

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

David Dillenberger[†]

September 23, 2009

Abstract

Experimental evidence suggests that individuals are more risk averse when they perceive risk that is gradually resolved over time. We address these findings by studying a decision maker (DM) who has recursive, non-expected utility preferences over compound lotteries. DM has preferences for one-shot resolution of uncertainty (PORU) if he always prefers any compound lottery to be resolved in a single stage. We establish an equivalence between dynamic PORU and static preferences that are identified with commonly observed behavior 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, PORU captures “loss aversion with narrow framing”.

Keywords: Recursive preferences over compound lotteries, resolution of uncertainty, Allais paradox, narrow framing, negative certainty independence.

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

^{*}I am grateful to Faruk Gul and Wolfgang Pesendorfer for their invaluable advice during the development of the paper. I thank Roland Benabou, Eric Maskin, Stephen Morris and Klaus Nehring for their helpful discussions and comments. The co-editor and four anonymous referees provided valuable comments that improved the paper significantly. 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. This paper is based on the first chapter of my doctoral dissertation at Princeton University.

[†]Department of Economics, University of Pennsylvania, 160 McNeil Building, 3718 Locust Walk Philadelphia, Pennsylvania 19104-6297. E-mail: ddill@sas.upenn.edu

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 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 assume 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 links 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* and *recursivity*. Time neutrality says that DM does not care about the time in which uncertainty is resolved as long as resolution happens in a single stage. Recursivity is a substitution property, which says that if DM prefers a single-stage lottery p to a single-stage lottery q , then he also prefers to substitute q with p in any two-stage lottery containing q as an outcome. Under these assumptions, any two-stage lottery is subjectively transformed into a simpler, one-stage lottery. In particular, there exists a single preference relation over the set of one-stage lotteries that fully determines DM’s preferences over the richer domain of two-stage lotteries.

In order to link behavior in both domains, we introduce the following two properties: the

¹All 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 use both notions interchangeably.

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 and it captures the idea that the frequency at which the outcomes of a random process are evaluated is a relevant economic variable. This notion formalizes an idea first raised by Palacios-Huerta [1999] (to be further discussed in the literature review section).
- *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 to Kahneman and Tversky’s [1979] “certainty effect” hypothesis, though it does not imply that people weigh probabilities non-linearly. The restrictions NCI imposes on preferences are just enough to explain commonly observed behavior in the common-ratio version of the Allais paradox (also known as “common-ratio effect with a certain prize”). In particular, NCI allows the vNM-independence axiom to fail when the certainty effect is present.

Proposition 1 establishes that NCI and PORU are equivalent. 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 Proposition 1, since it provides new theoretical predictions for dynamic behavior, based on robust (static) empirical evidence.

In an extended model, we allow DM to take intermediate actions (between the two stages of a lottery) 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 non-expected utility preferences for which, when applied recursively, a perfect information system is always the most valuable. Proposition 2 shows that this property, that we term *preferences for perfect information*, is equivalent to PORU. As a corollary, NCI is both a necessary and sufficient condition to have preferences for perfect information.

PORU implies that preferences over one-stage lotteries are quasi-concave. Adding the requirement that they are also quasi-convex, we confine our attention to the class of betweenness preferences (Chew [1983], Dekel [1986]), one of the most commonly used non-expected utility models. Within this class, axiom NCI has its own static implications. First, it is equivalent to a geometric condition, which we term the *steepest middle slope*, that is imposed on the map of indifference curves in every unit probability triangle. The steepest middle slope property indicates that for every triple of prizes $x_3 > x_2 > x_1$, the marginal rate of substitution between a probability shift from x_2 to x_3 and a probability shift from x_2 to x_1 is maximized on the indifference curve through the lottery that yields x_2 for certain.² Second, betweenness preferences yield an implicit expected utility representation of the form $V(p) = \sum_x u(x, V(p)) p(x)$, where the utility index u depends on both the prize x and the utility level $V(p)$. Suppose that u is at least twice differentiable with respect to both arguments (which implies that the function V is Fréchet differentiable). Under this smoothness assumption, NCI turns out to be equivalent to the vNM-independence axiom.

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 with its single-stage counterpart. Similarly to the standard risk premium, the gradual resolution premium is measured in monetary terms. In the case where DM is both risk averse and displays PORU, these two forces magnify each other. This observation can explain why people often purchase dynamic insurance contracts, such as periodic insurance for electrical appliances and cellular phones, at much more than actuarially fair rates.

The gradual resolution premium can be significant, in the sense that if the resolution process is “long” enough, individuals might be extremely reluctant to take risks. To illustrate this, we first extend our results to preferences over arbitrary n -stage lotteries. The parameter n 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 terminal prizes) whose value is arbitrarily close to that 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.” Our result shows that preferences

²As defined by Machina [1982], this marginal rate of substitution is the amount of probability that should be shifted from x_2 to x_3 per a unit shift of probability from x_2 to x_1 in order to keep DM indifferent.

that display PORU may lead to excessively conservative investment strategies.

1.1. Related literature

Confining his attention to binary single-stage lotteries and to preferences from the rank-dependent utility class (RDU, Quiggin [1982]), Segal ([1987], [1990]) discusses sufficient conditions under which the desirability of a two-stage lottery decreases as the two stages become less degenerate. We show in Proposition 3 that these conditions cannot be extended to the general case, that is, the only RDU preferences that satisfy PORU are expected utility. 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 numerous applications. Ang, Bekaert and Liu [2005] use recursive disappointment aversion preferences to study a dynamic portfolio choice designed to obtain the most preferred probability distribution over final wealth levels. The general theory we suggest provides a way to understand 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.

Schmidt [1998] develops a static model of expected utility with certainty preferences. His notion of certainty preferences is very close to axiom NCI. In his model, the value of any non-degenerate lottery is the expectation of a utility index over prizes, u , whereas the value of the degenerate lottery that yields the prize x for sure is $v(x)$. The certainty effect is captured by requiring $v(x) > u(x)$ for all x . Schmidt’s model violates both continuity and monotonicity with respect to first-order stochastic dominance, while in this paper we confine our attention to preferences that satisfy both properties.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, that is, a dynamic aggregation rule that argues that when making a series of choices, individuals “bracket” them by making each choice in isolation.⁴ Benartzi and Thaler [1995] were the first to use this

³Continuity and monotonicity ensure that the certainty equivalent of each lottery is well defined. This fact is used when applying the recursive structure of Segal’s model.

⁴Narrow framing is an example of people’s tendency to evaluate risky decisions separately. This tendency is illustrated in Tversky and Kahneman [1981], and further studied in Kahneman and Lovallo [1993] and Read, Loewenstein, and Rabin [1999] among others. Barberis and Huang [2007] present an extensive survey of this approach.

approach to suggest explanations for several economic “anomalies”, such as the equity premium puzzle (Mehra and Prescott [1985]). Barberis and Huang [2007] 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.

Our model can be used to address similar phenomena. The combination of the folding-back procedure and a specific form of 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 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.

Kőszegi and Rabin [2009] study a model in which utility additively depends on both current consumption and on recent changes in (rational) beliefs about present and future consumption, where the latter component displays loss aversion. In their setting, they identify narrow framing with preference over such fluctuations in beliefs. They also show that people prefer to get information clumped together (similar to PORU) rather than apart. Aside from the same conceptual differences between the two approaches, their set of results concerning information preferences is confined to the case where consumption happens only in the last period and is binary. This corresponds in our setup to lotteries over only two monetary prizes. Our results are valid for lotteries with arbitrary (finite) support.

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 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 results. Section 3 comments on the implications of our model on preferences over information systems. In section 4, we elaborate on the static implications of our model and provide examples. Section 5 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 magnifying effect. In section 6, we relate our approach to the notion of loss aversion with narrow framing. 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}^1 be the set of all simple lotteries (probability measures with finite support) over X . That is, each $p \in \mathcal{L}^1$ is a function $p : X \rightarrow [0, 1]$, satisfying $\sum_{x \in X} p(x) = 1$, and we restrict our analysis to the case where in any given lottery, the number of prizes with non-zero probability is finite. Let $S(p) = \{x | p(x) > 0\}$. For each $p, q \in \mathcal{L}^1$ and $\alpha \in (0, 1)$, the mixture $\alpha p + (1 - \alpha)q \in \mathcal{L}^1$ is the simple lottery that yields each prize x with probability $\alpha p(x) + (1 - \alpha)q(x)$. We denote by $\delta_x \in \mathcal{L}^1$ the degenerate lottery that gives the prize x with certainty, that is, $\delta_x(x) = 1$. Note that for any lottery $p \in \mathcal{L}^1$ we have $p = \sum_{x \in X} p(x) \delta_x$.

Correspondingly, let \mathcal{L}^2 be the set of all simple lotteries over \mathcal{L}^1 . That is, each $Q \in \mathcal{L}^2$ is a function $Q : \mathcal{L}^1 \rightarrow [0, 1]$, satisfying $\sum_{p \in \mathcal{L}^1} Q(p) = 1$. For each $P, Q \in \mathcal{L}^2$ and $\lambda \in (0, 1)$, the mixture $R = \lambda P + (1 - \lambda)Q \in \mathcal{L}^2$ is the two-stage lottery for which $R(p) = \lambda P(p) + (1 - \lambda)Q(p)$. We denote by $D_p \in \mathcal{L}^2$ the degenerate, in the first stage, compound lottery that gives lottery p in the second stage with certainty, that is, $D_p(p) = 1$. Note that for any lottery $Q \in \mathcal{L}^2$ we have $Q = \sum_{q \in \mathcal{L}^1} Q(q) D_q$. We think of each $Q \in \mathcal{L}^2$ as a dynamic two-stage process where, in the first stage, a lottery q is realized with probability $Q(q)$, and, in the second stage, a prize is obtained according to q .

Two special subsets of \mathcal{L}^2 are $\Gamma = \{D_p | p \in \mathcal{L}^1\}$, the set of degenerate lotteries in \mathcal{L}^2 and $\Lambda = \{Q \in \mathcal{L}^2 | Q(p) > 0 \Rightarrow p = \delta_x \text{ for some } x \in X\}$, the set of lotteries in \mathcal{L}^2 , outcome of which are degenerate in \mathcal{L}^1 . Note that both Γ and Λ are isomorphic to \mathcal{L}^1 .

Let \succeq be a continuous (in the topology of weak convergence) preference relation over \mathcal{L}^2 . Let \succeq_Γ and \succeq_Λ be the restriction of \succeq to Γ and Λ respectively. On \succeq we impose the following axioms:

A0 (*more is better*): $\forall x, y \in X, x \geq y \Leftrightarrow D_{\delta_x} \succeq D_{\delta_y}$

A1 (*time neutrality*): $\forall p \in \mathcal{L}^1, D_p \sim \sum_{x \in X} p(x) D_{\delta_x}$

A2 (*recursivity*): $\forall q, p \in \mathcal{L}^1, \text{ all } Q \in \mathcal{L}^2, \text{ and } \lambda \in (0, 1),$

$$D_p \succeq D_q \iff \lambda D_p + (1 - \lambda) Q \succeq \lambda D_q + (1 - \lambda) Q$$

A0 is a weak monotonicity assumption. 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 degenerate two-stage lotteries (that is, 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.

Lemma 1: *If \succeq satisfies A0, A1 and A2, then both \succeq_Γ and \succeq_Λ are monotone (with respect to the relation of first-order stochastic dominance).⁵*

Proposition (Segal [1990]): *\succeq satisfies A0, A1 and A2 if and only if, there exists a continuous function $V : \mathcal{L}^1 \rightarrow \mathbb{R}$, such that the certainty equivalent function $c : \mathcal{L}^1 \rightarrow X$ is given by $V(\delta_{c(p)}) = V(p)$ for all $p \in \mathcal{L}^1$, and for all $P, Q \in \mathcal{L}^2$:*

$$P \succeq Q \iff V\left(\sum_{p \in \mathcal{L}^1} P(p) \delta_{c(p)}\right) \geq V\left(\sum_{p \in \mathcal{L}^1} Q(p) \delta_{c(p)}\right).$$

Note that under A0, A1 and A2, the preference relation $\succeq_\Gamma = \succeq_\Lambda$ 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. Since only the function V appears in the formula above, we slightly abuse notation by writing $V(Q)$ for the value of the two-stage lottery Q . Lastly, since under the above assumptions $V(p) = V(D_p) = V(\sum_{x \in X} p(x) D_{\delta_x})$ 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. Our static properties are imposed on preference relations over

⁵ A0, A1 and A2 imply that both \succeq_Γ and \succeq_Λ satisfy the axiom of degenerate independence, ADI (Grant, Kajii and Polak [1992]). Simple induction arguments show that ADI is equivalent to monotonicity with respect to the relation of first-order stochastic dominance.

sets that are isomorphic to \mathcal{L}^1 (such as \succeq_Γ and \succeq_Λ). We denote by \succeq_1 such a generic preference relation and assume throughout that it is continuous and monotone.

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.

Define $\rho : \mathcal{L}^2 \rightarrow \mathcal{L}^1$ to be the reduction operator that maps a compound lottery to its reduced single-stage counterpart, that is, $\rho(Q) = \sum_{q \in \mathcal{L}^1} Q(q) q$. Note that by A1, $D_{\rho(Q)} \sim \sum_{x \in X} \left[\sum_{q \in \mathcal{L}^1} Q(q) q(x) \right] D_{\delta_x}$.

Definition 1: The preference relation \succeq displays *preference for one-shot resolution of uncertainty* (PORU) if $\forall Q \in \mathcal{L}^2, D_{\rho(Q)} \succeq Q$. If $\forall Q \in \mathcal{L}^2, Q \succeq D_{\rho(Q)}$, 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.⁶

2.2.2. The ratio Allais paradox and axiom NCI

In a generic Allais-type questionnaire (also known as “common-ratio effect with a certain prize”) subjects choose between A and B , where $A = \delta_{3000}$ and $B = 0.8\delta_{4000} + 0.2\delta_0$. They also choose between C and D , where $C = 0.25\delta_{3000} + 0.75\delta_0$ and $D = 0.2\delta_{4000} + 0.8\delta_0$. The majority of subjects tend to *systematically* violate expected utility by choosing the pair A and D .⁷

Since Allais’s [1953] original work, numerous versions of his questionnaire have appeared, many of which contain one lottery that does not involve any risk.⁸ Kahneman and Tversky

⁶Halevy [2007] provides some evidence in favor of PORU. In his paper, subjects were asked to state their reservation prices for four different compound lotteries. His results (which are discussed in section 4.2.1 of his paper) show that the reservation prices of the two degenerate two-stage lotteries ($V1$ and $V4$, members of Λ and Γ respectively) were approximately the same and larger than the reservation price of the gradually resolved lottery ($V3$).

⁷This example is taken from Kahneman and Tversky [1979]. Of 95 subjects, 80% choose A over B , 65% choose D over C , and more than half choose the pair A and D .

⁸Camerer [1995] is an extensive survey of the experimental evidence against expected utility, including the “common consequence effect” and “common ratio effect” that are related to Allais paradox.

[1979] 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. This reasoning 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 suggest a property that is motivated by similar insights and captures the certainty effect without implying that people weigh probabilities non-linearly. Consider the following axiom on \succeq_1 :

Negative Certainty Independence (NCI): $\forall p, q, \delta_x \in \mathcal{L}^1$ 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 of p . If we define $c(p|\lambda, q)$, the conditional certainty equivalent of a lottery p , as the solution to $\lambda p + (1 - \lambda)q \sim_1 \lambda \delta_{c(p|\lambda, q)} + (1 - \lambda)q$, then the axiom implies that $c(p|\lambda, q) \geq c(p)$ for all $p, q \in \mathcal{L}^1$ and $\lambda \in (0, 1)$. The implication of this axiom on responses to the Allais questionnaire above is as follows: If you choose the non-degenerate lottery B , then you must also choose D . This prediction is empirically rarely violated in versions of the Allais questionnaire that involved positive outcomes.^{9,10} As we have mentioned before, 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 . NCI, however, does not imply any probabilistic distortion. This observation becomes relevant in experiments similar to the one reported in Conlisk [1989, p.398], who studies the robustness of Allais-type behavior to boundary effects. Conlisk considers a slight perturbation of prospects similar to A, B, C and D above, so that (i) each of the new prospects, A', B', C' and D' , yields all three prizes with strictly positive probability, and (ii) in the resulting “displaced Allais question” (namely choosing between A' and B' and then choosing between C' and D'), the only pattern of choice that is consistent with expected utility is either the pair A' and C' or the pair B' and D' . Although violations of expected utility become significantly less frequent and are no longer systematic (a result that supports the claim that violations can be explained by the certainty effect), a nonlinear probability function predicts that this increase in consistency would be the result

⁹Conlisk [1989], for example, replicates the two basic Allais questions. About half of his subjects (119 out of 236) violate expected utility. The fraction of violations that are of the B and C type is $16/119 \simeq 0.13$.

¹⁰It is worth mentioning that there is also some empirical and experimental evidence that conflicts with NCI. For example, NCI will be inconsistent with the “reflection effect”, that is, a common ratio effect with negative numbers (Kahneman and Tversky [1979], Machina [1987]). I thank a referee for pointing this out.

of fewer subjects choosing A' over B' , and not because more subjects choose C' over D' . In fact, the latter occurred, which is consistent with NCI.

Proposition 1: *Under $A0$, $A1$ and $A2$, \succeq_1 satisfies NCI if and only if \succeq displays PORU.*

Proof (only if): Suppose \succeq_1 satisfies NCI. We need to show that an arbitrary two-stage lottery, Q is never preferred to its single-stage counterpart, $D_{\rho(Q)}$. Without loss of generality, assume that there are l outcomes in the support of Q . Using $A0$ – $A2$ we have:

$$Q = \sum_{i=1}^l Q(q^i) D_{q^i} \stackrel{(A0, A2)}{\sim} \sum_{i=1}^l Q(q^i) D_{\delta_{c(q^i)}} \stackrel{(A1)}{\sim} D_{\sum_{i=1}^l Q(q^i) \delta_{c(q^i)}}$$

and by repeatedly applying NCI,

$$\begin{aligned} \sum_{i=1}^l Q(q^i) \delta_{c(q^i)} &= Q(q^1) \delta_{c(q^1)} + (1 - Q(q^1)) \left(\sum_{i \neq 1} \frac{Q(q^i)}{(1 - Q(q^1))} \delta_{c(q^i)} \right) \stackrel{(NCI)}{\preceq_1} \\ Q(q^1) q^1 + (1 - Q(q^1)) &\left(\sum_{i \neq 1} \frac{Q(q^i)}{(1 - Q(q^1))} \delta_{c(q^i)} \right) = \\ Q(q^2) \delta_{c(q^2)} + (1 - Q(q^2)) &\left(\frac{Q(q^1)}{(1 - Q(q^2))} q^1 + \sum_{i \neq 1, 2} \frac{Q(q^i)}{(1 - Q(q^2))} \delta_{c(q^i)} \right) \stackrel{(NCI)}{\preceq_1} \\ Q(q^1) q^1 + Q(q^2) q^2 + \sum_{i \neq 1, 2} Q(q^i) &\delta_{c(q^i)} = \dots = \\ Q(q^l) \delta_{c(q^l)} + (1 - Q(q^l)) &\left(\sum_{i \neq l} \frac{Q(q^i)}{(1 - Q(q^l))} q^i \right) \stackrel{(NCI)}{\preceq_1} \sum_{i=1}^l Q(q^i) q^i = \rho(Q) \end{aligned}$$

Therefore,

$$Q = \sum_{i=1}^l Q(q^i) D_{q^i} \sim D_{\sum_{i=1}^l Q(q^i) \delta_{c(q^i)}} \preceq D_{\rho(Q)}.$$

(if): Suppose \succeq_1 does not satisfy NCI. Then there exists $p, q = \sum_x q(x) \delta_x, \delta_y \in \mathcal{L}^1$ and $\lambda \in (0, 1)$, such that $p \succeq_1 \delta_y$ and $\lambda \delta_y + (1 - \lambda)q \succ_1 \lambda p + (1 - \lambda)q$. By monotonicity, $\lambda \delta_{c(p)} + (1 - \lambda)q \succ_1 \lambda p + (1 - \lambda)q$. Let $Q := \lambda D_p + (1 - \lambda) \sum_x q(x) D_{\delta_x}$ and note that,

$$Q \sim \lambda D_{\delta_{c(p)}} + (1 - \lambda) \sum_x q(x) D_{\delta_x} \succ \sum_x [\lambda p(x) + (1 - \lambda)q(x)] D_{\delta_x} \sim D_{\rho(Q)}$$

Violating PORU. ■

Proposition 1 ties together two notions that are defined on different domains. The equivalence of PORU and NCI 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 affects behavior in a systematic way. On the other hand, empirical and experimental studies involving dynamic choices and experimental studies on preferences for uncertainty resolution are still rather rare. Proposition 1 thus provides new theoretical predictions for dynamic behavior, based on robust (static) empirical evidence.

3. PORU and the value of information

Suppose now that before the second stage lottery is played, but after the realization of the first stage lottery, the decision maker can take 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 A0–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 has no implications for the comparison between getting complete and partial information. Safra and Sulganik [1995] left open the question of whether there are static preference relations, other than expected utility, for which, when applied recursively, perfect information is always the most valuable. We show below that this property is equivalent to PORU. As a corollary, such preferences for perfect information are fully characterized by NCI.

Formally, fix an interval of monetary prizes $X \subset R$. Let $S = \{s_1, \dots, s_N\}$ be a finite set of possible states of nature. Each state $s \in S$ occurs with probability p_s . Let $J = \{j_1, \dots, j_M\}$ be a finite set of signals, and let $A = \{a_1, \dots, a_H\}$ be a finite set of actions. Let $u : A \times S \rightarrow X$ be a function that gives the deterministic outcome $u(a, s)$ (an element of X) if action $a \in A$ is taken and the realized state is $s \in S$. The collection $\Omega = \{S, J, A, (p_s)_{s \in S}, u\}$ is called an information environment.

Let $\pi : S \times J \rightarrow [0, 1]$ be a function such that $\pi(s, j)$ is the conditional 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 signals DM might get). The function π is called an information system.

For any $s \in S$, denote the updated probability of s after the signal $j \in J$ is obtained by

$p(s|j) = \frac{\pi(s,j)p_s}{\sum_{s' \in S} \pi(s',j)p_{s'}}$. A full information system, I , is a function such that for all $s \in S$ there exists $j(s) \in J$ with $p(s|j(s)) = 1$. The null information system, ϕ , is a function such that $p(s|j) = p_s$ for all $s \in S$ and $j \in J$.

Let $p^j(a) \in \mathcal{L}^1$ be the second-stage lottery if signal j is obtained and action $a \in A$ is taken, that is, $p^j(a) = \sum_{s \in S} p(s|j) \delta_{u(a,s)}$. For $a_j \in \arg \max_{a \in A} V(p^j(a))$, let $p^{j*} := p^j(a_j)$. Let $V(\pi) := V\left(\sum_{j \in J} \left(\sum_{s \in S} \pi(s,j) p_s\right) D_{p^{j*}}\right)$ be the value of the optimal compound lottery, that is, the compound lottery assigning probability $\alpha_j(\pi) = \sum_{s \in S} \pi(s,j) p_s$ to p^{j*} . Note that $V(\phi) = \max_{a \in A} V\left(\sum_{s \in S} p_s \delta_{u(a,s)}\right)$, and that $V(I) = V\left(\sum_{s \in S} p_s \delta_{u(a(s),s)}\right)$, where $a(s)$ is an optimal action if you know that the prevailing state is s , that is, $a(s) \in \arg \max_{a \in A} u(a,s)$.

Definition 2: \succeq displays *preferences for perfect information* if for every information environment Ω and any information system π , $V(I) \geq V(\pi)$.

Proposition 2: If \succeq satisfies A0 – 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 every information environment Ω and any information system π , $V(\pi) \geq V(\phi)$.

Since any temporal lottery corresponds to an information environment in which for all $a \in A$, $u(a,s) = v(s) \in X$, 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 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 system is of the “one-shot resolution” type and it has no instrumental value. By combining Proposition 1 and Proposition 2 we get:

Corollary 1: If \succeq satisfies A0 – A2, then \succeq displays preferences for perfect information if and only if \succeq_1 satisfies NCI.

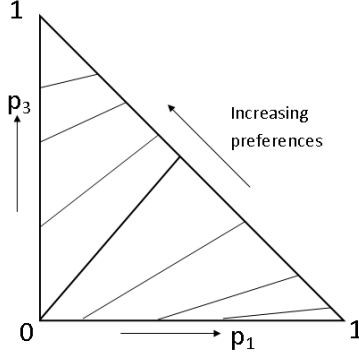


Figure 1: The probability triangle (showing linear indifference curves). The bold indifference curve through the origin demonstrates the steepest middle slope property (Lemma 3).

4. Static implications

4.1. NCI in the probability triangle

Fix three prizes $x_3 > x_2 > x_1$. All lotteries over these prizes can be represented as points in a two-dimensional space, $\Delta := \{p = (p_1, p_3) \mid p_1, p_3 \geq 0, p_1 + p_3 \leq 1\}$, as in figure 1. The origin $(0, 0)$ represents the lottery δ_{x_2} . The probability of the high prize, $p(x_3) = p_3$, is measured on the vertical axis, and the probability of the low prize, $p(x_1) = p_1$, is measured on the horizontal axis. The probability of obtaining the middle prize is $p(x_2) = p_2 = 1 - p_1 - p_3$. Given these conventions, monotonicity implies that preferences increase in the northwest direction. The properties below are geometric restrictions that NCI (hence PORU) imposes on the map of indifference curves in any probability triangle Δ , that corresponds to some triple $x_3 > x_2 > x_1$.

Lemma 2 (quasi-concavity): *If \succeq_1 satisfies NCI, then V is quasi-concave, that is, $V(\alpha p + (1 - \alpha)q) \geq \min\{V(p), V(q)\}$.*

Corollary 2: *If \succeq_1 satisfies NCI, then all indifference curves in Δ are convex.*

Let $\mu(p)$ be the slope, relative to the (p_1, p_3) coordinates, of the indifference curve at lottery p . $\mu(p)$ is the marginal rate of substitution between a probability shift from x_2 to x_3 and a probability shift from x_2 to x_1 . As explained by Machina [1982], changes in the slope express local changes in attitude towards risk: the greater the slope, the more (local) risk-averse DM is. Denote by $\mu^+(p)$ the right derivative of the indifference curve at p and by $\text{int}(\Delta)$ the interior of Δ . Let $\mathcal{I}_{(p)} := \{q \in \Delta \mid q \sim p\}$.

Definition 3: We say that V satisfies the *steepest middle slope property* if

(i) the indifference curve through the origin is linear, that is, $q \in \mathcal{I}_{((0,0))}$ implies $\mu(q) = \mu^+((0,0)) := \mu(\mathcal{I}_{((0,0))})$

(ii) the indifference curve through the origin is the steepest, that is, $\mu(\mathcal{I}_{((0,0))}) \geq \mu^+(q)$ for all $q \in \text{int}(\Delta)$.¹¹

Lemma 3 (steepest middle slope): *If \succeq_1 satisfies NCI, then V satisfies the steepest middle slope property.*

Examples of preferences that satisfy NCI will be given in section 4.2.1. For now, we use both lemmas to argue that two broad and widely used classes of preferences, rank-dependent utility (RDU, Quiggin [1982]) and quadratic utility (Chew, Epstein and Segal [1991]), do not satisfy NCI unless they coincide with expected utility.

Order the prizes $x_1 < x_2 < \dots < x_n$. The functional form for RDU is:

$$V(\sum_{i=1}^n p(x_i) \delta_{x_i}) = g(p(x_1)) u(x_1) + \sum_{i=2}^n u(x_i) \left[g\left(\sum_{j=1}^i p(x_j)\right) - g\left(\sum_{j=1}^{i-1} p(x_j)\right) \right]$$

where $g : [0, 1] \rightarrow [0, 1]$ is increasing, $g(0) = 0$ and $g(1) = 1$. If $g(p) = p$ then RDU reduces to expected utility.

The functional form for quadratic utility is:

$$V(\sum_{i=1}^n p(x_i) \delta_{x_i}) = \sum_{i=1}^n \sum_{j=1}^n \varphi(x_i, x_j) p(x_i) p(x_j)$$

where $\varphi(x_i, x_j) : X \times X \rightarrow \mathbb{R}$ is some symmetric function. If $\varphi(x_i, x_j) = \frac{u(x_i) + u(x_j)}{2}$ then quadratic utility reduces to expected utility.

Proposition 3: *If \succeq_1 satisfies NCI and is a member of either the RDU class or the quadratic utility class, then V is an expected utility functional.*¹²

¹¹By Corollary 2, all the right derivatives exist (see Rockafellar [1970], p.214).

¹²Segal [1990, section 5] used a different, but equivalent, way to write the functional form for RDU, using the transformation $f(p) = 1 - g(1 - p)$. He claims that within this model, if f is convex and its elasticity is non-decreasing, then the desirability of a two-stage lottery of the form $\alpha D_{\delta_y} + (1 - \alpha) D_{\beta \delta_y + (1 - \beta) \delta_x}$ decreases as the two stages become less degenerate. Similar results are stated in Segal [1987, theorem 4.2]. This condition is not sufficient to imply global PORU. For example, let $f(p) = p^2$, which satisfies Segal's conditions and $u(x) = x$. Take 3 prizes, 0, 1, and 2 and note that

$$V\left(\frac{1}{2}D_{\delta_1} + \frac{1}{2}D_{\frac{\sqrt{2}-1}{\sqrt{2}}\delta_0 + \frac{1}{\sqrt{2}}\delta_2}\right) = 1 > 0.853 = V\left(\frac{\sqrt{2}-1}{2\sqrt{2}}\delta_0 + \frac{1}{2}\delta_1 + \frac{1}{2\sqrt{2}}\delta_2\right)$$

Confining his attention to smooth preferences, in the sense that the function V is Fréchet differentiable, Machina [1982] suggests the following fanning out property: For all $p, q \in \Delta$, if p first-order stochastically dominates q , then $\mu(p) \geq \mu(q)$. If for all such $p \neq q$ we have $\mu(p) > \mu(q)$ then we say that \succeq_1 satisfies the proper fanning out property. Lemma 3 immediately implies that if \succeq_1 satisfies NCI, then \succeq_1 does not satisfy the proper fanning out property. This observation does not contradict the usual explanation of fanning out as a resolution to Allais paradox. Typical Allais experiments with positive outcomes (as the one described in section 2) provide evidence of behavior in the lower-right sub-triangle. In this region, NCI is consistent with fanning out.¹³

4.2. Betweenness

For the rest of the section, assume that \succeq_1 is quasi-convex, that is, $\forall p, q \in \mathcal{L}^1$, $V(\alpha p + (1 - \alpha)q) \leq \max\{V(p), V(q)\}$. The conjunction of quasi-convexity with quasi-concavity (Lemma 1) yields:

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 lotteries. It yields the following representation:

Proposition (Chew [1983], Dekel [1986]): \succeq_1 satisfies A3 iff there exists a function $u : X \times [0, 1] \rightarrow [0, 1]$, 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 $p \succeq q \Leftrightarrow V(p) \geq V(q)$, where $V(p)$ is defined implicitly as the unique $v \in [0, 1]$ that solves:

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

The betweenness axiom (A3), along with monotonicity, implies that indifference curves in any unit probability triangle are positively sloped straight lines. In particular, for any lottery $p \in \Delta$ such that $V(p) = v$,

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

¹³The behavioral evidence supporting fanning out is generally weaker in the upper-left sub triangle, the more-preferred region, than in the lower region (see Camerer [1995]).

In the case of betweenness, the steepest middle slope property reduces to the following condition: For every triple $x_3 > x_2 > x_1$ and for all $v \in (V(\delta_{x_1}), V(\delta_{x_3}))$,

$$\mu(V((0, 0)) | x_3, x_2, x_1) \geq \mu(v | x_3, x_2, x_1)$$

Our next result shows that within the betweenness class, this geometric condition is also sufficient to PORU.

Definition 4: \succeq_2 is *betweenness-recursive* if it satisfies A0–A2 and its \succeq_1 -type restrictions satisfy A3.

Proposition 4: *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 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.

4.2.1. 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 linear mixed-fan hypothesis. This set consists of all preferences whose indifference curves, in any unit probability triangle, are straight lines and

have the following pattern: Moving northwest, they first get steeper (fanning out) in the lower sub-triangle and then get flatter (fanning in) in the upper sub-triangle. The switch between fanning out and fanning in always occurs at the indifference curve that passes through the origin, so that the steepest middle slope property is satisfied.

Gul's [1991] model of disappointment aversion accommodates preferences that satisfy the linear mixed-fan hypothesis. Gul derives the local utility function

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

with $\beta > 0$ and $\phi : X \rightarrow \mathbb{R}$ 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. A prize x is a disappointing outcome in lottery p if $V(p) > V(\delta_x)$ and is an elating outcome otherwise. When calculating the (implicit) expected utility of a lottery, all disappointing outcomes get a uniformly greater weight.¹⁴ For Gul's preferences, the sign of β , the coefficient of disappointment aversion, unambiguously determines whether preferences satisfy PORU or PGRU (see Artstein-Avidan and Dillenberger [2006]).

4.2.2. NCI and differentiability

In most economic applications, it is assumed that individuals' preferences are "smooth". Confine our attention to the betweenness class, and suppose that $u : X \times [0, 1] \rightarrow [0, 1]$ is sufficiently differentiable with respect to both arguments. In this case, the function V is (continuously) Fréchet differentiable (Epstein [1993], Wang [1993]).¹⁵ The following result demonstrates that coupling this smoothness assumption with NCI leads us back to expected utility.

Proposition 5: *Suppose $u(x, v)$ is at least twice differentiable with respect to both 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 \bar{v} and denote by $x(\bar{v})$ the unique x satisfying

¹⁴Although 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.

¹⁵The notion of smoothness we consider here is the one assumed in Neilson [1992]. For a formal definition of Fréchet differentiability, see Machina [1982]. Roughly speaking, Fréchet differentiability means that $V(p)$ changes continuously with p and that V can be locally approximated by a linear functional. The Economic Meaning of Fréchet differentiability is discussed in Safra and Segal [2002].

$\bar{v} = u(x, \bar{v})$. Combining the geometric characterization (Proposition 4 item (iii)) of NCI with differentiability implies that for any $x > x(\bar{v}) > w$, the derivative with respect to v of the slope of an indifference curve on the corresponding probability triangle must vanish at \bar{v} . We use the fact that this statement is true for any $x > x(\bar{v})$ and that \bar{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 $x < x(\bar{v}) < b$ 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.¹⁶

5. Gradual resolution premium

We now extend our results to finite-stage lotteries.

5.1. Extension to n -stage lotteries

Fix $n \in \mathbb{N}$ and denote the space of finite n -stage lotteries by \mathcal{L}^n . The parameter n describes the frequency with which an individual updates information in a fixed time interval. The extension of our setting to \mathcal{L}^n is the following (a formal description is given in the appendix): equipped with a continuous and increasing function $V : \mathcal{L}^1 \rightarrow \mathbb{R}$, 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.

5.2. Definitions

For any $p \in \mathcal{L}^1$, denote by $e(p)$ the expectation of p , that is, $e(p) = \sum_x xp(x)$. Let $G(p, x) := \sum_{z \geq x} p(z)$. We say that lottery p second-order stochastically dominates lottery q , and denote it by $p \text{ sosd } q$, if for all $t < K$, $\sum_{k=0}^t [G(p, x_{k+1}) - G(q, x_{k+1})][x_{k+1} - x_k] \geq 0$, where

¹⁶Neilson [1992] provides sufficient conditions for smooth (in the sense of Proposition 5) 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 Proposition 5.

$x_0 < x_1 < \dots < x_K$ and $\{x_0, x_1, \dots, x_K\} = S(p) \cup S(q)$. DM is risk averse if $\forall p, q \in \mathcal{L}^1$ with $e(p) = e(q)$, p sosl q implies $p \succeq_1 q$.

For any $p \in \mathcal{L}^1$, the risk premium of p , denoted by $\text{rp}(p)$, is the number satisfying $\delta_{e(p)-\text{rp}(p)} \sim_1 p$. $\text{rp}(p)$ is the amount that DM would pay to replace p with its expected value. By definition, $\text{rp}(p) \geq 0$ whenever DM is risk averse.¹⁷

Definition 5: Fix $p \in \mathcal{L}^1$, and let $\mathcal{P}(p) := \{Q \mid \rho(Q) = p\}$. For any $Q \in \mathcal{P}(p)$, the *gradual resolution premium* of Q , denoted by $\text{grp}(Q)$, is the number satisfying $\langle 1, \delta_{c(p)-\text{grp}(Q)} \rangle \sim Q$.

$\text{grp}(Q)$ is the amount that DM would pay to replace Q with its single-stage counterpart. By definition, PORU implies $\text{grp}(Q) \geq 0$. Since $c(p) = e(p) - \text{rp}(p)$, we can, equivalently, define $\text{grp}(Q)$ as the number satisfying $\langle 1, \delta_{e(p)-\text{rp}(p)-\text{grp}(Q)} \rangle \sim Q$.¹⁸

Observe that the signs of $\text{rp}(p)$ and $\text{grp}(Q)$, 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 $\beta \geq 0$ and $\phi : X \rightarrow \mathbb{R}$ is concave (Gul's [1991] theorem 3). However, for sufficiently small $\beta \geq 0$ and sufficiently ϕ , one can find a lottery p with $\text{rp}(p) < 0$, whereas $\beta \geq 0$ is sufficient for $\text{grp}(Q) \geq 0$ for all $Q \in \mathcal{P}(p)$. On the other hand, if $\lambda'(v) > 0$ and $\lambda(v) > 1$ for all v ,¹⁹ then the local utility function

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

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 $\text{rp}(p) \geq 0 \forall p \in \mathcal{L}^1$. However, these preferences do not satisfy NCI,²⁰ meaning that there exists $Q \in \mathcal{P}(p)$ with $\text{grp}(Q) < 0$.

¹⁷Weak 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.

¹⁸The gradual resolution premium is measured in monetary units. It is different from the timing premium for early resolution, as suggested by Chew and Epstein [1989], which is defined in terms of probabilities.

¹⁹The condition that $\lambda(v)$ is non-decreasing is both necessary and sufficient for u to be a local utility function. See Nehring [2005].

²⁰Look 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 item (iii) in Proposition 4.

5.3. The magnifying effect

In the case where DM is both risk averse and displays PORU, these two forces magnify each other. By varying the parameter n , we change the frequency at which DM updates information. Our next result demonstrates that high frequency of information updates (sufficiently large value of n) alone 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. For purposes of clarity, we state our result in terms of Gul's disappointment aversion preferences (see section 4).

Proposition 6: *Consider disappointment aversion preferences with some $\phi : X \rightarrow \mathbb{R}$ and $\beta > 0$. For any $\varepsilon > 0$ and for any lottery $p = \sum_{j=1}^m p(x_j) \delta_{x_j}$, there exists $T < \infty$ and a multi-stage lottery $Q^T \in P(p)$ such that $V(Q^T) < \min_{x_j \in S(p)} \phi(x_j) + \varepsilon$.*

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 i^{th} toss is “heads” and $z_i = 0$ if it is “tails”. Let the terminal nodes of the n -stage lottery be:

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

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,

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

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.

5.4. Application, an insurance problem

Understanding the magnifying effect, 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. An 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.”

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 \hat{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 $k = 0, \dots, n$. Applying Gul’s formula, one obtains:

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

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,²¹ we have $1 - p < 1 - \frac{V_{\beta,n}(\hat{p})}{n} < z$, meaning that DM would not buy insurance at all *if he could avoid being aware of the gradual resolution of uncertainty*.²² 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.

²¹The condition is: $1 + \beta < \min \left\{ \frac{p^n}{p^n + n(1-p) - 1}, \frac{p^n z}{(1-z)(1-p^n) - p(1-p^{n-1}) - 1} \right\}$.

²²Nayyar [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 .

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

Proposition 7: *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 $\lim_{\beta \rightarrow \infty} \text{grp}(\beta|p,n) = 0$
- (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.

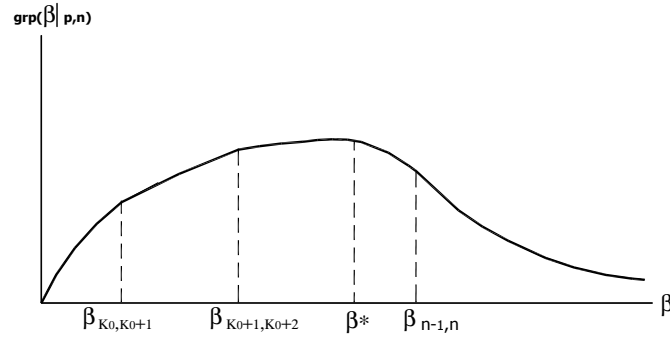


Figure 2: $\text{grp}(\beta|p,n)$. $\beta_{k,k+1}$ is the value of β where $h(\beta|p,n)$ decreases from $(n-k)$ to $(n-(k+1))$. $\text{grp}(\beta|p,n)$ is non-differentiable in each such $\beta_{k,k+1}$. k_0 is the smallest natural number that solves: $\max_{k' > n(1-p)} \frac{n-k'}{n}$

Recall that in Gul's model, the sign of the parameter β determines whether preferences display PORU or PGRU. In its original context, greater β implies greater disappointment aversion (as well as greater risk aversion). PORU can be interpreted as dynamic disappointment aversion. As suggested by Palacios-Huerta [1999], being exposed to the resolution process bears the risk of perceiving intermediate outcomes as disappointing or elating. Individuals who are more sensitive to disappointments suffer from getting partial information and, therefore, prefer all uncertainty to be resolved in a single point in time. Under this interpretation, it seems intuitive to expect the gradual resolution premium to be an increas-

ing function of β . This intuition is wrong and, in fact, item (ii) remains valid independent of the decision problem under consideration. 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 the value of the worst sub-lottery that by itself approaches 0. Since $\text{grp}(\beta | p, n)$ is a continuous function and is strictly positive on the positive reals, there must exist a finite β , denoted β^* in figure 2, in which $\text{grp}(\beta | p, n)$ 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 $\text{grp}(\beta | p, n)$ is single-peaked. Behaviorally speaking, moderately disappointment-averse individuals are more inclined to pay a higher premium than individuals who are either approximately disappointment-indifferent or extremely disappointment-averse.

6. PORU, “loss aversion with narrow framing” and the final-wealth hypothesis

Loss aversion with narrow framing is a combination of two motives: loss aversion, that is, people’s tendency to be more sensitive to losses than to gains, and a dynamic aggregation rule, narrow framing, that argues that when making a series of choices, individuals “bracket” them by making each choice in isolation (an extensive review of this approach is given in Barberis and Huang [2007]). When applied to behavior in financial markets, narrow framing means that individuals evaluate long-term investments based on 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 [2007] 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.

Our model can be used to address similar phenomena. The combination of a specific form of 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 [2008] 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 [2008] 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);²³ 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. Our interpretation is that the value of a lottery depends not only on its uncertainty, but also on the way this uncertainty is resolved over time. In particular, if most uncertainty resolves gradually and is evaluated frequently enough, then it generally cannot be compounded into a single lottery. Our model permits first-order risk aversion over each realized gamble (as, for example, in the case of Gul’s model discussed in the previous section), and hence a Rabin-type critique does not apply. In other words, the mere existence of other risks is not enough to eliminate the effect of first-order risk aversion. 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.²⁴ We, on the other hand, maintain the assumption that terminal wealth matters, and identify narrow framing as a parameter that measures the frequency with which DM evaluates lotteries.²⁵ 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 6 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, provided that the

²³First 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]).

²⁴Kőszegi and Rabin [2006, 2008], and Matthey [2008] offer models in which the reference point is determined endogenously.

²⁵There might be circumstances in which it would be difficult to relate the parameter n to any observable physical lottery. For example, suppose there is a flow of information that DM can privately decide how often to observe. In such situations, n can be interpreted as a subjective temporal effect, that is, as the resolution sensitivity of an individual. When we write an n -stage lottery, the implicit assumption is that this is the multi-stage lottery as perceived by DM and that he does not further compound risks. The interpretation of n as a preference parameter is perhaps more compelling if it is accompanied by a formal modeling of the information flow and the actions available to DM, which we do not pursue in this paper.

parameter n is sufficiently large. Second, we set aside the question of why individuals are sensitive to the way uncertainty is resolved (i.e. why they narrow frame),²⁶ and construct a model that reveals the (context independent) behavioral implications of such considerations.

7. Appendix

7.1. 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 \triangleright_p , "predecessor of", be a partial order on T with $x \triangleright_p y$ if x precedes y . For any node $t \in T$, let $PRE(t) = \{x : x \triangleright_p t\}$ 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 $t \triangleright_{ip} t'$, if $t \in PRE(t')$ and $\nexists t'' \in PRE(t')$ such that $t \in PRE(t'')$. An initial node is any $t \in T$ with $PRE(t) = \emptyset$. A pair (T, \triangleright_p) is a tree if it has a single initial node, and if for all $t \in T$, $PRE(t)$ is totally ordered by \triangleright_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, $s \triangleright_{is} t \iff t \triangleright_{ip} s$. Let $F(t) = \{x : x \triangleright_{is} t\}$. Let $(g_t)_{t \in T}$ be a collection of probability distributions, one for each node, over $F(t)$. If $F(z) = \emptyset$, 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 $k \in \{1, 2, \dots, n\}$, 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 $l \in \{1, 2, \dots, n\}$, 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 $l, l' \in \{1, 2, \dots, n\}$, $c \sim_n \Gamma_q^{l'}$.

$A2' : \text{Fix } t^* \in T^n. \text{ Suppose that for all } t \in T / \{t^*\}, F(t) \text{ is the same in both } Q^n \text{ and } Q^{n'}. \text{ If}$

²⁶Barberis 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.

Q^n yields the lottery q in t^* and $Q^{n'}$ yields the lottery q' in t^* , then $Q^n \succeq_n Q^{n'} \iff \Gamma_q^n \succeq_n \Gamma_{q'}^n$.

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 $k = n-1, n-2, \dots, 1$ and for all $t \in T^k$, let

$$W^{n+1-k}(Q^{n+1-k}(t)) = V\left(\sum_{s \in F(t)} g_t(s) \delta_{c(Q^{n+1-(k+1)}(s))}\right)$$

where $c(Q^l(s)) \in X$ is the certainty equivalent of $Q^l(s)$.

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 $n+1-k$, $k \in \{2, 3, \dots, n-1\}$ that originates from some $t^* \in T^k$. That is, 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 $p = \sum_{s \in F(t^*)} g_{t^*}(s) \delta_{c(Q^{n+1-(k+1)}(s))}$. Define the set $\mathcal{P}(p)$ just as in section 5.

Definition 6: \succeq_n display PORU if for all $Q^n, Q^{n'} \in \mathcal{L}^n$ and $p \in \mathcal{L}^1$ as described above, $W^{n+1-k}(Q_{Q^n}^{n+1-k}(t^*)) = V(p)$ and $W^{n+1-k}(Q_{Q^{n'}}^{n+1-k}(t^*)) = W(Q^2)$ for some $Q^2 \in \mathcal{P}(p)$ imply $Q^n \succeq_n Q^{n'}$.

Proposition 1': under $A0, A1'$ and $A2'$, Proposition 1 remains intact.

For brevity, we omit the detailed proof. It simply involved a repeated use of $A0, A1'$ and $A2'$ to transform the problem into the framework of section 2.

7.2. Proofs

Proof of Proposition 2

Since any temporal lottery corresponds to some information environment in which $u(a, s) = v(s) \in X$ for all $a \in A$, showing that (i) is necessary for (ii) is immediate. To show sufficiency, fix an information environment $\Omega = \{S, J, A, (p_s)_{s \in S}, u\}$. Let Q and p^j be two intermediate lotteries, where p^j assigns probability $p(s|j)$ to the outcome $u(a(s), s)$, and the compound lottery Q assigns probability $\alpha_j(\pi)$ to p^j , that is, $Q = \sum_{j \in J} \alpha_j(\pi) D_{\delta_{u(a(s), s)}}$. 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 $V(\pi) \leq V(Q)$.

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 system, I . Thus by (i) we have

that $V(I) \geq V(Q)$. Combining the two inequalities establishes the result.

Similarly, it is obvious that PGRU is necessary for ϕ being the least valuable information system. To show sufficiency, let $\underline{a} = \arg \max_a V(\sum_{s \in S} p_s \delta_{u(a,s)})$. Let Q and p^j be two intermediate lotteries, where p^j assigns probability $p(s|j)$ to the outcome $u(\underline{a}, s)$, and the compound lottery Q assigns probability $\alpha_j(\pi)$ to p^j , that is, $Q = \sum_{j \in J} \alpha_j(\pi) D_{\delta_{u(\underline{a}, s)}}$. By definition, $V(p^j) \leq V(p^{j*})$ for all j , and therefore, by monotonicity, $V(Q) \leq V(\pi)$.

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 $V(\phi) \leq V(Q)$. Combining the two inequalities establishes the result. ■

Proof of Lemma 2: Suppose not. Then there exist $p, q \in \mathcal{L}^1$ and $\alpha \in (0, 1)$ such that

$$\begin{aligned} V(\alpha D_p + (1 - \alpha) D_q) &= V(\alpha \delta_{c(p)} + (1 - \alpha) \delta_{c(q)}) \geq \\ \min \{V(\delta_{c(p)}), V(\delta_{c(q)})\} &> V(\delta_{c(\alpha p + (1 - \alpha) q)}) = V(\alpha p + (1 - \alpha) q) \end{aligned}$$

where the weak inequality is implied by monotonicity. Contradiction. ||

Proof of Lemma 3 (i): By monotonicity and continuity, there exists $q = (q, (1 - q)) \in \mathcal{I}_{((0,0))}$. By applying NCI twice, $q = \beta q + (1 - \beta) q \succeq \beta q + (1 - \beta) (0, 0) \succeq \beta (0, 0) + (1 - \beta) (0, 0) = (0, 0)$ for all $\beta \in [0, 1]$. Since $q \in \mathcal{I}_{((0,0))}$, the result follows.

(ii): Suppose not. Let q' be a lottery such that $\mu(\mathcal{I}_{((0,0))}) < \mu^+(q')$. Take $p \in \mathcal{I}_{((0,0))}$ and look at the triangle with vertices $(0, 0)$, p , q' . Using the triangle proportional sides theorem, for α sufficiently close to 1 we have $\alpha q' + (1 - \alpha) (0, 0) \succ \alpha q' + (1 - \alpha) p$. Contradiction. ||

Proof of Proposition 3:

(i) There is no preference relation of the rank-dependent utility class that satisfies PORU.²⁷ Fix $x_3 > x_2 > x_1$. By the RDU formula (page 15), the slope of indifference curves over Δ is given by

$$\mu(p) = \frac{g'(p_1) [u(x_2) - u(x_1)]}{g'(1 - p_3) [u(x_3) - u(x_2)]} := c \frac{g'(p_1)}{g'(1 - p_3)}$$

Note that along the hypotenuse, $H = \{\alpha(1, 0) + (1 - \alpha)(0, 1) \mid \alpha \in [0, 1]\}$ the slope is constant and equal c . At $(0, 0)$, the slope is $c \frac{g'(0)}{g'(1)}$. By Lemma 3, $\frac{g'(0)}{g'(1)} = 1$ and for all $\alpha, \beta < \frac{1}{1+c}$, $\mu(\alpha, c\beta) = \frac{c^2}{\mu(\beta, c\alpha)}$. Therefore, by Lemma 2, $\mu(\alpha, c\beta) = \mu(\beta, c\alpha) = c$. Since $\mu(\frac{1}{1+c}, \beta) = c$ for all $\beta \in [0, \frac{c}{1+c}]$, we use Lemma 2 to show that for all $\alpha \in [\frac{1}{1+c}, \frac{1+2c}{(1+c)^2}]$,

²⁷For simplicity we show the proof for the case that g is differentiable. Similar arguments apply to the case that g is only continuous.

$\beta \in [-\frac{c}{1+c} + c\alpha, 1 - \alpha]$ implies that $\mu(\alpha, \beta) = c$. Continue in this fashion, using Lemma 2 repeatedly, to show that for any lottery $p \in \Delta$, $\mu(p) = c$. Therefore, $g(p) = kp$ and since $g(1) = 1$, we have $k = 1$ so $g(p) = p$ and preferences are expected utility.

(ii) There is no preference relation of the quadratic utility class that satisfies PORU.

Fix $x_3 > x_2 > x_1$. By the quadratic utility formula, $\mu(p)$ equals

$$\frac{p_1 [\varphi(x_1, x_2) - \varphi(x_1, x_1)] + p_3 [\varphi(x_2, x_3) - \varphi(x_1, x_3)] + (1 - p_1 - p_3) [\varphi(x_2, x_2) - \varphi(x_1, x_2)]}{p_1 [\varphi(x_1, x_3) - \varphi(x_1, x_2)] + p_3 [\varphi(x_3, x_3) - \varphi(x_2, x_3)] + (1 - p_1 - p_3) [\varphi(x_2, x_3) - \varphi(x_2, x_2)]}$$

Note that if $\mu(m, 1 - m) = \mu(x, 1 - x) = k$, then for all $\alpha \in [0, 1]$,

$$\mu(\alpha m + (1 - \alpha)x, \alpha(1 - m) + (1 - \alpha)(1 - x)) = k$$

Lotteries p and q lie on the same expansion path if there is a common sub-gradient to the indifference curves at p and q . Chew *et al.* [1991] show that for any quadratic utility, all expansion paths are straight lines and perspective, that is, they have a common point of intersection, which could be infinity if they are parallel lines. An implication of this projective property is that for all $m \in (0, 1)$ there exists either (i) $x \in (0, 1)$ such that $\mu^+(m, 0) = \mu^+(0, x)$ or (ii) $y \in (0, 1)$ such that $\mu^+(m, 0) = \mu(y, 1 - y)$.²⁸ For case (i), let $\alpha_{m,x}^* \in (0, 1)$ solves $\alpha(m, 0) + (1 - \alpha)(0, x) \in \mathcal{I}_{((0,0))}$. By Lemmas 1 and 2,

$$\mu^+(0, 0) \leq \mu(\alpha_{m,x}^* m, (1 - \alpha_{m,x}^*) x) = \mu^+(0, x) \leq \mu^+(0, 0)$$

and similarly for case (ii). Therefore all indifference curves are linear and parallel, hence preferences are expected utility. ■

Proof of Proposition 4

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 $f_p(v) = \sum_i \alpha_i f_{p^i}(v)$. Let $\mathcal{P}(p) := \{Q \mid \rho(Q) = p\}$. 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 $V(p) \geq V(Q) \forall Q \in \mathcal{P}(p)$ is equivalent to showing that $f_p(V(Q)) \geq 0 \forall Q \in \mathcal{P}(p)$. 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:

²⁸The existence of either (i) or (ii) can be verified directly without relying on the projective property. The details can be provided upon request.

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

Claim 1:

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

iff

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

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

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

so $V(p) < V(Q)$.

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

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

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 $V(p\delta_{x_3} + (1-p)\delta_{x_1}) := V(p) = pu(x_3, V(p)) + (1-p)u(x_1, V(p)) = u(x_2, V(p))$. 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 $|S(p)| = n-1$ that belongs to an indifference set $I_v := \{p' : \sum_x u(x, v) p'_x = v\}$ in a $(n-1)$ -dimensional unit simplex $\Delta(n)$. Assume further that for some $x_v \in (w, b)$ with $x_v \notin S(p)$, $\langle 1, \delta_{x_v} \rangle \in I_v$ ²⁹. By monotonicity and continuity,³⁰ 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 way 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 $\alpha_1, \dots, \alpha_J$ 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, $V(q^j, v') \geq u(x_v, v')$ for all $j \in J$ and for all $v' \in V(\Delta(n))$ and therefore also

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

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

Proof: see Proposition 1.

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

²⁹The analysis would be the same, though with messier notations, even if $|S(p)| = n$, i.e., if $x \in S(p)$.

³⁰These two assumptions guarantee that no indifference set terminates in the relative interior of any $k \leq n-1$ dimensional unit simplex.

Proof of Proposition 5

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 $\bar{v} \in (0, 1)$. By monotonicity and continuity there exists $x(\bar{v}) \in (w, b)$ such that $\bar{v} = V(\delta_{x(\bar{v})})$. Take any $x > x(\bar{v})$ and note that $\mu(V|x, x(\bar{v}), w) = \left[\frac{u(x(\bar{v}), v)}{u(x, v) - u(x(\bar{v}), v)} \right]$, 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 $\bar{v} \in (0, V(\delta_x))$, Proposition 4 implies that $\mu(V|x, x(\bar{v}), w)$ is maximized at $v = \bar{v}$. A necessary condition is:

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

Or,³¹ using $\bar{v} = u(x(\bar{v}), \bar{v})$ and denote by u_i the partial derivative of u with respect to its i^{th} argument,

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

Note that by continuity and monotonicity of $u(x, v)$ in its first argument, for all $x \in (x(\bar{v}), b)$ there exists $p \in (0, 1)$ such that $p\delta_w + (1-p)\delta_x \sim_1 \delta_{x(\bar{v})}$, or $u(x, \bar{v})(1-p) = u(x(\bar{v}), \bar{v}) = \bar{v}$. Therefore, and using again Proposition 4, (1) is an identity for $x \in (x(\bar{v}), b)$, so we can take the partial derivative of both sides with respect to x and maintain equality. We get:

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

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

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

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

³¹second order conditions would be :

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

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

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

Note that the term

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

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

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

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 $\bar{v} \in (0, 1)$ and the corresponding $x(\bar{v}) \in (w, b)$ (with $\bar{v} = u(x(\bar{v}), \bar{v})$). Take any $x < x(\bar{v})$ and note that $\hat{\mu}(V|b, x(\bar{v}), x) = \left[\frac{u(x(\bar{v}), v) - u(x, v)}{1 - u(x(\bar{v}), v)} \right]$, 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 $\bar{v} \in (V(\delta_x), b)$, by using Proposition 4 we have:

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

or,

$$(u_2(x(\bar{v}), \bar{v}) - u_2(x, \bar{v})) [1 - \bar{v}] = -u_2(x(\bar{v}), \bar{v}) [\bar{v} - u(x, \bar{v})] \quad (2)$$

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

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

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

Since \bar{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

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

which is again well defined since

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

and

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

is finite, and hence the whole integral is finite.

So far we have:

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

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:

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

We further require:

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

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

(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

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

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:

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

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

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

Substituting (7) into (6) we have:

$$u(x, v) - v = [z(x) - z(x(v))] m(v) \quad (8)$$

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

$$v - z(x(v)) m(v) = 0 \quad (9)$$

and

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

or

$$1 = m(v) + v - z(x(v)) m(v) = m(v) \quad (10)$$

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 Proposition 6

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 compound 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:

$$z_i = \begin{cases} 1 & [\text{“success”}] \\ 0 & [\text{“failure”}] \end{cases} \quad i = 1, 2, 3, \dots$$

The terminal nodes are:

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

Claim:

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

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,

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

Therefore

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

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 $V(Q^{T_1}) \in (\phi(y), \phi(y) + \frac{\varepsilon}{2})$.

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 $V(Q^{T_1+T_2}) \in (\phi(y), \phi(y) + \varepsilon)$. 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 $p^k < \frac{3+4k}{4+4k}$, $k = 1, 2, \dots$, and note that $p^k \rightarrow 1$. \parallel

Now take any finite lottery $\sum_{j=1}^m p_j \delta_{x_j}$ and order its prizes as $x_1 < x_2 < \dots < x_m$. 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, \dots, x_m with a value arbitrarily close to $\phi(x_2)$. Conclude by mixing it with x_1 and repeat the construction above. \blacksquare

Proof of Proposition 7

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

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

The denominator of $\Delta V^{(k)}(\beta | p, n)$ 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 immediate. 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 ($V(\hat{p})$) is affected in two ways when β increases: First, given a threshold $h(\beta | p, n)$, the value is (smoothly) 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 $\frac{\sum_{k=1}^n \binom{n}{k} p^k (1-p)^{n-k} k}{1 + \beta(1-p)^n} \xrightarrow[\beta \rightarrow \infty]{} 0$.

To show the existence of β^* (item (iii)), pick $\beta' > 0$ such that $\text{grp}(\beta' | p, n) = \epsilon > 0$. Since $\lim_{\beta \rightarrow \infty} \text{grp}(\beta | p, n) = 0$, there exists $\bar{\beta} := \max \{ \beta \mid \text{grp}(\beta | p, n) = \frac{\epsilon}{2} \}$ and $\bar{\beta} < \infty$. Thus

$\text{grp}(\beta | p, n)$ is a continuous function on the compact interval $[0, \bar{\beta}]$, and hence achieves its maximum on this domain. For single-peakness, we have the following two claims:

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

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

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

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 $\frac{\partial}{\partial \beta} \Delta V^{(k)}(\beta | p, n)$ 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

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

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 $\frac{\partial}{\partial \beta} \Delta V^{(k)}(\beta | p, n)$ has no positive roots, and otherwise it has exactly one positive root. ||

Note that $\Delta V(\beta | p, n)$ is a continuous function that is not differentiable in the points where $h(\beta | p, n)$ changes. For $k = 2, 3, \dots, n-1$, 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: $\lim_{\beta \rightarrow -\beta_{k,k+1}} \frac{\partial}{\partial \beta} \Delta V^{(k)}(\beta | p, n) > \lim_{\beta \rightarrow +\beta_{k,k+1}} \frac{\partial}{\partial \beta} \Delta V^{(k+1)}(\beta | p, n)$

Proof: Apart from at $\beta = 0$, where $\Delta V^{(k)}(0 | p, n) = \Delta V^{(k+1)}(0 | p, n) = 0$, 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 $\frac{\partial}{\partial \beta} \Delta V^{(k)}(0 | p, n) < \frac{\partial}{\partial \beta} \Delta V^{(k+1)}(0 | p, n)$, since this implies that at $\beta_{k,k+1}$, $\Delta V^{(k+1)}(\beta | p, n)$ crosses $\Delta V^{(k)}(\beta | p, n)$ from above. Now:

$$\begin{aligned} \frac{\partial}{\partial \beta} \Delta V^{(k)}(0 | p, n) &= np \frac{(p^k - pp^k - p^n \binom{n-1}{n-k} (1-p)^k)}{p^k} \text{ and} \\ \frac{\partial}{\partial \beta} \Delta V^{(k+1)}(0 | p, n) &= np \frac{(p^{k+1} - pp^{k+1} - p^n \binom{n-1}{n-k-1} (1-p)^{k+1})}{p^{k+1}} \end{aligned}$$

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

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:

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

and

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

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

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

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

so claim 2 follows as well.

$k = n$:

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

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

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

which is clearly single peaked on $(0, \infty)$ (claim 1), and, again by similar calculations:

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

which is claim 2.

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

References

- [1] Allais, Maurice. 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, Andrew, Geert Bekaert, and Jun Liu. 2005. “Why Stocks May Disappoint.” *Journal of Financial Economics*, 76: 471-508.
- [3] Artstein-Avidan, Shiri, and David Dillenberger. 2006. “Don’t tell me anything until you know for sure.” Unpublished.
- [4] Barberis, Nicholas, and Ming Huang. 2007. “The Loss Aversion / Narrow Framing Approach to the Equity Premium Puzzle.” *Handbook of the Equity Risk Premium*, Vol. 1, Rajnish Mehra ed., Elsevier Science & Technology Books.
- [5] Barberis, Nicholas, Ming Huang, and Richard H. Thaler. 2006. “Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case for Narrow Framing.” *American Economic Review*, 96: 1069-1090.
- [6] Bellemare, Charles, Michaela Krause, Sabine Kröger, and Chendi Zhang. 2005. “Myopic Loss Aversion: Information Feedback vs. Investment Flexibility.” *Economics Letters*, 87: 319-324.

- [7] Benartzi, Shlomo and Richard Thaler. 1995. "Myopic loss aversion and the equity premium puzzle." *Quarterly Journal of Economics*, 110: 73–92.
- [8] Blackwell, David. 1953. "Equivalent Comparison of Experiments." *Annals of Mathematics and Statistics*, 24: 265-272.
- [9] Camerer, Colin. 1995. "Individual Decision Making." In: *The Handbook of Experimental Economics*, John H. Kagel and Alvin E. Roth, editors, Princeton University Press.
- [10] Chew, Soo Hong. 1983. "A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory: Resolving the Allais paradox." *Econometrica*, 51, 1065-1092
- [11] Chew, Soo Hong. 1989. "Axiomatic Utility Theories with the Betweenness Property." *Annals of Operations Research*, 19: 273-298.
- [12] Chew, Soo Hong, and Larry G. Epstein. 1989. "The Structure of Preferences and Attitudes Towards the Timing of The Resolution of Uncertainty." *International Economic Review*, 30: 103-117.
- [13] Chew, Soo Hong, Larry G. Epstein, and Uzi Segal. 1991. "Mixture Symmetry and Quadratic Utility." *Econometrica* 59, 139-164.
- [14] Conlisk, John. 1989. "Three Variations on the Allais Example." *American Economic Review*, 79: 392-407.
- [15] Dekel, Eddie. 1986. "An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom." *Journal of Economic Theory*, 40: 304-318.
- [16] Epstein, Larry G. 1993. "Behavior under risk: Recent developments in theory and applications." In: Laffont, J. J. (ed.) *Advances in Economic Theory*. Cambridge: University Press.
- [17] Epstein, Larry G., and Stanley E. Zin. 1989. "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: a Theoretical Framework." *Econometrica*, 57: 937-969.
- [18] Gollier, Christian. 2001. "The Economics of Risk and Time." The MIT Press, 0-262-07215-7.
- [19] Gneezy, Uri, and Jan Potters. 1997. "An experiment on risk taking and evaluation periods." *Quarterly Journal of Economics*, 112: 632–645.

- [20] Grant, Simon, Atsushi Kajii, and Ben Polak. 1992. "Many good choice Axioms: When can many-good lotteries be treated as money lotteries?" *Journal of Economic Theory*, vol. 56 (2), 313-337
- [21] Grant, Simon, Atsushi Kajii, and Ben Polak. 1998. "Intrinsic Preference for Information." *Journal of Economic Theory*, 83: 233-259.
- [22] Grant, Simon, Atsushi Kajii, and Ben Polak. 2000. "Temporal Resolution of Uncertainty and Recursive Non-Expected Utility Models." *Econometrica*, 68: 425-434.
- [23] Gul, Faruk. 1991. "A Theory of Disappointment Aversion." *Econometrica*, 59: 667-686.
- [24] Halevy, Yoram. 2007. "Ellsberg Revisited: An Experimental Study." *Econometrica*, 75: 503-536.
- [25] Haigh, Michael S., and John A. List. 2005. "Do professional traders exhibit myopic loss aversion? An experimental analysis." *Journal of Finance*, 60: 523-534.
- [26] Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision Under Risk." *Econometrica*, 47: 263-291.
- [27] Kahneman, Daniel, and Dan Lovallo. 1993. "Timid choices and bold forecasts: A cognitive perspective on risk-taking." *Management Science*, 39, 17-31.
- [28] Knight, Frank. 1921. "*Risk, Uncertainty and Profit*." Boston, Mass: Houghton Mifflin. Reprint, London: London School of Economics, 1946.
- [29] Köszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences.", *Quarterly Journal of Economics*, 4:1133-1165.
- [30] Köszegi, Botond, and Matthew Rabin. 2009. "Reference-Dependent Consumption Plans." *American Economic Review*, 99(3): 909-936.
- [31] Kreps, David M., and Evan L. Porteus. 1978. "Temporal Resolution of Uncertainty and Dynamic Choice Theory." *Econometrica*, 46, 185-200.
- [32] Machina, Mark J. 1982. "Expected Utility' Analysis without the Independence Axiom." *Econometrica*, 50: 277-323.
- [33] Machina, Mark J. 1987. "Choice under uncertainty: Problems solved and unsolved." *Journal of Economic Perspectives*, 1, 121-154.

- [34] Matthey, Astrid. 2008. "Yesterday's Expectation of Tomorrow Determines What You Do Today: The Role of Reference-Dependent Utility from Expectations." Jena Economic Research Paper No. 2008-003. Available at SSRN: <http://ssrn.com/abstract=920997>
- [35] Mehra, Rajnish, and Edward C. Prescott. 1985. "The Equity Premium: A Puzzle." *Journal of Monetary Economics*, 15: 145-161.
- [36] Nayyar, Shivani. 2004. "Disappointment Aversion and the Insurance Trap." Unpublished.
- [37] Nehring, Klaus. 2005. "Notes on Expected Utility with Psychological Consequences: The Logic of Betweenness-Based Risk Preferences." Paper presented at Risk, Uncertainty and Decision 2005 Conference, Heidelberger Akademie der Wissenschaften, Heidelberg.
- [38] Neilson, William S. 1992. "A Mixed Fan Hypothesis and its Implications for Behavior Towards Risk." *Journal of Economic Behavior and Organization*, 19: 197-211.
- [39] Palacios-Huerta, Ignacio. 1999. "The Aversion to the Sequential Resolution of Uncertainty." *Journal of Risk and Uncertainty*, 18: 249-269.
- [40] Quiggin, John. 1982. "A Theory of Anticipated Utility." *Journal of Economic Behavior and Organization*, 3: 323-343.
- [41] Rabin, Matthew. 2000. "Risk Aversion and Expected-Utility Theory: A Calibration Theorem." *Econometrica*, 68: 1281-1292.
- [42] Rabin, Matthew, and Richard H. Thaler. 2001. "Anomalies: Risk Aversion." *Journal of Economic Perspectives*, 15: 219-232.
- [43] Read, Daniel, George Loewenstein, and Matthew Rabin. 1999. "Choice bracketing." *Journal of Risk and Uncertainty* 19, 171-197.
- [44] Rockafellar, Tyrrell R. 1970. "Convex Analysis." Princeton University Press, Princeton, NJ. Reprint: 1997.
- [45] Safra, Zvi, and Eyal Sulganik. 1995. "On the Nonexistence of Blackwell's Theorem-Type Results with General Preference Relations." *Journal of Risk and Uncertainty*, 10: 187-201.
- [46] Safra, Zvi, and Uzi Segal. 2002. "On the Economic Meaning of Machina's Fréchet Differentiability Assumption." *Journal of Economic Theory*, 104(2): 450-461.

- [47] Safra, Zvi, and Uzi Segal. 2008. "Calibration Results for Non-Expected Utility Theories." *Econometrica*, 76: 1143 - 1166.
- [48] Schlee, Edward. 1990. "The Value of Information in Anticipated Utility Theory." *Journal of Risk and Uncertainty*, 3: 83–92.
- [49] Schmidt, Ulrich. 1998. "A Measurement of the Certainty Effect." *Journal of Mathematical Psychology*, 42: 32-47.
- [50] Segal, Uzi. 1987. "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach." *International Economic Review*, 28(1): 175–202.
- [51] Segal, Uzi. 1990. "Two-stage Lotteries without the Reduction Axiom." *Econometrica*, 58: 349-377.
- [52] Segal, Uzi, and Avia Spivak. 1990. "First order versus second order risk aversion." *Journal of Economic Theory*, 51: 111-125.
- [53] Tversky, Amos, and Daniel Kahneman. 1981. "The framing of decisions and the psychology of choice." *Science* 211, 453-458.
- [54] Wang, Tan. 1993. " L_P -Fréchet differentiable preference and "local utility" analysis." *Journal of Economic Theory*, 61, 139-159.