution of Uncertainty and Allais-Type Behavior *

David Dillenberger[†]

April 9, 2008

Abstract

Experimental evidence suggests that individuals are more risk averse when they perceive risk gradually. We address these findings by studying a decision maker (DM) who has recursive preferences over compound lotteries and who cares about the way uncertainty is resolved over time. DM has preferences for one-shot resolution of uncertainty if he always prefers any compound lottery to be resolved in a single stage. We establish an equivalence between dynamic preferences for one-shot resolution of uncertainty and static preferences that are identified with the behavior observed in Allais-type experiments. The implications of this equivalence on preferences over information systems are examined. We define the gradual resolution premium and demonstrate its magnifying effect when combined with the usual risk premium. In an intertemporal context, preferences for one-shot resolution of uncertainty capture narrow framing.

1. Introduction

Experimental evidence suggests that individuals are more risk averse when they perceive risk that is gradually resolved over time. In an experiment with college students, Gneezy and Potters [1997] found that subjects invest less in risky assets if they evaluate financial outcomes more frequently. Haigh and List [2005] replicated the study of Gneezy and Potters with professional traders and found an even stronger effect. These two studies allow for flexibility in adjusting investment according to how often the subjects evaluate the returns. Bellemare, Krause, Kröger, and Zhang [2005] found that even when all subjects have the same investment flexibility, variations in the frequency of information feedback alone affects

*I am grateful to Faruk Gul and Wolfgang Pesendorfer for their invaluable advice during the development of the paper. I am particularly indebted to my main advisor, Wolfgang Pesendorfer, for his continuous support and guidance. I thank Roland Benabou, Eric Maskin, Stephen Morris and Klaus Nehring for their helpful discussions and comments. I have also benefited from suggestions made by Shiri Artstein-Avidan, Amir Bennatan, Bo'az Klartag, Charles Roddie and Kareen Rozen. Special thanks to Anne-Marie Alexander for all her help and encouragement.

†Department of Economics, Princeton University, Princeton, NJ 08544. E-mail: ddillenb@princeton.edu.

investment behavior systematically. All their subjects had to commit in advance to a fixed equal amount of investment for three subsequent periods. Group A was told that they would get periodic statements (i.e. would be informed about the outcome of the gamble after every draw), whereas group B knew that they would hear only the final yields of their investment. The average investment in group A was significantly lower than in group B. The authors conclude that "information feedback should be the variable of interest for researchers and actors in financial markets alike." Such interdependence between the way individuals observe the resolution of uncertainty and the amount of risk they are willing to take is not compatible with the standard model of decision making under risk, which is a theory of choice among probability distributions over final outcomes.[1]

In this paper, we make the assumption that the value of a lottery depends not only on its uncertainty, but also on the way this uncertainty is resolved over time. Using this assumption, we provide a choice theoretic framework that can address the experimental evidence above, while pinpointing the required deviations from the standard model. We exploit the structure of the model to identify the link between the temporal aspect of risk aversion, a static attitude towards risk, and intrinsic preferences for information.

In order to facilitate exposition, we mainly consider a decision maker (DM) whose preferences are defined over the set of two-stage lotteries, namely lotteries over lotteries over outcomes. Following Segal [1990], we replace the *reduction of compound lotteries axiom* (an axiom that imposes indifference between compound lotteries and their reduced single-stage counterparts) with the following two assumptions: *time neutrality*, which says that DM does not care about the time in which the uncertainty is resolved as long as resolution happens in a single stage, and *recursivity*, which says that the ranking of second-stage lotteries is unaffected by the first stage. Under these assumptions, any two-stage lottery is *subjectively* transformed into a simpler, one-stage lottery. In particular, there is a single preference relation defined over the set of one-stage lotteries that fully determines preferences over the richer domain of two-stage lotteries.

As a first step to link behavior in both domains, we introduce and formally define the following two properties: the first is dynamic while the second is static.

- *Preferences for one-shot resolution of Uncertainty* (PORU). DM has PORU if he always prefers any two-stage lottery to be resolved in a single stage. PORU implies an aversion to receiving partial information. This notion formalizes an idea first raised by Palacios-Huerta [1999] (to be further discussed in the literature review section). Such preferences

[1] All lotteries discussed in this paper are objective, that is, the probabilities are known. Knight [1921] proposed distinguishing between "risk" and "uncertainty" according to whether the probabilities are given to us objectively or not. Despite this distinction, we will interchangeably use both notions to express the same thing.

capture the idea that "the frequency at which the outcomes of a random process are evaluated" is a relevant economic variable.

- *Negative certainty independence* (NCI). NCI states that if DM prefers lottery p to the (degenerate) lottery that yields the prize x for certain, then this ranking is not reversed when we mix both options with any common third lottery q. This axiom is similar, but it is less demanding than Kahneman and Tversky's [1979] "Certainty effect" hypothesis, since it does not imply that people weight probabilities non-linearly. NCI imposes weak restrictions on preferences, just enough to explain commonly observed behavior in Allais-type experiments.

Theorem 1, our main result, establishes a tight connection between the two behavioral properties just described; NCI is a sufficient condition to PORU, and within the class of betweenness-satisfying preferences (Dekel [1986]), it is also necessary.

On the one hand, numerous replications of the Allais paradox in the last fifty years prove NCI to be one of the most prominently observed preference patterns. On the other hand, empirical and experimental studies involving dynamic choices and experimental studies on preference for uncertainty resolution are still rather rare. The disproportional amount of evidence in favor of each property strengthens the importance of theorem 1, since it provides new theoretical predictions for dynamic behavior, based on robust (static) empirical evidence.

Within the betweenness class, axiom NCI has its own static implications. First, it is equivalent to the following geometrical condition that is imposed on the map of indifference curves in every unit probability triangle (a diagram that represents the set of all lotteries over three fixed prizes):

- *Steepest middle slope* property: for every triple $x_3 > x_2 > x_1$, the indifference curve that passes through the origin (the lottery that yields x_2 for certain) is the steepest.

Since this geometrical condition is relatively easy to verify, it proves to be an applicable tool. Second, in theorem 2 we show that NCI is incompatible with the assumption that preferences are at least twice differentiable. When coupled with such a smoothness assumption, NCI turns out to be equivalent to the vNM-independence axiom.

In an extended model, we allow DM to take (just before the second-stage lottery is acted out, but after the realization of the first-stage lottery) intermediate actions that might affect his ultimate payoff. The primitive in such a model is a preference relation over *information systems*, which is induced from preferences over compound lotteries. An immediate consequence of Blackwell's [1953] seminal result is that in the standard expected utility class, DM always prefers to have perfect information before making the decision, which allows him to choose the optimal action corresponding to the resulting state. Safra and Sulganik [1995] left open the question of whether there are other preference relations for which, when applied recursively, a perfect information system is always the most valuable. We show that this property is equivalent to PORU. As a corollary, axiom NCI fully characterizes, within the betweenness class, such preferences for information.

The idea that individuals prefer one-shot resolution of uncertainty can be quantified. The *gradual resolution premium* of any compound lottery is the amount that DM would pay to replace that lottery by its single-stage counterpart. Similarly to the regular risk premium (the amount that DM would pay to replace one-stage lottery by its expected value), the gradual resolution premium is measured in monetary terms. The signs of these two variables need not agree, that is, positive risk premium does not imply and is not implied by positive gradual resolution premium. In the case where DM is both risk averse and displays PORU, however, these two forces magnify each other. We use this observation to explain why people often purchase dynamic insurance contracts, such as periodic insurance for electrical appliances and cellular phones, at much more than the actuarially fair rates.

The gradual resolution premium can be very significant, in the sense that if the resolution process is "long" enough, it might imply an extreme degree of risk aversion. To illustrate this, we first extend our results to preferences over arbitrary n-stage lotteries. We interpret the parameter n as the "resolution sensitivity" of an individual. It describes the frequency with which an individual updates information in a fixed time interval. Qualitatively, the results remain intact; DM who has preferences for one-shot resolution of uncertainty prefers to replace each (compound) sub-lottery with its single-stage counterpart. We then look at preferences of the disappointment aversion class (Gul [1991]). Such preferences satisfy NCI, and therefore, in a dynamic context, PORU. We show that for any one-stage lottery, there exists a multi-stage lottery (with the same probability distribution over the terminal prizes) whose value approximately equals the value of getting the worst prize for sure. While referring to the problem of repeated investment, Gollier [2001] states that "the central theoretical question of the link between the structure of the utility function and the horizon-riskiness relationship remained unsolved." The result above shows that preferences that display PORU may lead to excessively conservative investment strategies.

1.1. Related literature

Palacios-Huerta [1999] was the first to raise the idea that the form of the timing of resolution of uncertainty might be an important economic variable. By working out an example, he demonstrates that DM with Gul's [1991] disappointment aversion preferences will be averse to the sequential resolution of uncertainty, or, in the language of this paper, will be displaying

PORU. He also discusses a lot of potential applications. Ang, Bekaert and Liu [2005] use recursive disappointment aversion preferences to study a dynamic portfolio choice designed to maximize final wealth. The general theory we suggest provides an insightful way to understand exactly which attribute of Gul's preferences accounts for the resulting behavior. It also makes a clear distinction between two notions of disappointment: The common static notion of disappointment, as it appears in the literature, and the dynamic version implied by PORU (see section 3).

Loss aversion with narrow framing (also known as "myopic loss aversion") is a combination of two motives: loss aversion (Kahneman and Tversky [1979]), that is, people's tendency to be more sensitive to losses than to gains, and narrow framing (Kahneman and Tversky [1984]), that is, a dynamic aggregation rule that argues that when making a series of choices, individuals "bracket" them by making each choice in isolation. Benartzi and Thaler [1995] were the first to use this approach to suggest explanations for several economic "anomalies", such as the equity premium puzzle (Mehra and Prescott [1985]). Barberis and Huang [2005] and Barberis, Huang and Thaler [2006] generalize Benartzi and Thaler's work by assuming that DM derives utility directly from the outcome of a gamble over and above its contribution to total wealth.

The model presented here can be used to address similar phenomena. The combination of the folding-back procedure and a specific form of non-smooth atemporal preferences implies that individuals behave *as if* they intertemporally perform narrow framing. The gradual resolution premium quantifies this effect. The two approaches are conceptually different: Loss aversion with narrow framing brings to the forefront the idea that individuals evaluate any new gamble separately from its cumulative contribution to total wealth, while we maintain the assumption that terminal wealth matters, and identify narrow framing as a subjective temporal effect. In addition, we set aside the question of why individuals are sensitive to the way uncertainty is resolved (i.e. why they narrow frame), and construct a model that reveals the (context independent) behavioral implications of such considerations.

Rabin [2000] and Safra and Segal [2006] use calibration results to criticize a broad class of models of decision making under risk. They point out that modest risk aversion over small stakes gambles necessarily implies absurd levels of risk aversion over large stakes gambles. Our model resists these critiques. If most uncertainty resolves gradually, then it cannot be compounded into a single lottery. Our model implies first order risk aversion over each realized gamble, and therefore neither Rabin's nor Safra and Segal's arguments apply.

In this paper, we study time's effect on preferences by distinguishing between "one-shot" and "gradual" resolution of uncertainty. A different, but complementary, approach is to study intrinsic preferences for "early" or "late" resolution of uncertainty. This research agenda was initiated by Kreps and Porteus [1978], and later extended by Epstein and Zin [1989] and Epstein and Chew [1989] among others. Grant, Kajii and Polak [1998, 2000] connect preferences for the timing of resolution of uncertainty to intrinsic preferences for information. We believe that both aspects of intrinsic time preferences play a role in most real life situations. For example, an anxious student might prefer to know as soon as possible his final grade in an exam, but still prefers to wait (impatiently) rather than to get the grade of each question separately. The motivation to impose time neutrality is to demonstrate the role of the "one-shot" versus "gradual" effect, which has been neglected in the literature to date.

The remainder of the paper is organized as follows: we start section 2 by establishing our basic framework, after which we introduce the main behavioral properties of the paper and state our main characterization result. In section 3, we elaborate on the static implications of our model and provide examples. Section 4 first extends our results to preferences over compound lotteries with an arbitrarily finite number of stages. We then define the gradual resolution premium and illustrate its effect. In section 5, we relate our approach to the notion of loss aversion with narrow framing. Section 6 comments on the implications of our model on preferences over information systems. Section 7 is devoted to an application of our model to the area of investment under uncertainty. We present our concluding remarks in section 8. Most proofs are relegated to the appendix.

2. The model

2.1. Groundwork

Consider an interval $[w, b] = X \subset R$ of monetary prizes. Let $\mathcal{L}(X)$, or simply $\mathcal{L}^1$, be the set of all simple lotteries (lotteries with a finite number of outcomes) over X. Typical elements of $\mathcal{L}^1$ are denoted by p, q and r. If p, $q \in \mathcal{L}^1$ and $\alpha \in (0,1)$, then the mixture $\alpha p + (1-\alpha) q \in \mathcal{L}^1$ is the lottery that yields each x with probability $\alpha p_x + (1-\alpha) q_x$. We denote by δ_x the lottery that gives the prize x with certainty.

Denote by $\mathcal{L}(\mathcal{L}(X))$, or simply by $\mathcal{L}^2$, the set of all simple lotteries over $\mathcal{L}^1$. A typical element of $\mathcal{L}^2$ is $Q = \langle \alpha_1, q^1; ...; \alpha_l, q^l \rangle$ with $\alpha_j \geqslant 0$, $\sum_{j=1}^{l} \alpha_j = 1$ and $q^j \in \mathcal{L}^1$, $j = 1, 2, ..., l$. We call elements of $\mathcal{L}^2$ two-stage lotteries. We think of each $Q \in \mathcal{L}^2$ as a dynamic two-stage process where, in the first stage, a lottery q^j is chosen with probability α_j, and, in the second stage, a prize is obtained according to q^j.

Two special subsets of $\mathcal{L}^2$ are $\Gamma = \{\langle 1, q \rangle \mid q \in \mathcal{L}^1\}$ and $\Delta = \{\langle \alpha_{x_i}, \delta_{x_i} \rangle_{i=1}^{m},\ x_i \in X\}$. All lotteries in Γ and Δ are fully resolved in a single stage; in every member of Γ, no uncertainty is resolved in the first stage, whereas the uncertainty of every lottery in Δ is fully resolved in the first stage. Note that both Γ and Δ are isomorphic to $\mathcal{L}^1$.

Let $\mathcal{V}$ denote the set of all continuous and strictly monotone preference relations over (sets isomorphic to) $\mathcal{L}^1$, with a generic element $\succeq_1$. Each $\succeq_1 \in \mathcal{V}$ is represented by some continuous function $V : \mathcal{L}^1 \to \mathbb{R}$.[2]

Given V, the certainty equivalent of lottery p is a prize $\mathrm{c}_V(p)$ satisfying $p \sim_1 \delta_{\mathrm{c}_V(p)}$, where $\sim_1$ is the indifference relation induced from $\succeq_1$. By continuity and monotonicity, $\mathrm{c}_V : \mathcal{L}^1 \to X$ is well defined.

Let $\succeq$ be a preference relation over $\mathcal{L}^2$. Let $\succeq_\Gamma$ and $\succeq_\Delta$ be the restriction of $\succeq$ to Γ and Δ respectively. We assume throughout the paper that both $\succeq_\Gamma$ and $\succeq_\Delta$ are in $\mathcal{V}$. On $\succeq$ we impose the following axioms:

A1 (*time neutrality*): $\forall q \in \mathcal{L}^1$, $\langle 1, q \rangle \sim \langle q_{x_i}, \delta_{x_i} \rangle_{i=1}^m$

A2 (*recursivity*):

$$\left\langle \alpha_1, q^1; ..; \alpha_i, q^i; ..; \alpha_l, q^l \right\rangle \succeq \left\langle \alpha_1, q^1; ..; \alpha_i, \widetilde{q}^i; ..; \alpha_l, q^l \right\rangle \Longleftrightarrow \left\langle 1, q^i \right\rangle \succeq \left\langle 1, \widetilde{q}^i \right\rangle$$

By postulating **A1**, we assume that DM does not care about the time in which the uncertainty is resolved as long as it happens in a single stage. **A2** assumes that preferences are recursive. It states that preferences over two-stage lotteries respect the preference relation over single-stage lotteries, in the sense that two compound lotteries that differ only in the outcome of a single branch are compared exactly as these different outcomes would be compared separately.

[2] (i) A preference relation $\succeq$ on a set Z is a complete and transitive binary relation on Z.
(ii) A real valued function V represents the preference relation $\succeq$ on a set Z if for all $z_1, z_2 \in Z$, $z_1 \succeq z_2 \Leftrightarrow V(z_1) \geq V(z_2)$
(iii) Continuity is in the topology of weak convergence.
(iv) Monotonicity is with respect to the relation of first-order stochastic dominance.

Proposition (Segal [1990]): $\succeq$ *satisfies* **A1** *and* **A2** *iff it can be represented by a continuous function* $W : \mathcal{L}^2 \longrightarrow \mathbb{R}$ *of the following form:*

$$W\left(\left\langle \alpha_1, q^1; ...; \alpha_l, q^l \right\rangle\right) = V\left(\alpha_1 \delta_{c_V(q^1)} + ... + \alpha_l \delta_{c_V\left(q^l\right)}\right)$$

Note that under $A1$ and $A2$, the preference relation $\succeq_1 = \succeq_\Gamma = \succeq_\Delta$ fully determines $\succeq$. The decision maker evaluates two-stage lotteries by first calculating the certainty equivalent of every second-stage lottery using the preferences represented by V, and then calculating (using V again) the first-stage value by treating the certainty equivalents of the former stage as the relevant prizes. As only the function V matters, we drop its index from the certainty equivalents in the remainder of the paper. Furthermore, we slightly abuse notation by writing $V(Q)$, instead of $W(Q)$, for the value of the two-stage lottery Q. Lastly, since under the above assumptions $V(p) = V(\langle 1, p \rangle) = V(\langle q_{x_i}, \delta_{x_i} \rangle_{i=1}^m)$ for all $p \in \mathcal{L}^1$, we simply write $V(p)$ for this common value.

2.2. Main properties

We now introduce and motivate our two main behavioral assumptions. The first is dynamic, whereas the second is static.

2.2.1. Preference for one-shot resolution of uncertainty

We model an individual, DM, whose concept of uncertainty is multi-stage and who cares about the way uncertainty is resolved over time. In this section, we define consistent preferences to have all uncertainty resolved in "one-shot" rather than "gradually" or vice versa.

Fix $p \in \mathcal{L}^1$ and denote its support by $S(p)$, that is, $S(p) = \{x \, | p_x > 0\}$. Let

$$\mathcal{P}(p) := \left\{ \left\langle \alpha_i, p^i \right\rangle_{i=1}^K \in \mathcal{L}^2 \, \middle| K \in \mathbb{N} \text{ and } \forall x \in S(p) = \bigcup_i S(p_i), \, p_x = \textstyle\sum_{i=1}^K \alpha_i p_x^i \right\}.$$

$\mathcal{P}(p)$ is the set of all two-stage lotteries that induce the same probability distribution over final outcomes as p does. For example, if p is a lottery that gives the prize x_1 with probability 0.3 and the prize x_2 with the remaining probability, then the two-stage lottery $Q = \langle 0.6, q; 0.4, r \rangle$, where q gives both prizes with equal probability and r yields x_2 for sure, is in $\mathcal{P}(p)$. Let

$$\mathcal{P}^O(p) := \{Q \, | Q \in \mathcal{P}(p) \cap (\Gamma \cup \Delta)\} = \left\{ \langle 1, p \rangle, \ \ \langle p_x, \delta_x \rangle_{x \in S(p)} \right\}$$

$\mathcal{P}^O(p)$ contains all elements of $\mathcal{P}(p)$ that are resolved in a single stage.

Definition: $\succeq$ displays *preference for one-shot resolution of uncertainty* (PORU) if $\forall p \in \mathcal{L}^1$ and $\forall Q \in \mathcal{P}(p)$, $R \in \mathcal{P}^O(p)$ implies $R \succeq Q$. If, subject to the same qualifiers, $R \in \mathcal{P}^O(p)$ implies $Q \succeq R$, then $\succeq$ displays *preference for gradual resolution of uncertainty* (PGRU).

PORU implies an aversion to receiving partial information. If uncertainty is not fully resolved in the first stage, DM prefers to remain fully unaware till the final resolution is available. PGRU implies the opposite. As we will argue in later sections, these notions render "the frequency at which the outcomes of a random process are evaluated" a relevant economic variable.

2.2.2. The Allais paradox and axiom NCI

In a generic Allais-type questionnaire,[3] subjects choose between A and B, where $A = \delta_{300}$ and $B = 0.8\delta_{400} + 0.2\delta_0$. They also choose between C and D, where $C = 0.25\delta_{300} + 0.75\delta_0$ and $D = 0.2\delta_{400} + 0.8\delta_0$. The majority of subjects tend to *systematically* violate expected utility by choosing the pair A and D.

Since Allais's [1953] original work, numerous versions of his questionnaire have appeared, most of which contain one lottery that does not involve any risk. Kahneman and Tversky use the term "certainty effect" to explain the commonly observed behavior. Their idea is that individuals tend to put more weight on certain events in comparison with very likely, yet uncertain, events. Although verbally it might appear to be intuitive reasoning, it is behaviorally translated into a nonlinear probability-weighting function, $\pi : [0,1] \rightarrow [0,1]$, that individuals are assumed to use when evaluating risky prospects. In particular, this function has a steep slope near –or even a discontinuity point at– 0 and 1. As we remark below, this implication has its own limitations. We thus suggest a direct behavioral property that is motivated by similar insights, but is less restrictive. Consider the following axiom on $\succeq_1$:

Negative Certainty Independence (NCI):[4] $\forall p, q, \delta_x \in \mathcal{L}^1$ and $\lambda \in [0,1]$, $p \succeq_1 \delta_x$ implies $\lambda p + (1-\lambda) q \succeq_1 \lambda \delta_x + (1-\lambda) q$.

The axiom states that if the sure outcome x is not enough to compensate DM for the risky prospect p, then mixing it with any other lottery, thus eliminating its certainty appeal, will not result in the mixture of x being more attractive than the corresponding mixture

[3] Also known as "common-ratio effect with a certain prize."

[4] We use the word "negative" since this axiom can, equivalently, be stated as: $\forall p, q, \delta_x \in \mathcal{L}^1$ and $\lambda \in [0,1)$, $\delta_x \not\succ_1 p$ *implies* $\lambda \delta_x + (1-\lambda) q \not\succ_1 \lambda p + (1-\lambda) q$. Here $\succ_1$is the asymmetric part of $\succeq_1$, and $\not\succ_1$ is its negation.

of p. The implication of this axiom on responses in Allais questionnaire above is: If you choose B, then you must choose D. This prediction is empirically rarely violated (see for example "pattern 2" in Conlisk [1989]). As stated above, the intuition behind NCI is that the sure outcome loses relatively more (or gains relatively less) than any other lottery from the mixture with the other lottery, q, but it does not imply any probabilistic distortion. This becomes relevant in experiments like those of Conlisk [1989], who studies the robustness of Allais-type behavior to small perturbations of the questionnaire which remove boundary effects. Although violations in that case were no longer systematic, a nonlinear probability function, as suggested above, predicts that this increase in consistency would be the result of fewer subjects choosing (the slightly perturbed) A over B, and not because more subjects choose (the slightly perturbed) C over D. In fact, the latter occurred, which is consistent with NCI.

Proposition 1: *Under* **A1** *and* **A2**, *if* $\succeq_1$ *satisfy NCI, then* $\succeq$ display PORU

Proof : We need to show that an arbitrary two-stage lottery, $\langle \alpha_1, q^1; ...; \alpha_l, q^l \rangle$, is never preferred to its single-stage counterpart, $\left\langle 1, \sum_{i=1}^{l} \alpha_i q^i \right\rangle$. Using **A1** *and* **A2** we have:

$$\langle \alpha_1, q^1; ...; \alpha_l, q^l \rangle \overset{(A_2)}{\sim} \left\langle \alpha_1, \delta_{c(q^1)}; ...; \alpha_l, \delta_{c\left(q^l\right)} \right\rangle \overset{(A_1)}{\sim} \left\langle 1, \sum_{i=1}^{l} \alpha_i \delta_{c(q^i)} \right\rangle$$

And by repeatedly applying NCI,

$$\begin{aligned}
&\textstyle\sum_{i=1}^{l} \alpha_i \delta_{c(q^i)} = \alpha_1 \delta_{c(q^1)} + (1-\alpha_1)\left(\sum_{i\neq 1} \frac{\alpha_i}{(1-\alpha_1)} \delta_{c(q^i)}\right) \overset{(\text{NCI})}{\preceq_1} \\
&\alpha_1 q^1 + (1-\alpha_1)\left(\sum_{i\neq 1} \frac{\alpha_i}{(1-\alpha_1)} \delta_{c(q^i)}\right) = \\
&\alpha_2 \delta_{c(q^2)} + (1-\alpha_2)\left(\frac{\alpha_1}{(1-\alpha_2)} q^1 + \sum_{i\neq 1,2} \frac{\alpha_i}{(1-\alpha_2)} \delta_{c(q^i)}\right) \overset{(\text{NCI})}{\preceq_1} \\
&\alpha_1 q^1 + \alpha_2 q^2 + \textstyle\sum_{i\neq 1,2} \alpha_i \delta_{c(q^i)} = ... = \\
&\alpha_l \delta_{c\left(q^l\right)} + (1-\alpha_l)\left(\sum_{i\neq l} \frac{\alpha_j}{(1-\alpha_l)} q^i\right) \overset{(\text{NCI})}{\preceq_1} \textstyle\sum_{i=1}^{l} \alpha_i q^i
\end{aligned}$$

Therefore, $\langle \alpha_1, q^1; ...; \alpha_l, q^l \rangle \sim \left\langle 1, \sum_{i=1}^{l} \alpha_i \delta_{c(q^i)} \right\rangle \preceq \left\langle 1, \sum_{i=1}^{l} \alpha_i q^i \right\rangle$ ■

The idea behind proposition 1 is simple: the second step of the folding-back procedure involves mixing all certainty equivalents of the corresponding second-stage lotteries. Applying NCI repeatedly implies that each certainty equivalent suffers from the mixture at least as much as the original lottery that it replaces would.

Proposition 1 states that NCI is a sufficient condition for PORU. To show necessity, we need to impose more structure. For the rest of the section, we confine our attention to a class of preferences $\succeq \in \mathcal{V}$ that satisfy the betweenness axiom.

A3 (*single-stage betweenness*) $\forall\ p, q \in \mathcal{L}^1$ and $\alpha \in [0,1]$, $\ p \succeq_1 q$ implies $p \succeq_1 \alpha p + (1-\alpha)\, q \succeq_1 q$

A3 is a weakened form of the vNM-independence axiom. It implies neutrality toward randomization among equally-good prizes/lotteries. It yields the following representation:

Proposition (Dekel [1986]): $\succeq_1 \in V$ *satisfies* **A3** *iff there exists a local utility function* $u : X \times [0,1] \to [0,1]$*, which is continuous in both arguments, strictly increasing in the first argument and satisfies* $u\,(w, v) = 0$ *and* $u\,(b, v) = 1$ *for all* $v \in [0,1]$*, such that for all* $p \in \mathcal{L}^1$*,* $V(p)$ *is defined implicitly by:*

$$V(p) = \sum_{x \in X} u\,(x, V(p))\, p_x$$

NCI in the probability triangle

The betweenness axiom (**A3**), along with monotonicity, implies that indifference curves in any unit probability triangle are positively sloped straight lines. To demonstrate this result using the representation theorem, note that for any lottery p over a given triple $x_3 > x_2 > x_1$, $V(p) = p_1 u(x_1, V(p)) + (1 - p_1 - p_3) u(x_2, V(p)) + p_3 u(x_3, V(p))$. The slope of any indifference curve in the corresponding two-dimensional space, $\Delta := \{(p_1, p_3) \mid p_1, p_3 \geq 0,\ p_1 + p_3 \leq 1\}$ is given by:

$$\mu\,(V|x_3, x_2, x_1) = \frac{u(x_2, v) - u(x_1, v)}{u(x_3, v) - u(x_2, v)}$$

which is positive and independent of the vector of probabilities. By definition, the slope represents the marginal rate of substitution between p_3 and p_1, and as explained by Machina [1982], changes in the slope express local changes in attitude towards risk: the greater the slope, the more risk averse DM is.

Definition: $\succeq_1$ has the *steepest middle slope* property if for every triple $x_3 > x_2 > x_1$ and for all $v \in (V(\delta_{x_1}), V(\delta_{x_3}))$,

$$\mu\,(V(\delta_{x_2})\,|x_3, x_2, x_1) \geq \mu\,(V|x_3, x_2, x_1)$$

that is to say, this property holds if for every three prizes $x_3 > x_2 > x_1$, the indifference curve through δ_{x_2} is the steepest.

Observe that NCI implies the steepest middle slope property. To see this, let $I_{V(\delta_{x_2})} := \{p' \in \Delta : p' \sim_1 \delta_{x_2}\}$ and let $\overline{\mu} := \mu\left(V(\delta_{x_2})\,|x_3, x_2, x_1\right)$. Take any lottery $p \in I_{V(\delta_{x_2})}$. For any $\lambda \in [0,1]$ and $q \in \Delta$, both $\lambda p + (1-\lambda)q$ and $\lambda\delta_x + (1-\lambda)q$ are in Δ (a convex set) and by the triangle proportional sides theorem, the line segment that connects them has a slope that equals $\overline{\mu}$. But NCI requires that $\lambda p + (1-\lambda)q \succeq_1 \lambda\delta_x + (1-\lambda)q$ and since indifference curves are upward sloping, the indifference curve that passes through $\lambda p + (1-\lambda)q$ must have a slope no greater than $\overline{\mu}$. Since λ and q were arbitrary, the result follows.

2.3. Characterization

Definition : $\succeq_2$ is *betweenness-recursive* if it satisfies **A1** – **A2** and its restrictions to $\mathcal{V}$ satisfy **A3**.

Theorem 1: *For any betweenness-recursive preferences, the following three statements are equivalent:*

(i) $\succeq$ *displays PORU.*

(ii) $\succeq_1$ *satisfies NCI.*

(iii) $\succeq_1$ *has the steepest middle slope property.*

A characterization of PGRU is analogously obtained by reversing the weakly preferred sign in NCI, and replacing steepest with flattest in (iii) .

The detailed proof is in the appendix. The main step in it is to establish, using certain properties of preferences from the betweenness class, that PORU is equivalent to the following condition:

$$\mathbf{C}_1 : \left[\sum\nolimits_{x\in S(p)} u\,(x, v)\,p_x - u\,(\mathrm{c}\,(p)\,, v)\right] \geqslant 0 \;\forall p \in \mathcal{L}^1 \text{ and } \forall v \in V\!\left(\mathcal{L}^1\right)$$

where $V(\mathcal{L}^1) := \{V|\exists p \in \mathcal{L}^1 \text{ with } v = V(p)\}$.[5] We interpret $\mathbf{C}_1$ by exploiting the main idea behind the construction of the local utility function, $u\,(x, v)$. As explained by Dekel [1986],

[5] The specific normalization $V\left(\mathcal{L}^1\right) = [0, 1]$ is inessential for this result.

and demonstrated in figure 1, one can think of $u(x,v)$ as a collection of functions that are derived in the following way: Fix an indifference hyperplane with a value v (denoted by I_v in figure 1a) and construct a collection of parallel hyperplanes relative to it. This collection can be taken to represent some expected utility preferences with an associated Bernoulli function $u_v(x)$. For every lottery p, we can then calculate $V(p,v) := E_p[u_v(x)]$, its expected utility relative to the value v (figure 1b). Repeat this construction for every value of v (which is bounded above and below, since X is bounded) to get the collection of functions $\{u_v(x)\}_{v \in V(\mathcal{L}^1)}$ that are equal to $u(x,v)$. $\mathbf{C}_1$ then implies that DM becomes the most risk averse at the true lottery value. That is, if relative to $V(p)$, the true utility level, DM is just indifferent between p and the certain prize c(p), then relative to any other value v, he (weakly) prefers the lottery. The graphical illustration of $\mathbf{C}_1$ in the probability triangle is precisely item (iii) in the theorem (figure 1c), whereas item (ii) is its direct behavioral interpretation. The proof is completed by ensuring sufficiency of item (iii) to $\mathbf{C}_1$ and using proposition 1.

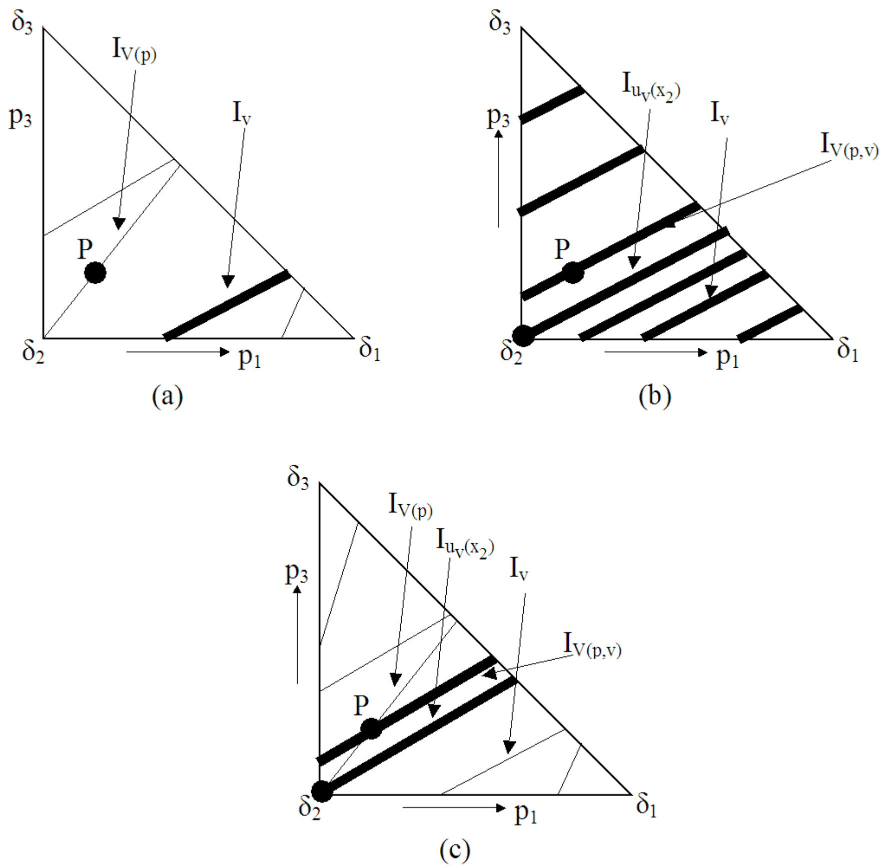

Figure 1: 1a: Fixing an indifference curve of level v. 1b: Constructing the local utility function $u_v(x)$. 1c: Putting them together, $E_p[u(x,V(p))] = u(x_2,V(p))$, but $E_p[u(x,v)] > u(x_2,v)$ for $v \neq V(p)$.

Theorem 1 ties together three notions that are defined on different domains: PORU is a dynamic property, NCI is a static property, and the third item is a geometrical condition, which applies to single-stage lotteries with at most three prizes in their support. The core of the theorem is the equivalence of PORU and NCI, which suggests that being prone to Allais-type behavior and being averse to the gradual resolution of uncertainty are synonymous. This assertion justifies the proposed division of the space of two-stage lotteries into the one-shot and gradually resolved lotteries. On the one hand, numerous replications of the Allais paradox in the last fifty years prove that the availability of a certain prize in the choice set is important and affects behavior in a systematic way. Moreover, we have no firm evidence of a consistent attitude towards lotteries, all of which lie in the interior of a probability triangle. On the other hand, empirical and experimental studies involving dynamic choices and experimental studies on preference for uncertainty resolution are still rather rare. Theorem 1 thus provides new theoretical predictions for dynamic behavior, based on robust (static) empirical evidence.

The applicability of the steepest middle slope property stems from its simplicity. In order to detect violation of PORU, one need not construct the (potentially complicated) exact choice problem. Rather, it is sufficient to introspect the slopes of one-dimensional indifference curves. This, in turn, is a relatively simple task, at least once a local utility function is given.

3. Static implications

3.1. NCI and differentiability

In most economic applications, it is assumed that individuals' preferences, and therefore the utility functions that represent them, are not only continuous, but also at least twice differentiable.[6] The following result demonstrates that among the betweenness class, smoothness and NCI are inconsistent, in the sense that coupling them leads us back to expected utility.

Theorem 2: *Suppose $u(x,v)$ is at least twice differentiable with respect to both its arguments, and that all derivatives are continuous and bounded. Then preferences satisfy NCI if and only if they are expected utility.*

Expected utility preferences are characterized by the independence axiom that implies NCI. To show the other direction, we fix $\overline{v}$ and denote by $x(\overline{v})$ the unique x satisfying $\overline{v} = u(x,\overline{v})$. Combining the geometrical characterization (theorem 1 item (iii)) of NCI with

[6] Debreu [1972] provides, for any $k > 0$, a formal definitions of k^{th}- order differentiable preferences.

differentiability implies that for any $x > x(\overline{v}) > w$, the derivative with respect to v of the slope of an indifference curve on the corresponding probability triangle must vanish at $\overline{v}$. We use the fact that this statement is true for any $x > x(\overline{v})$ and that $\overline{v}$ is arbitrary to get a differential equation with a solution on $\{(x,v)\,|v < u(x,v)\}$ given by $u(x,v) = h^1(v)\,g^1(x) + f^1(v)$, and $h^1(v) > 0$. We perform a similar exercise for $x < x(\overline{v}) < b$ to uncover that on the other region, $\{(x,v)\,|v < u(x,v)\}$, $u(x,v) = h^2(v)\,g^2(x) + f^2(v)$, and $h^2(v) > 0$. Continuity and differentiability then imply that the functional form is equal in both regions, therefore for all x, $u(x,v) = h(v)\,g(x) + f(v)$, and $h(v) > 0$. The uniqueness theorem for betweenness representations establishes the result.

3.2. Examples

Expected utility preferences are a trivial example of preferences that in a dynamic context satisfy PORU; DM with such preferences is just indifferent to the way uncertainty is resolved. The following is an important class of preferences for which, when applied recursively, PORU is a meaningful concept:

Preferences that satisfy the mixed-fan hypothesis. This set consists of all preferences whose indifference curves, in any unit probability triangle, have the following pattern: Moving northwest, they first get steeper ("fanning out") in the lower-right sub triangle (the less-preferred region), and then get flatter ("fanning in") in the upper-left sub triangle (the more-preferred region). Before giving examples from this class, we first state sufficient restrictions on the local utility function to satisfy the mixed-fan hypothesis.[7]

Denote by $L(x) := \{p \in \mathcal{L}^1 : \delta_x \succsim_1 p\}$ the lower contour set of $x \in X$.

Sufficient conditions for mixed fan: *If $u(x,v)$ is a local utility function of the form*

$$u(x,v) - v = \begin{cases} u^1(x,v) & V^{-1}(v) \in L(x) \\ u^2(x,v) & V^{-1}(v) \notin L(x) \end{cases}$$

with the following restrictions:

(1) $\frac{\partial}{\partial x}\frac{\partial}{\partial v}u^1(x,v) \leqslant 0$,

(2) $\frac{\partial}{\partial x}\frac{\partial}{\partial v}u^2(x,v) \leqslant 0$, *and*

(3) $\inf_x \frac{\partial}{\partial v}u^1(x,v) \geqslant \sup_x \frac{\partial}{\partial v}u^2(x,v)$

then preferences satisfy the mixed-fan hypothesis.

[7] Neilson [1992] provides sufficient conditions for smooth (in the sense of theorem 2) betweenness preferences to satisfy the mixed-fan hypothesis. The additional requirement, that the switch between "fanning out" and "fanning in" always occurs at the indifference curve that passes through the origin (the lottery that yields the middle prize for certain), renders those conditions empty, as is evident from theorem 2.

Chew [1989] axiomatizes semi-implicit weighted utility. The local utility function he considers is

$$u(x,v) - v = \begin{cases} \overline{w}(x)(x-v) & x > v \\ \underline{w}(x)(x-v) & x \leqslant v \end{cases}$$

with $\underline{w}(x) > 0$, $\overline{w}(x) > 0$, $\underline{w}'(x) \geqslant 0$, $\overline{w}'(x) \geqslant 0$. To ensure that these preferences satisfy the mixed-fan hypothesis, we add the restriction that $\inf_x \underline{w}(x) > \sup_x \overline{w}(x)$.

Gul [1991] proposes a theory of disappointment aversion. He derives the local utility function

$$u(x,v) = \begin{cases} \frac{\phi(x)+\beta v}{1+\beta} & \phi(x) > v \\ \phi(x) & \phi(x) \leqslant v \end{cases}$$

with $\beta > 0$ and $\phi : X \to \mathbb{R}$ increasing.

Gul's notion of disappointment aversion amounts to dividing the support of each lottery into two groups, the elated outcomes and the disappointed outcomes, and giving the disappointed outcomes a uniformly greater weight when calculating the expected utility of the lottery.[8] For these preferences, the sign of β, the coefficient of disappointment aversion, unambiguously determines whether preferences satisfy PORU or PGRU (see Artstein-Avidan and Dillenberger [2006]).

PORU can be interpreted as dynamic disappointment aversion. As suggested by Palacios-Huerta [1999], one may argue that being exposed to the resolution process bears the risk of perceiving intermediate outcomes as disappointing or elating, and if DM is more sensitive to disappointments, he would prefer to know only the final result. The term "disappointment aversion preferences" usually refers to Gul's static model. Our dynamic notion of disappointment aversion is translated into a strong restriction on indifference maps across probability triangles. Although Gul's model satisfies both, it is a boundary case. To emphasize the distinction between these two notions, we provide examples of other betweenness-satisfying preferences that were suggested as one-parameter generalizations of Gul's static model but, nevertheless, dynamically violate PORU. As implied by theorem 1, to track down violations of PORU, it is enough to show that neither of the preferences below satisfy the steepest middle slope property.[9]

[8] Although Gul's preferences imply probability transformation, this transformation is done endogenously. It is the value of each elated prize, and not its probability, which is explicitly down-weighted.

[9] Gul's preferences are one parameter (β) richer than expected utility preferences. The economic interpretation of β in a dynamic context is not evident. Indeed, one of Gul's axioms (axiom 4 in his paper) is necessary to identify β, but is unrelated to NCI. It is imposed in order to rule out further deviations from the expected utility model.

Nehring [2005] suggests preferences which are represented by an implicit utility function of the following form:

$$u(x,v)-v=\begin{cases}(\phi(x)-\phi(v))^{\alpha} & x>v\\ -\beta(\phi(v)-\phi(x))^{\alpha} & x\leqslant v\end{cases}$$

with $\alpha,\beta>0$. Gul's model corresponds to the case of $\alpha=1$, and disappointment aversion implies $\beta>1$.

Nehring interprets $u(x,v)-v$ as relative utilities (outcomes are evaluated psychologically relative to a certain reference point) and $\phi(x)$ as absolute utilities. He shows that such a class is uniquely characterized by the "bi-linearity" property: There exists a monotonic and continuous function $\pi:[0,1]\to[0,1]$ and a mapping $\phi:X\to\mathbb{R}$, such that preferences restricted to binary lotteries are represented by the function $V(p\delta_x+(1-p)\delta_y)=\pi(p)\phi(x)+(1-\pi(p))\phi(y)$, for $x>y$.
Unless $\alpha=1$ (and $\beta\geqslant1$), no member of this class of preferences satisfies NCI.

Routledge and Zin [2004] provide a different one-parameter extension of Gul's model, enabling the identification of outcomes as disappointing only when they lie sufficiently below the (implicit) certainty equivalent. They derive the representation:

$$\phi(\mathrm{c}(p))=\sum\nolimits_{x_i}p(x_i)\phi(x_i)-\beta\sum\nolimits_{x_i\leqslant\delta\mathrm{c}(p)}p(x_i)[\phi(\delta\mathrm{c}(p))-\phi(x_i)]$$

with $\delta\leqslant1$. Note that in Gul's model $\delta=1$ (where $\beta=0$ corresponds to expected utility). Unless $\delta=1$, these preferences also do not satisfy NCI.

4. Gradual resolution premium

For further purposes, we first extend our results to finite-stage lotteries.

4.1. Extension to n-stage lotteries

Fix $n\in\mathbb{N}$ and denote the space of finite n-stage lotteries by $\mathcal{L}^n$. We interpret the parameter n as the "resolution sensitivity" of an individual. It describes the frequency with which an individual updates information in a fixed time interval, which is a characteristic of preferences. The extension of our setting to $\mathcal{L}^n$ is the following (a formal description is given in the appendix): Occupied with a continuous and increasing function $V:\mathcal{L}^1\to\mathbb{R}$, DM evaluates any n-stage lottery by folding back the probability tree and applying the same V in each stage. Preferences for one-shot resolution of uncertainty implies that DM prefers to replace each (compound) sub-lottery with its single-stage counterpart. The equivalence between PORU and NCI remains intact. In what follows, we will continue simplifying notation by writing $V(Q)$ for the value of any multi-stage lottery Q. We sometimes write Q^n to emphasize that we consider an n-stage lottery.

4.2. Definitions

Denote by $\mathrm{e}(p)$ the expectation of a lottery $p \in \mathcal{L}^1$, that is, $\mathrm{e}(p) = \sum_x x p_x$. Let $G(p,x) := \sum_{z \geqslant x} p_z$. We say that lottery p second-order stochastically dominates lottery q, and denote it by p sosd q, if for all $t < K$, $\sum_{k=0}^{t} [G(p, x_{k+1}) - G(q, x_{k+1})][x_{k+1} - x_k] \geqslant 0$, where $x_0 < x_1 < ... < x_K$ and $\{x_0, x_1, ..., x_K\} = S(p) \cup S(q)$. DM is risk averse if $\forall p, q \in \mathcal{L}^1$ with $\mathrm{e}(p) = \mathrm{e}(q)$, p sosd q implies $p \succeq_1 q$.

For any $p \in \mathcal{L}^1$, the risk premium of p, denoted by $\mathrm{rp}(p)$, is the number satisfying $\delta_{\mathrm{e}(p)-\mathrm{rp}(p)} \sim_1 p$. $\mathrm{rp}(p)$ is the amount that DM would pay to replace p with its expected value. By definition, $\mathrm{rp}(p) \geqslant 0$ whenever DM is risk averse.[10]

Definition: Fix $p \in \mathcal{L}^1$. For any $Q \in \mathcal{P}(p)$, the *gradual resolution premium* of Q, denoted by $\mathrm{grp}(Q)$, is the number satisfying $\langle 1, \delta_{c(p)-\mathrm{grp}(Q)} \rangle \sim Q$.

$\mathrm{grp}(Q)$ is the amount that DM would pay to replace Q with its single-stage counterpart. By definition, PORU implies $\mathrm{grp}(Q) \geqslant 0$. Since $c(p) = \mathrm{e}(p) - \mathrm{rp}(p)$, we can, equivalently, define $\mathrm{grp}(Q)$ as the number satisfying $\langle 1, \delta_{\mathrm{e}(p)-\mathrm{rp}(p)-\mathrm{grp}(Q)} \rangle \sim Q$.[11]

Observe that the signs of the two variables above, $\mathrm{rp}(p)$ and $\mathrm{grp}(Q)$, need not agree. In other words, (global) risk aversion does not imply, and is not implied by, PORU. Indeed, Gul's symmetric disappointment aversion preferences (see section 3) are risk averse if and only if $\beta \geqslant 0$ and $\phi : X \to \mathbb{R}$ is concave (Gul's [1991] theorem 3). However, for sufficiently small $\beta \geqslant 0$ and sufficiently convex ϕ, one can find a lottery p with $\mathrm{rp}(p) < 0$, whereas $\beta \geqslant 0$ is sufficient for $\mathrm{grp}(Q) \geqslant 0$ for any $Q \in \mathcal{P}(p)$. On the other hand, if $\lambda'(v) > 0$ and $\lambda(v) > 1$ for all v,[12] then the local utility function

$$u(x,v) = \begin{cases} x & x > v \\ v - \lambda(v)(v - x) & x \leqslant v \end{cases}$$

[10] Weak risk aversion is defined as follows: For all p, $\delta_{e(p)} \succeq p$. This definition is not appropriate once we consider preferences that are not expected utility. The definition of the risk premium, on the other hand, is independent of the preferences considered.

[11] Similarly to the risk premium, the complete resolution premium is measured in monetary units. For this reason, these two premiums are different from the timing premium for early resolution, as suggested by Chew and Epstein [1989], which is measured in terms of probabilities.

[12] The condition that $\lambda(v)$ is non-decreasing is both necessary and sufficient for $u(,)$ to be a local utility function. See Nehring [2005].

has the property that $u(\cdot, v)$ is concave for all v. Therefore, DM is globally risk averse (Dekel's [1986] property 2), and hence $\text{rp}(p) \geqslant 0$ $\forall p \in \mathcal{L}^1$. However, these preferences do not satisfy NCI,[13] meaning that there exists a lottery p and $Q \in \mathcal{P}(p)$ for which $\text{grp}(Q) < 0$.

4.3. The magnifying effect

In the case where DM is both risk averse and has PORU, these two forces, as reflected in the two premiums previously defined, magnify each other. Understanding this, insurance companies, when offering dynamic insurance contracts, can require much greater premiums than the actuarially fair ones and still be sure of consumers' participation. This can explain why people often buy periodic insurance for moderately priced objects, such as electrical appliances and cellular phones, at much more than the actuarially fair rates. [14]

To illustrate, consider the following insurance problem: An individual with Gul's preferences, with a linear ϕ and a positive coefficient of disappointment aversion β, owns an appliance (e.g. a cellular phone) that he is about to use for n periods. The individual gets utility 1 in any period the appliance is used and 0 otherwise. In each period, there is an exogenous probability $(1-p)$ that the appliance will not work (it might be broken, fail to get reception, etc.). The individual can buy a periodic insurance, which guarantees the availability of the appliance, for a price $z \in (1-p, 1)$. Therefore, if he buys insurance for some period, he gets a certain utility of $(1-z)$, and otherwise he faces the lottery in which with a probability p he gets 1, and with the remaining probability he gets 0. For simplicity, assume that the price of a replacement appliance is 0, so that the individual either still has it from the last period or gets a new one for free in the beginning of any period.

Let $\widehat{p}$ be the probability distribution over final outcomes (without insurance). Denote by X the total number of periods in which the appliance works. Since X is a binomial random variable, $\Pr(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$, for $k = 0, ..., n$. Applying Gul's formula, one obtains:

$$V_{\beta,n}(\widehat{p}) = \frac{\sum_{k=h+1}^{n} \binom{n}{k} p^k (1-p)^{n-k} k + (1+\beta) \sum_{k=0}^{h} \binom{n}{k} p^k (1-p)^{n-k} k}{1 + \beta \sum_{k=0}^{h} \binom{n}{k} p^k (1-p)^{n-k}}$$

[13] Look at the slope of an indifference curve for values $x_3 > v > x_2 > x_1$. We have: $\mu(V|x_3, x_2, x_1) = \frac{\lambda(v)(x_2 - x_1)}{x_3 - v + \lambda(v)(v - x_2)}$. In this region, the slope is increasing in v if $x_3 > \frac{\lambda(v)(\lambda(v)-1)}{\lambda'(v)} + v$. For a given v, we can always choose arbitrarily large x_3 that satisfies the condition, and construct, by varying the probabilities, a lottery whose value is equal to v. Apply this argument in the limit where $v = x_2$ to violate condition (*iii*) of theorem 1.

[14] A popular example is given by Tim Harford ("The Undercover Economist", *Financial Times*, May 13, 2006): "There is plenty of overpriced insurance around. A popular cell phone retailer will insure your $90 phone for $1.70 a week—nearly $90 a year. The fair price of the insurance is probably closer to $9 a year than $90."

where $h(p,\beta,n)$ is the unique natural number such that all prizes greater than it are elated and all those smaller than it are disappointed.

Let Q be the corresponding gradual (n-stage) lottery *as perceived by DM.* Its value is:

$$V_{\beta,n}(Q) := \frac{1}{(1+\beta(1-p))^n} \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} (1+\beta)^{n-k} k$$

Using standard backward induction arguments, it can be shown that DM will buy insurance for all periods if $\beta > \frac{z-(1-p)}{(1-z)(1-p)} > 0$. In that case, $z < 1 - \frac{V_{\beta,n}(Q)}{n}$. Nevertheless, if β is not too high,[15] we have $1-p < 1 - \frac{V_{\beta,n}(\widehat{p})}{n} < z$, meaning that DM would not buy insurance at all *if he could avoid being aware of the gradual resolution of uncertainty.*[16] This observation explains why and how the attractiveness of a lottery depends not only on the uncertainty embedded in it, but also on the way this uncertainty is resolved over time.

Since $V_{\beta,n}(\widehat{p})$ decreases with β, $\text{rp}(\beta\,|p,n) := np - V_{\beta,n}(\widehat{p})$ is a strictly increasing function of β. The behavior of the gradual resolution premium, $\text{grp}(\beta\,|p,n) := V_{\beta,n}(\widehat{p}) - V_{\beta,n}(Q)$ is more subtle. We have the following result:

Proposition 2: *In the insurance problem described above:*

(*i*) *Strict PORU in the interior:* $\text{grp}(\beta\,|p,n) > 0\ \forall\beta \in (0,\infty)$

(*ii*) *Weak PORU in the extreme:* $\text{grp}(0\,|p,n) = 0$ *and* $\lim_{\beta\to\infty} \text{grp}(\beta\,|p,n) = 0$

(*iii*) *Single-peakness*: *There exists* $\beta^*(p,n) < \infty$ *such that either* $0 < \beta < \beta' < \beta^*$ *or* $\beta^* < \beta' < \beta$ *implies*

$$\text{grp}(\beta\,|p,n) < \text{grp}(\beta'\,|p,n) < \text{grp}(\beta^*\,|p,n)$$

See figure 2.

Recall that in Gul's model, the sign of the parameter β unambiguously determines whether preferences display PORU or PGRU. In its original context, greater β implies greater disappointment aversion (as well as greater risk aversion). Since we argued that PORU can be interpreted as dynamic disappointment aversion, it might seem intuitive to expect the gradual resolution premium to be an increasing function of β. This intuition is wrong and, in fact, item (ii) remains valid independent of the decision problem under consideration. In order to see this, note that $\text{grp}(\beta\,|p,n)$ is defined as the difference of two functions, both strictly decreasing with β. When $\beta = 0$, DM cares only about the expected value of the lottery. When β is sufficiently large, all prizes but 0 become elated, and hence the value of p converges to 0. Correspondingly, the value of the gradual lottery converges to

[15] The condition is: $1+\beta < \min\left\{\frac{p^n}{p^n+n(1-p)-1}, \frac{p^n z}{(1-z)(1-p^n)-p(1-p^{n-1})-1}\right\}$.

[16] Nayyar [2004] termed such a situation an "insurance trap". Note that DM still acts rationally given that without insurance he is forced to be exposed to Q^n rather than to p.

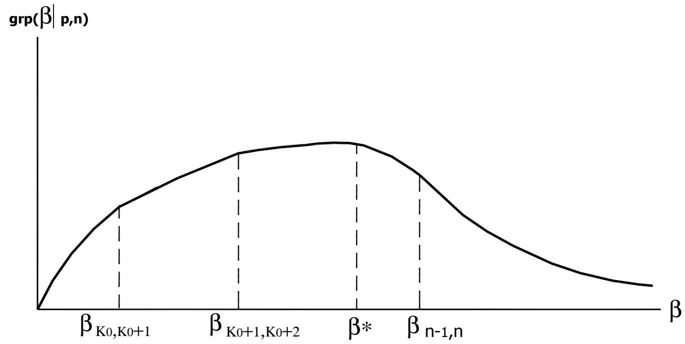

Figure 2: grp$(\beta\,|p,n)$. $\beta_{k,k+1}$ is the value of β where $h\,(\beta\,|p,n)$ decreases from $(n-k)$ to $(n-(k+1))$. grp$(\beta\,|p,n)$ is non-differentiable in each such $\beta_{k,k+1}$. k_0 is the smallest natural number that solves: $\max\limits_{k'>n(1-p)} \frac{n-k'}{n}$

the value of the worst sub-lottery that by itself approaches 0. Since item (i) reinforces the result of theorem 1 and states that grp$(\beta\,|p,n)$ is actually strictly positive on the positive reals, and since grp$(\beta\,|p,n)$ is a continuous function, there must exist a finite β, denoted β^* in figure 2, in which grp$(\beta\,|p,n)$ is maximized. Item (iii) sheds further light on the behavior of moderate disappointment-averse individuals. It suggests that $\beta^*\,(p,n)$ is unique, and that grp$(\beta\,|p,n)$ is single-peaked. Behaviorally speaking, a moderately disappointment-averse individual is more inclined to pay a higher premium, whereas individuals, who are either approximately disappointment-indifferent or very disappointment-averse, would not pay a substantial premium.

The analysis of the insurance problem suggests that, given n, extreme values of β neutralize the magnifying effect of the gradual resolution premium. In general, this premium can be very significant. By varying the parameter n, we change the frequency at which DM updates information. Our next result shows that high frequency of information updates might inflict an extreme cost on DM; a particular splitting of a lottery drives down its value to the value of the worst prize in its support.

Proposition 3: *Consider disappointment aversion preferences with some $\phi : X \to \mathbb{R}$ and $\beta > 0$. For any $\varepsilon > 0$, and for any lottery $p = \sum_{j=1}^{m} p_j \delta_{x_j}$, there exists $T < \infty$ and a multi-stage lottery $Q^T \in \mathcal{P}\,(p)$ such that $V\big(Q^T\big) < \min\limits_{x_j \in \text{supp}(p)} \phi\,(x_j) + \varepsilon$.*

Let p be a binary lottery that yields 0 and 1 with equal probabilities. Consider n tosses of an unbiased coin. Define a series of random variables $\{z_i\}_{i=1}^n$ with $z_i = 1$ if the ith toss is "heads" and $z_i = 0$ if it is "tails". Let the terminal nodes of the n-stage lottery be:

$$\begin{array}{cl} 1 & \text{if } \sum_{i=1}^n z_i > \frac{n}{2} \\ 0.5\delta_1 + 0.5\delta_0 & \text{if } \sum_{i=1}^n z_i = \frac{n}{2} \\ 0 & \text{if } \sum_{i=1}^n z_i < \frac{n}{2} \end{array}$$

Note that the value of this n-stage lottery, calculated using recursive disappointment aversion preferences, is identical to the value calculated using recursive expected utility and probability $\frac{0.5}{1+\beta 0.5} < 0.5$ for "heads" in each period. Applying the weak law of large numbers,

$$\Pr\left(\sum_{i=1}^n z_i < \frac{n}{2}\right) \to 1$$

and therefore, for n large enough, the value approaches $\phi(0)$. We use a similar construction to establish that this result holds true for any lottery.

Ignoring the dynamic aspect of risk aversion might be misleading. We have already argued that a substantial fraction of many insurance premiums we observe in daily life can be attributed to the gradual resolution premium. Proposition 3 proves that this effect is quantitatively important, if the parameter n is sufficiently large.

5. PORU, "loss aversion with narrow framing" and the final-wealth hypothesis

Loss aversion with narrow framing (also known as "myopic loss aversion") is a combination of two motives: loss aversion (Kahneman and Tversky [1979]), that is, people's tendency to be more sensitive to losses than to gains, and a dynamic aggregation rule, narrow framing (Kahneman and Tversky [1984]), that argues that when making a series of choices, individuals "bracket" them by making each choice in isolation. When applied to behavior in financial markets, narrow framing means that individuals tend to evaluate long-term investments according to their short-term returns. Benartzi and Thaler [1995] were the first to use this approach and suggest explanations for several economic “anomalies”, such as the equity premium puzzle (Mehra and Prescott [1985]). Barberis and Huang [2005] and Barberis, Huang and Thaler [2006] generalize Benartzi and Thaler's work by assuming that DM derives utility directly from the outcome of a gamble over and above its contribution to total wealth.

The model presented in this paper can be used to address the same phenomena addressed with the loss aversion with narrow framing approach. Both models assume time neutrality. The combination of a specific form of non-smooth atemporal preferences and the folding-back procedure accounts for PORU. In an intertemporal context, these two features are analogous to loss aversion and narrow framing, respectively. The gradual resolution premium is the cost an individual incurs from frequently evaluating the outcomes of a dynamic random process.

The loss aversion with narrow framing approach challenges the hypothesis that only final wealth matters. Rabin [2000] and Safra and Segal [2006] give a parallel critique on a broad class of smooth models of decision making under risk. These authors use calibration results to argue that modest risk aversion over small stakes gambles necessarily implies absurd levels of risk aversion over large stakes gambles. Both Safra and Segal [2006] and Barberis Huang and Thaler [2006] argue that if DM faces some background risk, then a similar problem persists even if preferences are non-differentiable (i.e. if preferences display first-order risk aversion[17]); merging new gambles with preexisting ones eliminates the effect of first-order risk aversion.

Our model is consistent with risk aversion over small stakes gambles and only moderate risk aversion over large stakes gambles even if individuals face background risks. If most risks resolve gradually, then they cannot be compounded into a single lottery. Our model then implies first order risk aversion over each realized gamble. In other words, the mere existence of other risks is not enough to apply Rabin-type critique. Such an argument is only compelling if DM compounds risks that are resolved over a long period.

The conceptual difference between the two approaches is twofold. First, loss aversion with narrow framing brings to the forefront the idea that individuals evaluate any new gamble separately from its cumulative contribution to total wealth. Both the reference points relative to which gains and losses are computed and the way they dynamically adjust are usually set exogenously.[18] We, on the other hand, maintain the assumption that terminal wealth matters, and identify narrow framing as a preference parameter. The similarity between "disappointment aversion" and "loss aversion" has already been pointed out in Gul [1991] and stimulates further comparisons between these two notions. The novel insight provided by proposition 3 is that the (temporal) effect of narrow framing can be achieved even without giving up the assumption that utility depends on overall wealth, and that this effect is quantitatively important. Second, we set aside the question of why individuals are sensitive

[17] First order risk aversion means that the premium a risk averse DM is willing to pay to avoid an actuarially fair random variable $t\tilde{\epsilon}$ is proportional, for small t, to t. It implies "kinked" indifference curves along the main diagonal in a states-of-the-world representation (Segal and Spivak [1990]).

[18] Köszegi and Rabin [2006] offer a model in which the reference point is determined endogenously.

to the way uncertainty is resolved (i.e. why they narrow frame),[19] and construct a model that reveals the (context independent) behavioral implications of such considerations.

6. PORU and the value of information

We now reconsider the case of two-stage lotteries ($n = 2$). Let us suppose that just before the second-stage lottery is played, but after the realization of the first-stage lottery, DM can take, in the face of the remaining uncertainty, some action that might affect his ultimate payoff. The primitive in such a model is a preference relation over information systems (as we formally define below), which is induced from preferences over compound lotteries. Assume throughout this section that preferences over compound lotteries satisfy **A1** and **A2**. An immediate consequence of Blackwell's [1953] seminal result is that in the standard expected utility class, DM always prefers to have perfect information before making the decision, which allows him to choose the optimal action corresponding to the resulting state. Schlee [1990] shows that if $\succeq_1$ is of the rank-dependent utility class (Quiggin [1982]), then the value of perfect information will always be non-negative. This value is computed relative to the value of having no information at all, and therefore Schlee's result is salient about the comparison between getting complete and partial information. Safra and Sulganik [1995] left open the question of whether there are preference relations, other than expected utility, for which perfect information is always the most valuable. We show below that such preferences are fully characterized by PORU. Combining this result with theorem 1 reveals its implication on betweenness-recursive preferences.

More formally, let $S = \{s_1, \ldots, s_N\}$ be some finite set of states. Each state $s \in S$ occurs with probability p_s. The outcome of a lottery will depend both on the resulting state and on an action DM has made. For this we let $A = \{a_1, ..., a_M\}$ be a finite set of actions. Let $u : A \times S \rightarrow R$ be a function that gives the outcome $u(a, s)$ if action $a \in A$ is taken and the realized state is $s \in S$. (This outcome corresponds to the final prize $x \in X$.)

The first-stage lottery can be thought of as a randomization over a set $J = \{j_1, ..., j_m\}$ of signals indexed by j (where signal j indicates that p^j was selected in the first-stage lottery). Let $\pi : S \times J \rightarrow [0, 1]$ be a function such that $\pi(s, j)$ is the probability of getting the signal $j \in J$ when the prevailing state is $s \in S$. We naturally require that for all $s \in S$, $\sum_{j \in J} \pi(s, j) = 1$ (so that when the prevailing state is s, there is some probability distribution on the signal DM might get). The function π is called an information structure.

[19] Barberis and Huang [2006] suggest two different underlying sources of narrow framing. The first is based on a non-consumption utility, such as regret, and the second relates narrow framing to the "accessibility" of the uncertainty people confront. As these authors mention, each such motive, if taken literally, predicts different duration of narrow framing.

It automatically induces a splitting of the lottery into two stages, where with probability $\alpha_j(\pi) = \sum_{s' \in S} \pi(s', j) p_{s'}$, p^j is the second-stage lottery.

A full information structure, I, is a function such that for all $s \in S$ there exists $j(s) \in J$ with $\Pr(s|j(s)) = \frac{\pi(s,j(s))p_s}{\sum_{s' \in S} \pi(s',j(s))p_{s'}} = 1$, and for all $j \neq j(s)$ one has $\Pr(s|j) = 0$. In other words, in the sum above defining α_j, there is only one summand. The null information structure, ϕ, is a function such that $\Pr(s|j) = \Pr(s)$ for all $s \in S$ and $j \in J$.

Define $a^*(s)$ as the optimal action if you know that the prevailing state is s, that is, $u(a^*(s), s) := \max_{a \in A} u(a, s)$. Let $V^p(I)$ be the value of the lottery that assigns probability p_s to the outcome $u(a^*(s), s)$. After a signal j has been given, DM chooses the best a under the circumstances, namely a that maximizes the value of the lottery that assigns probability p_s^j to gain the outcome $u(a, s)$. We let $V(p^{j*})$ stand for the value of the j^{th} lottery maximized over the choice of an action $a \in A$. Finally, let $V^p(\pi)$ be the value of the lottery where the action is taken after receiving signal j, that is, the compound lottery assigning probability $\alpha_j(\pi) = \sum_{s' \in S} \pi(s', j) p_{s'}$ to p^{j*}.

Definition: $\succeq$ displays *preferences for perfect information* if for any information structure π and for any payoff function u, $V^p(I) \geq V^p(\pi)$.

Proposition 4: *If* $\succeq$ *satisfies* **A1** *and* **A2***, then the two statements below are equivalent:*

(i) $\succeq$ *displays PORU*

(ii) $\succeq$ *displays preferences for perfect information.*

Analogously, PGRU holds if and only if for any information structure π and for any payoff function u, $V^p(\pi) \geq V^p(\phi)$

Showing that (i) is necessary for (ii) is immediate. For the other direction, we note that two forces reinforce each other: First, getting full information means that the underlying lottery is of the "one-shot resolution" type, since uncertainty is completely resolved by observing the signal. Second, better information enables better planning; using it, a decision maker with monotonic preferences is sure to take the optimal action in any state. The proof distinguishes between the two prime motives for getting full information: The former, which is captured by PORU, is intrinsic, whereas the latter, which is reflected via the monotonicity of preferences with respect to outcomes, is instrumental. The result for PGRU is similarly proven. The null information structure is of the "one-shot resolution" type and it has no instrumental value.

Corollary: *If $\succeq$ satisfies* **A1** *and* **A2***, then $\succeq$ displays preferences for perfect information* whenever $\succeq_1$ *satisfies NCI.*

Proposition 4 is independent of **A3**. By adding **A3** as a premise we get:

Corollary: *For any betweenness-recursive preferences, $\succeq$ displays preferences for perfect information* iff $\succeq_1$ *satisfies NCI.*

7. Application to investment under uncertainty

The concept of option value was initially demonstrated by Arrow and Fischer [1974], and later recognized in the works of McDonald and Siegel [1986], Pindyck [1991], and Dixit and Pindyck [1994]. These authors point out that if an investor has a choice over when to implement an (irreversible) investment decision, then investing according to the net present value (NPV) of a project is not adequate; waiting leaves room for new information that DM might use to make better decisions. In other words, the availability of future signals always favors delaying the investment.

This result rests on the assumption that decision makers are expected utility maximizers, and it ceases to hold once we relax that assumption. In particular, PORU suggest another effect that should be taken into account: The harmful effect of the gradual resolution of uncertainty, induced by an informative signal, can outweigh the benefit of getting more information from this same signal.

To illustrate, consider the following three-period investment problem.[20] In the first period, an investor decides whether or not to invest in a certain machine. The investment requires an immediate cost of C dollars. If he chooses to invest (option A), he will be able to produce in both the second and the third period. In the second period, the demand is certain and the investor is sure to receive variable profits $\pi^2 > 0$. In the third period, the demand is uncertain and the profits are determined by the realization of a finite random variable π^3. Denote by p the probability distribution over the third-period profits π_i^3, $i = 1, .., n$. If the investor decides not to invest in the first period (option B), he may still invest in the second period. In that case he waives π^2, but he is able, before making the investment decision, to learn the realization of a signal j that is correlated with π^3 and comes from a finite set $J = \{j_1, ..., j_m\}$. Thus, if he invests in the second period after receiving the signal j, his

[20]Based on an example given in Gollier [2001]

third-period's profits are distributed according to the conditional distribution p^j. Let α_j be the unconditional probability of getting the signal $j \in J$. The discount rate between any two successive periods is r.

Assume that DM has disappointment aversion preferences with linear ϕ and positive β. The value of option A, V^A, is given by:

$$V^A = -C + \frac{1}{1+r}\pi^2 + \frac{1}{(1+r)\,2}V(p)$$

Let $z_j := \max\left\{0, -C + \frac{1}{1+r}V(p^j)\right\}$. The value of option B is denoted V^B and is the unique value v that solves:

$$v = \frac{1}{1+r}\frac{\sum_{j:z_j>v} z_j\alpha_j + (1+\beta)\sum_{j:z_j\leq v} z_j\alpha_j}{1+\beta\sum_{j:z_j\leq v}\alpha_j}$$

Let $Q = \langle \alpha_j, p^j \rangle_{j=1}^m$ be the compound lottery such that for each i, $p_i = \sum_j \alpha_j p_i^j$. Denote its value by $V(Q)$. Further let $\Delta V := V^A - V^B$. The investor chooses option A if and only if:

$$\Delta V = \frac{1}{(1+r)^2}\left[\begin{array}{c} \underbrace{(1+r)\left(\pi^2 - rC\right)}_{\text{net present value}} \\ \\ \underbrace{-(1+\beta)\frac{\sum_{j:V(p^j)\leqslant\min[V(Q),C(1+r)]}\alpha_j\left(C - \frac{V\left(p^j\right)}{(1+r)}\right)}{1+\beta\sum_{j:V(p^j)\leqslant V(Q)}\alpha_j}}_{\text{option value}} \\ \\ +\quad\underbrace{(V(p) - V(Q))}_{\text{gradual resolution premium}} \end{array}\right] \geq 0$$

The first component is the regular NPV rule: Invest today if the forgone second-period profits are larger than the interest gained due to delaying the investment. This would be the decision criterion in the absence of a signal for the case $\beta = 0$, when DM is risk neutral and simply maximizes the NPV of the investment.

The second component is the flexibility value, or the option value, of delaying the investment. It reflects the idea that occupied with more information, DM can refrain from investing if he learns that the demand is likely to be too low. This term is positive and is an increasing function of β.

Since $Q \in \mathcal{P}(p)$, the last term is the gradual resolution premium, $\text{grp}(Q)$: investing today saves DM the need to be aware of the gradual resolution of uncertainty. This term is non-negative for $\beta \geqslant 0$.

As we mentioned above, if a standard expected utility maximizer prefers to invest in period two, even under the null-information system, then he clearly does so when the information system is finer. However, for strictly positive values of β this is not necessarily true. For example, suppose $\beta = 1$, $C = 50$, $r = 0.1$, $\pi^2 = 5$ and $\pi_i^3 \in \{0, 1000, 2000\}$. Let $p = \frac{1}{3}\delta_0 + \frac{1}{3}\delta_{1000} + \frac{1}{3}\delta_{2000}$, and $Q = \left\langle \delta_0, 1/3; \frac{1}{2}\delta_{1000} + \frac{1}{2}\delta_{2000}, 2/3 \right\rangle \in \mathcal{P}(p)$. Since $\pi^2 = rC$, in the absence of a signal, DM is simply indifferent between investing in period 1 and investing in period 2. The signal is useful in that if the investor learns that the quantity demanded is zero, he will choose not to invest. Nevertheless, this option value is not sufficient to compensate him for the compound lottery he must face in case he chooses to view a signal, and he strictly prefers option A. Therefore, the NPV rule should be twice adjusted, taking into account both the second and the third effects.

In the general case, it is not obvious which effect, the option value or PORU, dominates. Similarly to the assertion in proposition 2, there exists a finite value β^* in which the gradual resolution premium is maximized.[21] This observation implies the following:
(i) There exists $\overline{\beta}$, such that for all $\beta > \overline{\beta}$ the option value is dominant.
(ii) If the option value is dominant at β^*, so it is dominant for all $\beta > \beta^*$.
(iii) There exists $\underline{\beta}$, such that the option value is dominant for all $\beta \in \left(0, \underline{\beta}\right]$.

The setting above can be used to distinguish between decision makers with PORU and decision makers who have preferences for early or late resolution of uncertainty. The availability of an informative signal would induce decision makers with preferences for early resolution of uncertainty to choose option B. Independently of its instrumental value, a signal leads to an earlier, yet not complete, resolution of uncertainty. Therefore, the only possible confusion would be between the behavior of individuals with PORU and individuals who prefer late resolution of uncertainty. This confusion can be avoided by altering the resolution process in option A. Suppose that under that option, the uncertainty about period 3 returns would already have been resolved in period 2. Due to the time neutrality assumption, such a change has no effect on individuals with PORU. Individuals with preferences for late resolution of uncertainty, on the other hand, would be worse off under this alternative.

8. Conclusion

Searching for a better understanding of decision-making under risk, and disentangling decision makers' attitude towards risk and time have been two active fields of research in

[21] Since $\lim_{\beta \to \infty} (V(p) - V(Q)) = 0$, there exists $\widehat{\beta} := \max \left\{ \beta \left| (V(p) - V(Q)) \right|_\beta = \frac{\epsilon}{2} \right\}$ and $\widehat{\beta} < \infty$. Thus $(V(p) - V(Q))$ is a continuous function on the compact interval $\left[0, \widehat{\beta}\right]$, and hence achieves its maximum on this domain.

economics. This paper contributes to both fields. We study preferences over multi-stage lotteries and explicitly assume that the way uncertainty is resolved over time matters. Being exposed to the resolution process bears the risk of perceiving intermediate outcomes as disappointing or elating. Individuals who are more sensitive to disappointment suffer from getting partial information and, therefore, strictly prefer *ex-ante* all uncertainty to be resolved in a single point in time. Behaviorally, these individuals will display higher risk aversion if uncertainty is resolved gradually. We formally define such dynamic *preferences for one-shot resolution of uncertainty* (PORU), and show that they can be modeled using a single, static preference relation. Our main result states that to characterize PORU, one needs to impose on these static preferences a property, *negative certainty independence* (NCI), which is identified with Allais-type behavior, the most compelling argument against the independence axiom. In other words, being prone to Allais-type behavior and being averse to the gradual resolution of uncertainty are synonymous. This equivalence provides clear predictions for dynamic preferences, and calls for further experimental testing to be done. Our model also predicts a specific attitude towards information. Although we accommodate situations where people avoid information that is instrumental to their decision making, perfect information will never be rejected, and will always be preferred to any other information system.

The frequency with which an individual evaluates lotteries over time is a preference parameter in our model, and its effect is measured by the gradual resolution premium. The more often an individual updates information, the more sensitive he is to gradual resolution. We show that this effect can be quantitatively important, implying extreme degrees of risk aversion.

If most actual risks that individuals face are resolved gradually over time, then these risks cannot be compounded into a single lottery and, therefore, the gradual resolution premium should not be disregarded. Rabin and Thaler [2001] state that "...it is clear that loss aversion and the tendency to isolate each risky choice must both be key components of a good descriptive theory of risk attitudes." Our model shows that in an intertemporal context, both features, and especially the isolation component, can be addressed independently of studying framing effects.

9. Appendix

9.1. Extension to n-stage lotteries, a formal description

The following is a formal description of any compound lottery, or a probability tree. Let T be a finite set of (chance) nodes. Let $\rhd_p$, "predecessor of", be a partial order on T with $x \rhd_p y$ if x precedes y. For any node $t \in T$, let $PRE(t) = \{x : x \rhd_p t\}$ be the set of predecessors of t. For any $t, t' \in T$, we say that t is an immediate predecessor of t', and denote it by $t \rhd_{ip} t'$, if $t \in PRE(t')$ and $\nexists t'' \in PRE(t')$ such that $t \in PRE\left(t''\right)$. An initial node is any $t \in T$ with $PRE(t) = \varnothing$. A pair $(T, \ \rhd_p)$ is a tree if it has a single initial node, and if for all $t \in T$, $PRE(t)$ is totally ordered by $\rhd_p$ (so that each node t has no more than one immediate predecessor).

We say that T is of length n if each complete path in T is of length n. Denote by T^k the set of stage k's nodes. We have $\bigcup_{k=1}^{n+1} T^k = T$. A node s is an immediate successor of t iff t is an immediate predecessor of s, that is , $s \rhd_{is} t \Longleftrightarrow t \rhd_{ip} s$. Let $F(t) = \{x : x \rhd_{is} t\}$. Let $(\rho_t)_{t\in T}$ be a collection of probability distributions, one for each node, over $F(t)$. If $F(z) = \varnothing$, we say that z is a terminal node. Denote by T^{n+1} the set of all terminal nodes. We identify T^{n+1} as the set of ultimate prizes. For any $k \in \{1, 2, ..., n\}$, we identify $t \in T^k$ as a compound lottery, starting at time k, of length $n + 1 - k$. In order to agree with other notations in the text, we write any such lottery as $Q^{n+1-k}(t)$. Finally, let Γ^l be the set of lotteries of the following form: For all $j \neq l$, every $t \in T^j$ is a trivial node (i.e. $|F(t)| = 1$). In time l, a certain one-stage lottery is acted out.

Let $\succeq_n$ be a complete and transitive binary relation over $\mathcal{L}^n$, on which we impose the following axioms:

For any $l \in \{1, 2, .., n\}$, let Γ_q^l be the member of Γ^l with the single-stage lottery being q.

A1$'$: $\forall q \in \mathcal{L}^1$ *and for all* $l, l' \in \{1, 2, .., n\}$, $c \sim_n \Gamma_q^{l'}$.

A2$'$ *: Fix $t^* \in T^n$. Suppose that for all $t \in T / \{t^*\}$, $F(t)$ is the same in both Q^n and $Q^{n'}$. If Q^n yields the lottery q in t^* and $Q^{n'}$ yields the lottery q'in t^*, then $Q^n \succeq_n Q^{n'} \Longleftrightarrow \Gamma_q^n \succeq_n \Gamma_{q'}^n$.*

The implied value of any compound lottery is the following: For any $t \in T^n$, define $W^1(Q^1(t)) = V(Q^1(t))$, and recursively for $k = n-1, n-2..., 1$ and for all $t \in T^k$, let

$$W^{n+1-k}\left(Q^{n+1-k}(t)\right) = V\left(\left\langle \rho_t(s), \mathrm{c}\left(Q^{n+1-(k+1)}(s)\right)\right\rangle_{s\in F(t)}\right)$$

where $\mathrm{c}\big(Q^l(s)\big) \in X$ is the certaity equivalent of $\mathrm{c}\big(Q^l(s)\big)$

Lastly, and using the representation above, we extend the definition of PORU to this richer domain. Let Q^n, $Q^{n'} \in \mathcal{L}^n$ be two compound lotteries that are equal except in one sub-lottery of length $n + 1 - k$, $k \in \{2, 3, ..., n-1\}$ that originates from some $t^* \in T^k$. Formally, for all $t \in T$ such that $t^* \notin PRE(t)$, $F(t)$ is the same in both Q^n and $Q^{n'}$. Denote the associate (different) sub-lotteries by $Q_{Q^n}^{n+1-k}(t^*)$ and $Q_{Q^{n'}}^{n+1-k}(t^*)$, respectively. Let p be the lottery that for all $s \in F(t^*)$ gives the prize $c_V\left(Q^{n+1-(k+1)}(s)\right)$ with probability $\rho_{t^*}(s)$. Define the set $\mathcal{P}(p)$ just as before.

Definition: $\succeq_n$ display PORU if for all Q^n, $Q^{n'} \in \mathcal{L}^n$ and $p \in \mathcal{L}^1$ as described above, $W^{n+1-k}\left(Q^{n+1-k}_{Q^n}(t^*)\right) = V(p)$ and $W^{n+1-k}\left(Q^{n+1-k}_{Q^{n'}}(t^*)\right) = W(Q^2)$ for some $Q^2 \in \mathcal{P}(p)$ imply $Q^n \succeq_n Q^{n'}$.

Theorem 1': *under* $\mathbf{A1}', \mathbf{A2}', \mathbf{A3}$, *theorem 1 remains intact.*

For brevity, we omit the detailed proof. It simply involved a repeated use of $\mathbf{A1}', \mathbf{A2}'$, and **A3** to transform the problem into the framework of theorem 1.

9.2. Proofs

Proof of theorem 1

Define $f_p(v) = \sum_{x\in S(p)} [u(x,v) - v] p_x$. Thus $V(p)$ is the unique solution to $f_p(v) = 0$. Note that whenever $p = \sum_i \alpha_i p^i$ we have $f_p(v) = \sum_i \alpha_i f_{p^i}(v)$. Since $f_p(v) = 0$ has a unique solution and for all $x \in (w,b)$, $u(x, V(\delta_w)) > u(w, V(\delta_w))$ and $u(b, V(\delta_b)) > u(x, V(\delta_b))$, showing that $V(p) \geqslant V(Q)\ \forall Q \in \mathcal{P}(p)$ is equivalent to showing that $f_p(V(Q)) \geqslant 0\ \forall Q \in \mathcal{P}(p)$. To show the latter, we subtract from it $0 = \sum_i \alpha_i f_{p^i}(V(p^i))$, which does not change the expression, and regroup the terms as follows:

$$
\begin{aligned}
f_p(V(Q)) &= \sum_i \alpha_i f_{p^i}(V(Q)) \\
&= \sum_i \alpha_i \left[f_{p^i}(V(Q)) - f_{p^i}\left(V\left(p^i\right)\right)\right] \\
&= \sum_i \alpha_i \sum_{x\in S(p^i)} \left[(u(x,V(Q)) - V(Q)) - \left(u\left(x, V\left(p^i\right)\right) - V\left(p^i\right)\right)\right] p^i_x \\
&= \sum_i \alpha_i \sum_{x\in S(p^i)} \left[u(x,V(Q)) - u\left(x,V\left(p^i\right)\right)\right] p^i_x + \sum_i \alpha_i V\left(p^i\right) - V(Q) \\
&= \sum_i \alpha_i \sum_{x\in S(p^i)} u(x,V(Q))\, p^i_x - \sum_i \alpha_i \sum_{x\in S(p^i)} u\left(x,V\left(p^i\right)\right) p^i_x \\
&\quad + \sum_i \alpha_i V\left(p^i\right) - V(Q) \\
&= \sum_i \alpha_i \sum_{x\in S(p^i)} u(x,V(Q))\, p^i_x - \sum_i \alpha_i V\left(p^i\right) + \sum_i \alpha_i V\left(p^i\right) - V(Q) \\
&= \sum_i \alpha_i \left[\sum_{x\in S(p^i)} u(x,V(Q))\, p^i_x - u\left(c\left(p^i\right), V(Q)\right)\right]
\end{aligned}
$$

Claim 1:

$$\sum_i \alpha_i \left[\sum\nolimits_{x\in S(p^i)} u\left(x, V(Q)\right) p_x^i - u\left(\mathrm{c}\left(p^i\right), V(Q)\right)\right] \geqslant 0 \ \forall p \text{ and } \forall Q \in \mathcal{P}\left(p\right)$$

iff

$$\forall i, \ \left[\sum\nolimits_{x\in S(p^i)} u\left(x, V(Q)\right) p_x^i - u\left(\mathrm{c}\left(p^i\right), V(Q)\right)\right] \geqslant 0 \ \forall p \text{ and } \forall Q \in \mathcal{P}\left(p\right).$$

Proof: The "if" part is obvious. For the "only if" part, assume that for some j and for some $v \neq V(p^j)$, $u\left(\mathrm{c}\left(p^j\right), v\right) - \sum_{x\in S(p^j)} u\left(x, v\right) p_x^j > 0$. Pick $y \in X$ and $\alpha \in (0,1)$ such that $V\left\langle 1, \left(\alpha\delta_y + (1-\alpha)\,\delta_{\mathrm{c}(p^j)}\right)\right\rangle = v$ (by betweenness and continuity, such y and α exist.) Let $Q = \left\langle \alpha, \delta_y; (1-\alpha), p^j \right\rangle$ (hence $V(Q) = v$). Finally , let $p := \alpha\delta_y + (1-\alpha)\,p^j$. Note that $Q \in \mathcal{P}\left(p\right)$. By construction we have

$$f_p\left(v\right) = (1-\alpha)\left[\sum\nolimits_{x\in S(p^j)} u\left(x, v\right) p_x^j - u\left(\mathrm{c}\left(p^j\right), v\right)\right] < 0$$

so $V(p) < V(Q)$.‖

Since p was arbitrary, we get the following necessary and sufficient condition for PORU:

$$\mathbf{C}_1: \ \left[\sum\nolimits_{x\in S(p)} u\left(x, v\right) p_x - u\left(\mathrm{c}\left(p\right), v\right)\right] \geqslant 0 \ \forall p \text{ and } \forall v \in V\left(\mathcal{L}^1\right).$$

Claim 2: $\mathbf{C}_1$ iff for every triple $x_3 > x_2 > x_1$, the indifference curve through δ_{x_2} is the steepest.

Proof: (only if): Fix $x_3 > x_2 > x_1$. By continuity, for every such triple there exists a $p \in (0,1)$ such that $p\delta_{x_3} + (1-p)\delta_{x_1} \sim_1 \delta_{x_2}$. Therefore, the vertex $(0,0)$ that represents the lottery δ_{x_2} and the point $(1-p, p)$ lie on the same indifference curve. This indifference set is of the original preferences, and hence the value attached to it is $V(p\delta_{x_3} + (1-p)\delta_{x_1}) := V(p) = pu\left(x_3, V(p)\right) + (1-p)\,u\left(x_1, V(p)\right) = u\left(x_2, V(p)\right)$. By $\mathbf{C}_1$, for any other v, if we pass through $(1-p, p)$ the (artificial) indifference curve corresponding to the value v, it must lie weakly above the curve from the same collection that passes through $(0,0)$. Since the betweenness property implies that indifference curves are straight lines (so their slopes are constant), the result follows.

(if): Take a lottery p with $\left|S\left(p\right)\right| = n-1$ that belongs to an indifference set $I_v := \{p' : \sum_x u\left(x, v\right) p'_x = v\}$ in a $(n-1)$-dimensional unit simplex $\Delta\left(n\right)$. Assume further that for some $x_v \in (w, b)$ with $x_v \notin S(p)$, $\langle 1, \delta_{x_v} \rangle \in I_v$[22]. By monotonicity and continuity,[23] p can be written as a convex combination $\alpha r + (1-\alpha) w$, for some $\alpha \in (0,1)$ and $r, w \in I_v$ with $|S(r)| = |S(w)| = n-2$. By the same argument, both r and w can be written, respectively, as a convex combination of two other lotteries with size of support equal $n-3$ and that belong to I_v. Continue in the same fashion to get an index set J and a collection of lotteries, $\{q^j\}_{j \in J}$, such that for all $j \in J$, $|S(q^j)| = 2$ and $q^j \in I_v$. Note that by monotonicity, if $y, z \in S(q^j)$ then either $z > x_v > y$ or $y > x_v > z$. By construction, for some $\alpha_1, ..., \alpha_J$ with $\alpha_j > 0$ and $\sum_j \alpha_j = 1$, $\sum_j \alpha_j q^j = p$. Let $V(q, v) := \sum_x q_x u(x, v)$. By hypothesis, $V(q^j, v') \geqslant u(x_v, v')$ for all $j \in J$ and for all $v' \in V(\Delta(n))$ and therefore also

$$\begin{aligned} V(p, v') &= \sum_j \alpha_j V\left(q^j, v'\right) = \sum_x \sum_j \alpha_j q_x^j u(x, v') \\ &\geqslant \sum_j \alpha_j u(x_v, v') = u(x_v, v') = u(\mathrm{c}(p), v'). \| \end{aligned}$$

Claim 3: NCI and $\mathbf{C}_1$ are equivalent.

Proof:

$\mathbf{C}_1 \rightarrow$NCI: Assume $p \succeq_1 \delta_x$. Using the observation that for any two lotteries p and q, $V(p) \geqslant V(q)$ is equivalent to $f_p(V(q)) \geqslant 0$, we have $\sum_i p_{x_i} u(x_i, V(p) \geqslant u(x, V(p)$. By $\mathbf{C}_1$ and monotonicity, $\sum_i p_{x_i} u(x_i, v) \geqslant u(x, v)$ for all v, and in particular for $v = V(\lambda p + (1-\lambda) q)$.[24] Calculating the expected utility of the two lotteries $\lambda p + (1-\lambda) q$ and $\lambda \delta_x + (1-\lambda) q$ relative to the value $V(\lambda p + (1-\lambda) q)$ and using again the observation above, establishes the result.

NCI$\rightarrow \mathbf{C}_1$: Suppose not. Then there exists a lottery $p \sim$c(p) with $\left[\sum_{x \in S(p)} u(x, v) p_x - u(c(p), v)\right] < 0$ for some v. Pick $y \in X$ and $\alpha \in (0,1)$ such that $V(\alpha p + (1-\alpha)\delta_y) = v$. We have $v < \alpha u(c(p), v) + (1-\alpha) u(y, v) = V\left(\alpha \delta_{c(p)} + (1-\alpha)\delta_y, v\right)$, or $\alpha \delta_{c(p)} + (1-\alpha)\delta_y \succ_1 \alpha p + (1-\alpha)\delta_y$, contradicting NCI.‖

Note that by reversing the inequality in $\mathbf{C}_1$ and the weakly-prefer sign in NCI, we derive the analogous conditions for PGRU.■

[22] The analysis would be the same, though with messier notations, even if $|S(p)| = n$, i.e., if $x \in S(p)$.

[23] These two assumptions guarantee that no indifference set terminates in the relative interior of any $k \leqslant n-1$ dimensional unit simplex.

[24] If $p \sim \delta_x$, the assertion is evident. Otherwise, we need to find p^* that is both first order stochastically dominated by p and satisfies $p^* \sim \delta_x$, and use the monotonicity of $u(, v)$ with respect to its first argument. By continuity such p^* exists.

Proof of theorem 2

Since for expected utility preferences NCI is always satisfied, it is enough to demonstrate the result for lotteries with at most 3 prizes in their support.

For $x \in [w, b]$, denote by $V(\delta_x)$ the unique solution of $v = u(x, v)$. Without loss of generality, set $u(w, v) = 0$ and $u(b, v) = 1$ for all $v \in [0, 1]$. Fix $\overline{v} \in (0, 1)$. By monotonicity and continuity there exists $x(\overline{v}) \in (w, b)$ such that $\overline{v} = V\left(\delta_{x(\overline{v})}\right)$. Take any $x > x(\overline{v})$ and note that $\mu(V|x, x(\overline{v}), w) = \left[\frac{u(x(\overline{v}),v)}{u(x,v)-u(x(\overline{v}),v)}\right]$, the slope of the indifference curves on the space $\{(p_w, p_x) \mid p_w, p_x \geq 0,\ p_w + p_x \leq 1\}$, is continuous and differentiable as a function of v on $[0, V(\delta_x)]$.

Since $\overline{v} \in (0, V(\delta_x))$, theorem 1 implies that $\mu(V|x, x(\overline{v}), w)$ is maximized at $v = \overline{v}$. A necessary condition is:

$$\frac{\partial}{\partial v}\left[\frac{u(x(\overline{v}), \overline{v})}{u(x, \overline{v}) - u(x(\overline{v}), \overline{v})}\right] = 0$$

Or,[25] using $\overline{v} = u(x(\overline{v}), \overline{v})$ and denote by u_i the partial derivative of u with respect to its i^{th} argument,

$$u_2(x(\overline{v}), \overline{v})[u(x, \overline{v}) - \overline{v}] = [u_2(x, \overline{v}) - u_2(x(\overline{v}), \overline{v})]\overline{v} \tag{1}$$

Note that by continuity and monotonicity of $u(x, v)$ in its first argument, for all $x \in (x(\overline{v}), b)$ there exists $p \in (0, 1)$ such that $p\delta_w + (1-p)\delta_x \sim_1 \delta_{x(\overline{v})}$, or $u(x, \overline{v})(1-p) = u(x(\overline{v}), \overline{v}) = \overline{v}$. Therefore, and using again theorem 1, (1) is an identity for $x \in (x(\overline{v}), b)$, so we can take the partial derivative of both sides with respect to x and maintain equality. We get:

$$u_2(x(\overline{v}), \overline{v})\, u_1(x, \overline{v}) = u_{21}(x, \overline{v})\overline{v}$$

Since u is strictly increasing in its first argument, $u_1(x, \overline{v}) > 0$ and $\overline{v} > 0$. Thus: $\frac{u_{21}(x,\overline{v})}{u_1(x,\overline{v})} = \frac{u_2(x(\overline{v}),\overline{v})}{\overline{v}} = l(\overline{v})$ independent of x, or by changing order of differentiation: $\frac{\partial}{\partial v}[\ln u_1(x, \overline{v})]$ is independent of x.

Since $\overline{v}$ was arbitrary, we have the following differential equation on $\{(x, v) \mid v < u(x, v)\}$:

$$\frac{\partial}{\partial v}[\ln u_1(x, v)] = l(v)$$

[25] second order conditions would be :

$$\frac{u_{22}(x(\overline{v}), \overline{v})}{u_{22}(x, \overline{v})} < \frac{\overline{v}}{u(x, \overline{v})}\ (< 1)$$

By the fundamental theorem of calculus, the solution of this equation is:

$$\begin{aligned}
\frac{\partial}{\partial v}\left[\ln u_1(x,v)\right] &= l(v) \\
&\implies \ln u_1(x,v) = \ln u_1(x,0) + \textstyle\int_{s=0}^{v} l(s)\,ds \\
&\implies u_1(x,v) = u_1(x,0)\exp\left(\textstyle\int_{s=0}^{v} l(s)\,ds\right) \\
&\implies u(x,v) - u(x(v),v) = \exp\left(\textstyle\int_{s=0}^{v} l(s)\,ds\right)\int_{x(v)}^{x} u_1(t,0)dt \\
&\implies u(x,v) - v = \exp\left(\textstyle\int_{s=0}^{v} l(s)\,ds\right)\left(u(x,0) - u(x(v),0)\right)
\end{aligned}$$

Note that the term

$$\exp\left(\textstyle\int_{s=0}^{v} l(s)\,ds\right) = \exp\left(\int_{s=0}^{v}\frac{u_2(x(s),s)}{s}ds\right)$$

is well defined since by the assumption that all derivatives are continuous and bounded and that $u_1 > 0$, we use L'Hopital's rule and implicit differentiation to show that the term

$$\begin{aligned}
\lim_{s\to 0}\frac{u_2(x(s),s)}{s} &= \lim_{s\to 0} u_{21}(x(s),s)\,x'(s) + u_{21}(x(s),s) \\
&= \lim_{s\to 0} u_{21}(x(s),s)\,\frac{1-u_2(x(s),s)}{u_1(x(s),s)} + u_{21}(x(s),s)
\end{aligned}$$

is finite and hence $\left(\int_{s=0}^{v}\frac{u_2(x(s),s)}{s}ds\right)$ is finite as well.

To uncover $u(x,v)$ on the region $\{(x,v) \mid v > u(x,v)\}$, fix again some $\overline{v} \in (0,1)$ and the corresponding $x(\overline{v}) \in (w,b)$ (with $\overline{v} = u(x(\overline{v}),\overline{v})$). Take any $x < x(\overline{v})$ and note that $\widehat{\mu}(V|b,x(\overline{v}),x) = \left[\frac{u(x(\overline{v}),v)-u(x,v)}{1-u(x(\overline{v}),v)}\right]$, the slope of the indifference curves on the space $\{(p_x,p_b) \mid p_x,p_b \geq 0,\ p_x+p_b \leq 1\}$, is continuous and differentiable as a function of v on $[V(\delta_x),b]$.

Since $\overline{v} \in (V(\delta_x),b)$, by using theorem 1 we have:

$$\frac{\partial}{\partial v}\left[\frac{u(x(\overline{v}),\overline{v}) - u(x,\overline{v})}{1-u(x(\overline{v}),\overline{v})}\right] = 0$$

or,

$$\left(u_2(x(\overline{v}),\overline{v}) - u_2(x,\overline{v})\right)[1-\overline{v}] = -u_2(x(\overline{v}),\overline{v})\left[\overline{v} - u(x,\overline{v})\right] \tag{2}$$

Using the same argumentation from the former case, (2) holds for all $x \in (w,x(\overline{v}))$, so we can take the partial derivative of both sides with respect to x and maintain equality. We get:

$$-u_{21}(x,\overline{v})[1-\overline{v}] = u_1(x,\overline{v})u_2(x(\overline{v}),\overline{v})$$

Since u is strictly increasing in its first argument, $u_1(x,\overline{v}) > 0$ and $1-\overline{v} > 0$. Thus: $\frac{u_{21}(x,\overline{v})}{u_1(x,\overline{v})} = -\frac{u_2(x(\overline{v}),\overline{v})}{[1-\overline{v}]} = k(\overline{v})$ independent of x, or by changing order of differentiation: $\frac{\partial}{\partial v}[\ln u_1(x,\overline{v})]$ is independent of x.

Since $\overline{v}$ was arbitrary, we have the following differential equation on $\{(x,v) \mid v > u(x,v)\}$:

$$\frac{\partial}{\partial v}[\ln u_1(x,v)] = k(v)$$

Its solution is given by

$$\begin{aligned}
\frac{\partial}{\partial v}[\ln u_1(x,v)] &= k(v) \\
&\implies \ln u_1(x,1) - \ln u_1(x,v) = \textstyle\int_{s=v}^{1} k(s)\,ds \\
&\implies \ln u_1(x,v) = \ln u_1(x,1) - \textstyle\int_{s=v}^{1} k(s)\,ds \\
&\implies u_1(x,v) = u_1(x,1)\exp\left(\textstyle\int_{s=v}^{1} k(s)\,ds\right)^{-1} \\
&\implies u(x,v) - u(x(v),v) = \exp\left(\textstyle\int_{s=v}^{1} k(s)\,ds\right)^{-1} \displaystyle\int_{x}^{x(v)} u_1(t,1)dt \\
&\implies u(x,v) - v = -[u(x(v),1) - u(x,1)]\exp\left(\textstyle\int_{s=v}^{1} k(s)\,ds\right)^{-1}
\end{aligned}$$

which is again well defined since

$$\exp\left(\textstyle\int_{s=v}^{1} k(s)\,ds\right) = \exp\left(\textstyle\int_{s=v}^{1} - \displaystyle\frac{u_2(x(s),s)}{[1-s]}ds\right)$$

and

$$\begin{aligned}
\lim_{s\to 1} - \frac{u_2(x(s),s)}{[1-s]} &= \lim_{s\to 1} u_{21}(x(s),s)\,x'(s) + u_{21}(x(s),s) \\
&= \lim_{s\to 1} u_{21}(x(s),s)\frac{1-u_2(x(s),s)}{u_1(x(s),s)} + u_{21}(x(s),s)
\end{aligned}$$

is finite, and hence the whole integral is finite.

So far we have:

$$u(x,v) - v = \begin{cases} [u(x,0) - u(x(v),0)]\exp\left(\int_{s=0}^{v} \frac{u_2(x(s),s)}{s}ds\right) & x > x(v) \\ -[u(x(v),1) - u(x,1)]\left(\exp\left(\int_{s=v}^{1} - \frac{u_2(x(s),s)}{[1-s]}ds\right)\right)^{-1} & x < x(v) \end{cases} \tag{3}$$

We add the following restrictions:

(i) $u(b,v)=1$ for all $v\in[0,1]$, which implies:

$$[1-u(x(v),0)]\exp\left(\int_{s=0}^{v}\frac{u_2(x(s),s)}{s}ds\right)=1-v$$

(ii) $u(w,v)=0$ for all $v\in[0,1]$, which implies:

$$u(x(v),1)\left(\exp\left(\int_{s=v}^{1}-\frac{u_2(x(s),s)}{[1-s]}ds\right)\right)^{-1}=v$$

Substituting into (3) to get:

$$u(x,v)-v=\begin{cases}[u(x,0)-u(x(v),0)]\frac{1-v}{[1-u(x(v),0)]} & x>x(v)\\ -[u(x(v),1)-u(x,1)]\frac{v}{u(x(v),1)} & x<x(v)\end{cases} \tag{4}$$

We further require:

(iii) Continuity at $x=x(v)$. This is immediate since

$$\lim_{x\to_{-}x(v)}(u(x,v)-v)=\lim_{x\to_{+}x(v)}(u(x,v)-v)=0$$

(iv) Differentiability at $x(v)$ for all v:

$$u_1(x(v),0)\frac{1-v}{[1-u(x(v),0)]}=u_1(x(v),1)\frac{v}{u(x(v),1)}$$

or

$$\frac{u_1(x(v),1)}{u_1(x(v),0)}=\frac{[1-u(x(v),v)]}{[1-u(x(v),0)]}\frac{u(x(v),1)}{u(x(v),v)} \tag{5}$$

Let r$(x,v):=\frac{-u_{11}(x,v)}{u_1(x,v)}$. Given $v\in(0,1)$, note that

$$r(x,v)=\begin{cases}-\frac{u_{11}(x,0)}{u_1(x,0)} & x>x(v)\\ -\frac{u_{11}(x,1)}{u_1(x,1)} & x<x(v)\end{cases}$$

But since u is continuous and $r(x,v)$ is well defined, $r(x,v)$ must be continuous as well. Therefore, we require:

$$-\frac{u_{11}(x(v),0)}{u_1(x(v),0)}=-\frac{u_{11}(x(v),1)}{u_1(x(v),1)}$$

and since this is true for any v and the function $x(v)$ is onto, we have for all $x\in(w,b)$:

$$-\frac{u_{11}(x,0)}{u_1(x,0)}=-\frac{u_{11}(x,1)}{u_1(x,1)}$$

which implies that for some a and b, $u(x,1) = au(x,0) + b$. But $u(0,1) = u(0,0) = 0$ and $u(1,1) = u(1,0) = 1$, hence, by continuity, $b = 0$ and $a = 1$, or $u(x,1) = u(x,0) := z(x)$ for all $x \in [w,b]$. Plug into (4) to get:

$$u(x,v) - v = \begin{cases} [z(x) - z(x(v))] \frac{1-v}{[1-z(x(v))]} & x > x(v) \\ -[z(x(v)) - z(x)] \frac{v}{z(x(v))} & x < x(v) \end{cases} \tag{6}$$

and into (5) to get:

$$\frac{u_1(z(x))}{u_1(z(x))} = 1 = \frac{[1-v]}{[1-z(x(v))]} \frac{z(x(v))}{v}$$

or

$$\frac{v}{z(x(v))} = \frac{[1-v]}{[1-z(x(v))]} := m(v) \tag{7}$$

Substituting (7) into (6) we have:

$$u(x,v) - v = [z(x) - z(x(v))]\, m(v) \tag{8}$$

and using the boundary conditions, (i) and (ii), again we find that

$$u(w,v) - v = 0 - v = [0 - z(x(v))]\, m(v)$$

or

$$v - z(x(v))\, m(v) = 0 \tag{9}$$

and

$$u(b,v) - v = 1 - v = [1 - z(x(v))]\, m(v)$$

or

$$1 = m(v) + v - z(x(v))\, m(v) = m(v) \tag{10}$$

where the second equality is implied by (9). Therefore $m(v) = 1$ and using (7) and (8) we have

$$u(x,v) = z(x)$$

which implies that the local utility function is independent of v, hence preferences are expected utility.■

Proof of the sufficient conditions for mixed fan

Note that in the two-dimensional probability simplex, an indifference set is defined by

$v = p_{x_1}u(x_1, v) + (1 - p_{x_1} - p_{x_3})u(x_2, v) + p_{x_3}u(x_3, v)$. The slope of an indifference curve is then given by: $\frac{\partial p_3}{\partial p_1} = \frac{u(x_2,v)-u(x_1,v)}{u(x_3,v)-u(x_2,v)}$. Using theorem 1, it is evident that the requirement $\sum u(x, v)p_x \geqslant u(c(p), v)$ is equivalent to having the indifference curve through the $(0, 0)$ vertex being the steepest. Denote by v_i the solution to $\phi(x_i, v) - v = 0$. By monotonicity, for any triple $x_3 > x_2 > x_1$, and for any $v \in (v_1, v_3)$, $V^{-1}(v) \in L(x_3)$ and $V^{-1}(v) \notin L(x_1)$ but for the middle prize, x_2, both are possible. Let u_i denotes the partial derivative of u with respect to its i^{th} argument. The derivative (with respect to v) at $\widehat{v}$ of $\ln\left(\frac{dp_{x_3}}{dp_{x_1}}\right)$ is given by $\frac{[u_2(x_2,\widehat{v})-u_2(x_1,\widehat{v})]}{u(x_2,\widehat{v})-u(x_1,\widehat{v})} - \frac{[u_2(x_3,\widehat{v})-u_2(x_2,\widehat{v})]}{u(x_3,\widehat{v})-u(x_2,\widehat{v})}$. By assumptions (1)-(3), if $V^{-1}(\widehat{v}) \in L(x_2)$, this term is positive ("fanning out") whereas If $V^{-1}(\widehat{v}) \notin L(x_2)$, it is negative ("fanning in"). In particular, the indifference curve in the level $v = \phi^1(x_2, v) = \phi^2(x_2, v)$ is the steepest.■

Proof of proposition 2

Let $\triangle V(\beta | p, n)$:=grp$(\beta | p, n)$, anb for $k = 2, 3, ..., n-1$, denote $\triangle V(\beta | p, n)$ with $h(\beta | p, n) = n - k$ by $\triangle V^{(k)}(\beta | p, n)$. It can be shown that

$$\begin{aligned}&\triangle V^{(k)}(\beta | p, n)\\ = \; & np\beta(1-p)\frac{-(1-p)^{k-1}\left(-\beta\left(\sum_{j=0}^{n-(k+1)}\binom{j+k-2}{j}p^j\right) + p^{n-k}\left(\binom{n-2}{n-(k+1)}\beta + \binom{n-1}{n-k}\right)\right) + 1}{(1+\beta(1-p))\left(\beta\left(\sum_{j=k-1}^{n-1}\binom{j}{j-(k-1)}p^{j-(k-1)}\right)(1-p)^k + 1\right)}\end{aligned}$$

The denominator of $\triangle V^{(k)}(\beta | p, n)$ is always positive, whereas the coefficient $np\beta(1-p)$ is strictly positive for $\beta > 0$. At $\beta = 0$ the nominator is equal to $1 - \binom{n-1}{n-k}(1-p)^{k-1}p^{n-k}$ which is positive since $\binom{n-1}{n-k}(1-p)^{k-1}p^{n-k}$ is simply the probability of $n-k$ successes in $n-1$ trials of a Bernoulli random variable with parameter p. We then note that the nominator is also increasing with β. Indeed, this is the case if $\left(\sum_{j=0}^{n-(k+1)}\binom{j+k-2}{j}p^j\right) > p^{n-k}\binom{n-2}{n-(k+1)}$, which is true since $p < 1$ and $\sum_{j=0}^{n-(k+1)}\binom{j+k-2}{j} = \binom{n-2}{n-k-1}$. Therefore, item (i) is implied. Since $\beta = 0$ implies expected utility, the first part of item (ii) is immidiate. For the second part of item (ii), observe that as β increases, the value of the sequential lottery $(V(Q^n))$ is (smoothly) strictly decreasing and converges to 0, the value of the worst prize in its support. The value of the one stage lottery $(V(\widehat{p}))$ is affected in two ways when β increases: First, given a threshold $h(\beta | p, n)$, the value is (smootly) strictly decreasing with β. Second, $h(\beta | p, n)$ itself is a decreasing step-function of β. For β large enough, all prizes but 0 are elated and the value of the lottery is given by $\frac{\sum_{k=1}^{n}\binom{n}{k}p^k(1-p)^{n-k}k}{1+\beta(1-p)^n} \underset{\beta\to\infty}{\to} 0$.

To show the existence of β^* (item (iii)), pick $\beta' > 0$ such that grp$(\beta' | p, n) = \epsilon > 0$. Since $\lim_{\beta\to\infty}$ grp$(\beta | p, n) = 0$, there exists $\overline{\beta} := \max\left\{\beta \left| \text{grp}(\beta | p, n) = \frac{\epsilon}{2}\right.\right\}$ and $\overline{\beta} < \infty$. Thus grp$(\beta | p, n)$ is a continuous function on the compact interval $\left[0, \overline{\beta}\right]$, and hence achieves its maximum on this domain. For single-peakness, we have the following two claims:

Claim 1: *$\forall k=2,3,...,n-1$, $\triangle V^{(k)}(\beta\,|p,n)$ is either strictly increasing or single-peaked on* (0,∞).

Proof: By differentiating $\triangle V^{(k)}(\beta\,|p,n)$ with respect to β, one gets:

$$\begin{aligned}&\frac{\partial}{\partial\beta}\triangle V^{(k)}(\beta\,|p,n)\\ =\;& np\frac{C\beta^2+\left(2Ap^k(1-p)^k-2p^n\binom{n-2}{n-k-1}(1-p)^k\right)\beta+\left((1-p)p^k-p^n\binom{n-1}{n-k}(1-p)^k\right)}{p^k(-\beta+p\beta-1)^2\left(B\beta(-p+1)^k+1\right)^2}\end{aligned}$$

Where C is some constant, and $A:=\left(\sum_{j=0}^{n-(k+1)}\binom{j+k-2}{j}p^j\right)$.

The roots of $\frac{\partial}{\partial\beta}\triangle V^{(k)}(\beta\,|p,n)$ are the roots of the second-degree polynomial in β that appears in the nominator.

Evaluated at $\beta=0$, this polynomial is equal to $\left(p^k-pp^k-p^n\binom{n-1}{n-k}(1-p)^k\right)$. Note that

$$\left(p^k-pp^k-p^n\binom{n-1}{n-k}(1-p)^k\right)>0\Longleftrightarrow 1>\binom{n-1}{n-k}p^{n-k}(1-p)^{k-1}$$

which is true as claimed before.

In addition, the slope of that polynomial at $\beta=0$ is equal to the coefficient of β, $2Ap^k(1-p)^k-2p^n\binom{n-2}{n-k-1}(1-p)^k$, which is positive since $\left(\sum_{j=0}^{n-(k+1)}\binom{j+k-2}{j}p^j\right)>p^{n-k}\binom{n-2}{n-k-1}$.

To summarize, both the slope and the intercept of the polynomial in the nominator are positive at $\beta=0$. Therefore, if $C\geq 0$ then $\frac{\partial}{\partial\beta}\triangle V^{(k)}(\beta\,|p,n)$ has no positive roots, and otherwise it has exactly one positive root.$\|$

Note that $\triangle V(\beta\,|p,n)$ is a continuous function that is not differentiable in the points where $h(\beta\,|p,n)$ changes. For $k=2,3,...,n-1$, let $\beta_{k,k+1}$ be the value of β where $h(\beta\,|p,n)$decreases from $(n-k)$ to $(n-(k+1))$. Using the same notations as above, we claim that at the switch point, the slope of the resolution premium decreases.

Claim 2: $\lim_{\beta\to_-\beta_{k,k+1}}\frac{\partial}{\partial\beta}\triangle V^{(k)}(\beta\,|p,n)>\lim_{\beta\to_+\beta_{k,k+1}}\frac{\partial}{\partial\beta}\triangle V^{(k+1)}(\beta\,|p,n)$

Proof: Apart from at $\beta=0$, where $\triangle V^{(k)}(0\,|p,n)=\triangle V^{(k+1)}(0\,|p,n)=0$, it can be shown that the two curves cross at exactly one more point, given by

$$\beta_{k,k+1}=\frac{np-(n-k)}{\left(\sum_{j=0}^{n-(k+1)}(n-k-j)\binom{j+k-1}{j}p^j\right)(1-p)^{k+1}}$$

Note that $\beta_{k,k+1} > 0$ iff $p > \frac{n-k}{n}$. To prove the claim it will be sufficient to show that $\frac{\partial}{\partial\beta}\triangle V^{(k)}(0\,|p,n) < \frac{\partial}{\partial\beta}\triangle V^{(k+1)}(0\,|p,n)$, since this implies that at $\beta_{k,k+1}$, $\triangle V^{(k+1)}(\beta\,|p,n)$ crosses $\triangle V^{(k)}(\beta\,|p,n)$ from above. Now:

$\frac{\partial}{\partial\beta}\triangle V^{(k)}(0\,|p,n) = np\frac{\left(p^k - pp^k - p^n\binom{n-1}{n-k}(1-p)^k\right)}{p^k}$ and

$\frac{\partial}{\partial\beta}\triangle V^{(k+1)}(0\,|p,n) = np\frac{\left(p^{k+1} - pp^{k+1} - p^n\binom{n-1}{n-k-1}(1-p)^{k+1}\right)}{p^{k+1}}$

$$
\begin{aligned}
\text{so}\quad & \frac{\partial}{\partial\beta}\triangle V^{(k+1)}(0\,|p,n) > \frac{\partial}{\partial\beta}\triangle V^{(k)}(0\,|p,n) \\
& \iff \frac{1}{p^k} n(-p+1)^k p^n \left(p\binom{n-1}{-k+n} + p\binom{n-1}{-k+n-1} - \binom{n-1}{-k+n-1}\right) > 0 \\
& \iff p\binom{n-1}{-k+n} + p\binom{n-1}{-k+n-1} - \binom{n-1}{-k+n-1} > 0 \\
& \iff p > \frac{\binom{n-1}{-k+n-1}}{\binom{n-1}{-k+n} + \binom{n-1}{-k+n-1}} = \frac{(n-k)}{n}.\|
\end{aligned}
$$

To complete the proof we verify that both claims above are also valid for the two extreme cases, $k = 1$ (where only the best prize, n is elation) and $k = n$ (only the worst prize, 0 is disappointment).

$k = 1$: Using the same notation as used above we have:

$$\triangle V^{(1)}(\beta\,|p,n) = np\beta\left(\sum_{j=0}^{n-2} p^j\right)(p-1)^2 \frac{\beta+1}{(1+(1-p)\beta)(1+(1-p^n)\beta)}$$

and

$$\frac{\partial}{\partial\beta}\triangle V^{(1)}(\beta\,|p,n) = n(1-p)(p-p^n)\frac{(1-pp^n)\beta^2 + 2\beta + 1}{(-\beta + p^n\beta - 1)^2(-\beta + p\beta - 1)^2} > 0$$

for all $\beta \geq 0$ so $\triangle V^{(1)}(\beta\,|p,n)$ is strictly increasing with β (claim 1).

For the second claim, similar calculations as above establish that:

$$\frac{\partial}{\partial\beta}\triangle V^{(2)}(0\,|p,n) > \frac{\partial}{\partial\beta}\triangle V^{(1)}(0\,|p,n) \iff p > \frac{n-1}{n}$$

so claim 2 follows as well.

$k = n$:

$$\triangle V^{(n)}(\beta\,|p,n) = np^2\beta(1-p)\frac{\left(\sum_{j=1}^{n-1}\binom{n-1}{j}p^{j-1}(-1)^{j-1}\right)}{(1+\beta(1-p))(1+\beta(1-p)^n)}$$

Let $C = \left(\sum_{j=1}^{n-1} \binom{n-1}{j} p^{j-1} (-1)^{j-1} \right)$, so:

$$\frac{\partial}{\partial \beta} \triangle V^{(n)} (\beta | p, n) = Cnp^2 (p-1) \frac{\beta^2 (1-p)^{n+1} - 1}{(\beta(-p+1)^n + 1)^2 (-\beta + p\beta - 1)^2}$$

which is clearly single peaked *on* (0,∞) (claim 1), and, again by similar calculations:

$$\frac{\partial}{\partial \beta} \triangle V^{(n)} (0 | p, n) > \frac{\partial}{\partial \beta} \triangle V^{(n-1)} (0 | p, n) \Longleftrightarrow p > \frac{1}{n}$$

which is claim 2.

Combining claim 1 and claim 2 ensures that $\triangle V(\beta | p, n)$is single-peaked on (0,$\infty$).■

Proof of proposition 3

We first show that the claim is true for any lotteries of the form $p\delta_x + (1-p)\delta_y$, with $x > y$.

Case 1, $p = 0.5$:

Construct the compund lottery $Q^n \in \mathcal{P}(0.5\delta_x + 0.5\delta_y)$ as follows:

In each period $\Pr(\text{"success"}) = \Pr(\text{"failure"}) = 0.5$.

Define:

$$z_i = \begin{cases} 1 & [\text{"success"}] \\ 0 & [\text{"failure"}] \end{cases} \qquad i = 1, 2, 3, ..$$

The terminal nodes are:

$$\begin{array}{cl} \delta_x & \text{if } \sum_{i=1}^{n} z_i > \frac{n}{2} \\ 0.5\delta_x + 0.5\delta_y & \text{if } \sum_{i=1}^{n} z_i = \frac{n}{2} \\ \delta_y & \text{if } \sum_{i=1}^{n} z_i < \frac{n}{2} \end{array}$$

Claim:

$\lim_{n \to \infty} V(Q^n) = V(\delta_y) = \phi(y)$

Proof of claim: We use the fact that Value of the lottery using recursive Gul preferences and probability 0.5 for "success" in each period is equal to the value of the lottery using recursive expected utility and probability $\frac{0.5}{1+\beta 0.5}$ for "success" in each period.

Since $z_i's$ are i.i.d random variables, the weak law of large numbers implies:

$$\frac{\sum_{i=1}^{n} z_i}{n} \xrightarrow{p} \frac{0.5}{1 + \beta 0.5} < 0.5$$

or,

$$\Pr\left(\sum_{i=1}^{n} z_i < \frac{n}{2}\right) \to 1$$

Therefore

$$\begin{aligned} V(Q^n) &= \phi(x)\Pr\left(\sum_{i=1}^{n} z_i > \frac{n}{2}\right) + \\ &\quad \frac{0.5\phi(x) + (1+\beta)\,0.5\phi(y)}{1+\beta 0.5}\Pr\left(\sum_{i=1}^{n} z_i = \frac{n}{2}\right) + \\ &\quad \phi(y)\Pr\left(\sum_{i=1}^{n} z_i < \frac{n}{2}\right) \to \phi(y) \end{aligned}$$

case 2, $p < 0.5$:

Take $Q^{n+1} = \langle 2p, Q^n; 1-2p, \delta_y \rangle$, with Q^n as defined above.

case 3, $p > 0.5$:

Fix $\varepsilon > 0$. Using the construction in case 1, obtain Q^{T_1} with $V\left(Q^{T_1}\right) \in \left(\phi(y), \phi(y) + \frac{\varepsilon}{2}\right)$. Re-construct a lottery as above, but replace δ_y with Q^{T_1} in the terminal node. By the same argument, there exists T_2 and $V\left(Q^{T_1+T_2}\right) \in (\phi(y), \phi(y) + \varepsilon)$. Note that the underlying probability of y in $Q^{T_1+T_2}$ is 0.25. Therefore, by monotonicity, the construction works for any $p < 0.75$. Repeat in the same fashion to show that the assertion is true for $p^k < \frac{3+4k}{4+4k}$, $k = 1, 2, ..,$ and note that $p^k \to 1.\|$

Now take any finite lottery $\sum_{j=1}^{m} p_j \delta_{x_j}$ and order its prizes as $x_1 < x_2 < ... < x_m$. Repeat the construction above for the binary lottery x_{m-1}, x_m to make its value arbitrarily close to $\phi(x_{m-1})$. Then mix it appropriately with x_{m-2} and repeat the argument above. Continue in this fashion to get a multi-stage lottery over $x_2, ..., x_m$ with a value arbitrarily close to $\phi(x_2)$. Conclude by mixing it with x_1 and repeat the construction above.■

Proof of proposition 4

It is obvious that (i) is necessary for (ii). To show sufficiency, we introduce the intermediate lotteries Q and p^j, where the compound lottery Q assigns probability $\alpha_j(\pi)$ to p^j, and p^j assigns probability p_s^j to the outcome $u(a^*(s), s)$. Clearly, since for each state s and for any action a we have $u(a, s) \leq u(a^*(s), s)$, by monotonicity of the value of a lottery with respect to the relation of first-order stochastic dominance, $V(p^{j*}) \leq V(p^j)$, and hence, by the same reason, also $V^p(\pi) \leq V(Q)$.

However, now Q is simply the folding back of the two-stage lottery, which when played in one-shot is the lottery corresponding to full information structure, I. Thus by (i) we have that $V^p(I) \geq V(Q)$. Combining the two inequalities establishes the result.

Similarly, it is obvious that PGRU is necessary for ϕ being the least valuable information structure. To show sufficiency, define $V(a,p)$ as the value of a lottery in which with probability p_s you get the outcome $u(a,s)$. Let $\underline{a}= \arg\max_a V(a,p)$, then $V^p(\phi) = V(\underline{a},p)$. Let Q be a two-stage lottery that assigns probability $\alpha_j(\pi)$ to p^j and p^j assigns probability p_s^j to the outcome $u(\underline{a},s)$. By definition, $V(p^j) \leq V(p^{j*})$ for all j, and therefore, by monotonicity, $V(Q) \leq V^p(\pi)$. However, now Q is simply the folding back of the two-stage lottery, which when played in one-shot is the lottery corresponding to ϕ.Thus by (i) we have that $V^p(\phi) \leq V(Q)$. Combining the two inequalities establishes the result.■

References

[1] Allais, M., (1953), "Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine." *Econometrica,* 21, 503-546.

[2] Ang, A., Bekaert,G. and J. Liu., (2005), "Why Stocks May Disappoint." *Journal of Financial Economics*, 76, 471-508.

[3] Arrow, K.J., and C. Fischer, (1974), "Environmental Preservation, Uncertainty and Irreversibility." *Quarterly Journal of Economics,* 88, 312-19.

[4] Artstein-Avidan, S. and D. Dillenberger, (2006), "Don't tell me anything until you know for sure." mimeo.

[5] Barberis, N. and M. Huang, (2005), "The Loss Aversion / Narrow Framing Approach to the Equity Premium Puzzle." NBER Working Paper No. 12378.

[6] Barberis, N., Huang, M. and R. H. Thaler, (2006), "Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case for Narrow Framing." *American Economic Review*, 96, 1069-1090.

[7] Bellemare, C., Krause, M., Kröger, S.and C. Zheng, (2005), "Myopic Loss Aversion: Information Feedback vs. Investment Flexibility." *Economics Letters*, 87, 319-324.

[8] Benartzi, S. and R.Thaler, (1995), "Myopic loss aversion and the equity premium puzzle." *Quarterly Journal of Economics,* 110, 73–92.

[9] Blackwell, D., (1953), "Equivalent Comparison of Experiments." *Annals of Mathematics and Statistics,* 24, 265-272.

[10] Chew, S.H., (1989), "Axiomatic Utility Theories with the Betweenness Property." *Annals of Operations Research*, 19, 273-298.

[11] Chew, S.H. and L.G. Epstein, (1989), "The Structure of Preferences and Attitudes Towards the Timing of The Resolution of Uncertainty." *International Economic Review,* 30, 103-117.

[12] Conlisk, J., (1989), "Three Variations on the Allais Example." *The American Economic Review*, 79, 392-407.

[13] Debreu, G., (1972), "Smooth Preferences." *Econometrica*, 40, 603-615.

[14] Dekel, E., (1986), "An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom." *Journal of Economic Theory,* 40, 304-318.

[15] Dixit, A.K., and R.S. Pindyck, (1994), " Investment Under Uncertainty." Princeton: Princeton Univ. Press.

[16] Epstein, L.G. and S.E. Zin, (1989), "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: a Theoretical Framework." *Econometrica,* 57, 937-969.

[17] Gollier, C., (2001), "The Economics of Risk and Time." The MIT Press, 0-262-07215-7.

[18] Gneezy, U., J. Potters, (1997), "An experiment on risk taking and evaluation periods." *Quarterly Journal of Economics,* 112, 632–645.

[19] Grant,S., Kajii A., and B. Polak, (1998), "Intrinsic Preference for Information." *Journal of Economic Theory*, 83, 233-259.

[20] Grant,S., Kajii A., and B. Polak, (2000), "Temporal Resolution of Uncertainty and Recursive Non-Expected Utility Models." *Econometrica*, 68, 425-434.

[21] Gul, F., (1991), "A Theory of Disappointment Aversion." *Econometrica*, 59, 667-686.

[22] Haigh, M.S. and J.A. List, (2005), "Do professional traders exhibit myopic loss aversion? An experimental analysis." *The Journal of Finance*, 60, 523-534.

[23] Kahneman, D. and A.Tversky, (1979), "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica*, 47, 263-291.

[24] Knight, F., (1921), "Risk, Uncertainty and Profit." Boston, Mass: Houghton Mifflin. Reprint, London: London School of Economics, 1946.

[25] Köszegi, B. and M. Rabin, (2007), "A Model of Reference-Dependent Preferences", forthcoming in *Quarterly Journal of Economics.*

[26] Kreps, D. M. and E.L. Porteus, (1978), "Temporal Resolution of Uncertainty and Dynamic Choice Theory." *Econometrica,* 46, 185-200.

[27] Machina, M.J., (1982), "'Expected Utility' Analysis without the Independence Axiom." *Econometrica*, 50, 277-323.

[28] Mcdonald, R. and D.Siegel, (1986), "The Value of Waiting to Invest." *Quarterly Journal of Economics,* 101, 707-28.

[29] Mehra, R. and E.C. Prescott, (1985), "The Equity Premium: A Puzzle." *Journal of Monetary Economics*, 15, 145-161.

[30] Nayyar, S., (2004), "Disappointment Aversion and the Insurance Trap." mimeo.

[31] Nehring, K., (2005), "Notes on Expected Utility with Psychological Consequences: The Logic of Betweenness-Based Risk Preferences." mimeo.

[32] Neilson, W.S., (1992), "A Mixed Fan Hypothesis and its Implications for Behavior Towards Risk."*Journal of Economic Behavior and Organization,* 19, 197-211.

[33] Palacious-Huerta, I., (1999), "The Aversion to the Sequential Resolution of Uncertainty." *Journal of Risk and Uncertainty*,18, 249-269.

[34] Pindyck, R., (1991), "Irreversibility, Uncertainty and Investment." *Journal of Economic Literature,* 29, 1110-48.

[35] Quiggin, J., (1982), "A Theory of Anticipated Utility." *Journal of Economic Behavior and Organization,* 3, 323–343.

[36] Rabin, M., (2000), "Risk Aversion and Expected-Utility Theory: A Calibration Theorem." *Econometrica,* 68, 1281–1292.

[37] Rabin, M. and R. H. Thaler, (2001), "Anomalies: Risk Aversion." *Journal of Economic Perspectives*, 15, 219–232.

[38] Routledge, B. and S.E. Zin, (2004), "Generalized Disappointment Aversion and Asset prices." mimeo.

[39] Safra, Z. and E. Sulganik, (1995), "On the Nonexistence of Blackwell's Theorem-Type Results with General Preference Relations." *Journal of Risk and Uncertainty*, 10, 187-201.

[40] Safra, Z. and U. Segal, (2006), "Calibration Results for Non-Expected Utility Theories." working paper.

[41] Schlee, E., (1990), "The Value of Information in Anticipated Utility Theory." *Journal of Risk and Uncertainty,* 3, 83–92.

[42] Segal, U., (1990), "Two-stage Lotteries without the Reduction Axiom." *Econometrica,* 58, 349-377.

[43] Segal, U.and A. Spivak, (1990), "First order versus second order risk aversion." *Journal of Economic Theory*, 51, 111-125.