

ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

*An International Society for the Advancement of Economic
Theory in its Relation to Statistics and Mathematics*

<http://www.econometricsociety.org/>

Econometrica, Vol. 78, No. 6 (November, 2010), 1973–2004

PREFERENCES FOR ONE-SHOT RESOLUTION OF UNCERTAINTY AND ALLAIS-TYPE BEHAVIOR

DAVID DILLENBERGER

University of Pennsylvania, Philadelphia, PA 19104-6297, U.S.A.

The copyright to this Article is held by the Econometric Society. It may be downloaded, printed and reproduced only for educational or research purposes, including use in course packs. No downloading or copying may be done for any commercial purpose without the explicit permission of the Econometric Society. For such commercial purposes contact the Office of the Econometric Society (contact information may be found at the website <http://www.econometricsociety.org> or in the back cover of *Econometrica*). This statement must be included on all copies of this Article that are made available electronically or in any other format.

PREFERENCES FOR ONE-SHOT RESOLUTION OF UNCERTAINTY AND ALLAIS-TYPE BEHAVIOR

BY DAVID DILLENBERGER¹

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 who has recursive, nonexpected utility preferences over compound lotteries. The decision maker has preferences for one-shot resolution of uncertainty if he always prefers any compound lottery to be resolved in a single stage. We establish an equivalence between dynamic preferences for one-shot resolution of uncertainty and static preferences that are identified with 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.

KEYWORDS: Recursive preferences over compound lotteries, resolution of uncertainty, Allais paradox, 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 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 affect 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 concluded that “information feedback should be the variable of interest for researchers and actors in financial markets alike.”

¹I 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, Klaus Nehring, and Uzi Segal 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, Ehud Lehrer, George Mailath, 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.

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.²

In this paper, we assume that the value of a lottery depends directly on the way the 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 toward risk, and intrinsic preferences for information.

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 with the following two assumptions: time neutrality and recursivity. *Time neutrality* says that the DM does not care about the time in which uncertainty is resolved as long as resolution happens in a single stage. *Recursivity* says that if the 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 the DM's preferences over the domain of two-stage lotteries.

To link behavior in both domains, we introduce the following two properties: the first is dynamic while the second is static.

- *Preferences for one-shot resolution of uncertainty* (PORU). The DM has PORU if he always prefers any two-stage lottery to be resolved in a single stage. PORU implies an intrinsic aversion to receiving partial information. This notion formalizes an idea first raised by Palacios-Huerta (1999).

- *Negative certainty independence* (NCI). NCI states that if the 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 nonlinearly. The restrictions NCI imposes on preferences are just enough to explain commonly observed behavior in the common-ratio version of the Allais paradox with positive outcomes. In particular, NCI allows the von Neumann–Morgenstern (vNM) independence axiom to fail when the certainty effect is present.

²All 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.

Proposition 1 establishes that NCI and PORU are equivalent. On the one hand, numerous replications of the Allais paradox 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 the DM to take intermediate actions that might affect his ultimate payoff. The primitive in such a model is a preference relation over information systems, which is induced from preferences over compound lotteries. [Safrá and Sulganik \(1995\)](#) left open the question of whether there are nonexpected utility preferences for which, when applied recursively, perfect information is always the most valuable information system. Proposition 2 shows that this property, which 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.

We extend our results to preferences over arbitrary n -stage lotteries and show that PORU can be quantified. The gradual resolution premium of any compound lottery is the amount that the DM would pay to replace that lottery with its single-stage counterpart. We demonstrate that, for a broad class of preferences, the gradual resolution premium can be quantitatively important; for any one-stage lottery, there exists a multistage lottery (with the same probability distribution over terminal prizes) whose value is arbitrarily close to that of getting the worst prize for sure.

1.1. *Related Literature*

Confining his attention to binary single-stage lotteries and to preferences from the rank-dependent utility class ([Quiggin \(1982\)](#)), [Segal \(1987, 1990\)](#) discussed sufficient conditions under which the desirability of a two-stage lottery decreases as the two stages become less degenerate. Proposition 3 shows that these conditions cannot be extended to the general case, that is, the combination of rank-dependent utility and PORU implies 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 demonstrated that a 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 discussed numerous applications. The general theory we suggest provides a way to understand which attribute of Gul's preferences accounts for the resulting behavior. It also spells out the extent to which the analysis can be extended beyond Gul's preferences.

[Schmidt \(1998\)](#) developed 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 nondegenerate 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 approach to suggest explanations for several economic "anomalies," such as the equity premium puzzle (Mehra and Prescott (1985)). Barberis, Huang, and Thaler (2006) generalized Benartzi and Thaler's work by assuming that the 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 recursivity 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 we 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) studied a model in which utility additively depends both on 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 identified narrow framing with preference over such fluctuations in beliefs. They also showed 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

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) presented an extensive survey of this approach.

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 [Chew and Epstein \(1989\)](#), among others. [Grant, Kajii, and Polak \(1998, 2000\)](#) connected 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 result. [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. 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 nonzero probability is finite. Let $S(p) = \{x \mid p(x) > 0\}$. 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 $\Gamma = \{D_p \mid p \in \mathcal{L}^1\}$, the set of degenerate lotteries in \mathcal{L}^2 , and $\Lambda = \{Q \in \mathcal{L}^2 \mid Q(p) > 0 \Rightarrow p = \delta_x \text{ for some } x \in X\}$, the set of lotteries in \mathcal{L}^2 , the outcomes 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:

AXIOM A0—More Is Better: For all $x, y \in X$, $x \succeq y \Leftrightarrow D_{\delta_x} \succeq D_{\delta_y}$.

AXIOM A1—Time Neutrality: For all $p \in \mathcal{L}^1$, $D_p \sim \sum_{x \in X} p(x)D_{\delta_x}$.

AXIOM A2—Recursivity: For all $q, p \in \mathcal{L}^1$, all $Q \in \mathcal{L}^2$, and $\lambda \in (0, 1)$,

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

Axiom A0 is a weak monotonicity assumption. By postulating Axiom A1, we assume that the DM does not care about the time in which the uncertainty is resolved as long as it happens in a single stage. Axiom 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 Axioms 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): The relation \succeq satisfies Axioms 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 $P, Q \in \mathcal{L}^2$,

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

⁵Axioms 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.

Note that under Axioms **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 . Last, 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 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 a DM whose concept of uncertainty is multistage 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 Axiom **A1**, $D_{\rho(Q)} \sim \sum_{x \in X} [\sum_{q \in \mathcal{L}^1} Q(q)q(x)]D_{\delta_x}$.

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, the DM prefers to remain fully unaware until 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

⁶Halevy (2007) provided some evidence in favor of PORU. In his paper, subjects were asked to state their reservation prices for four different compound lotteries. The behavior of approximately 60% of his subjects was consistent with Axioms **A0–A2**. Furthermore, approximately 40% of his subjects were classified as having preferences that are consistent with the recursive, non-expected utility model. His results (which are discussed in Section 4.2.1 of his paper) show that within the latter group, 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 ($V3$).

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 (1979) used 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 nonlinearly. Consider the following axiom on \succeq_1 :

AXIOM NCI—Negative Certainty Independence: *For all $p, q, \delta_x \in \mathcal{L}^1$ 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 the 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 nondegenerate 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}

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) gave an extensive survey of the experimental evidence against expected utility, including the "common consequence effect" and "common ratio effect" that are related to the Allais paradox.

⁹Conlisk (1989), for example, replicated 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.

As we mentioned before, NCI does not imply any probabilistic distortion. This observation becomes relevant in experiments similar to the one reported by Conlisk (1989, p. 398), who studied the robustness of Allais-type behavior to boundary effects. Conlisk considered 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 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 Axioms 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 Axioms A0–A2, we have

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

and by repeatedly applying NCI,

$$\begin{aligned} \sum_{i=1}^l Q(q^i)\delta_{c(q^i)} &= Q(q^1)\delta_{c(q^1)} + (1 - Q(q^1))\left(\sum_{i \neq 1} \frac{Q(q^i)}{(1 - Q(q^1))} \delta_{c(q^i)}\right) \\ &\stackrel{\text{(NCI)}}{\succeq_1} Q(q^1)q^1 + (1 - Q(q^1))\left(\sum_{i \neq 1} \frac{Q(q^i)}{(1 - Q(q^1))} \delta_{c(q^i)}\right) \\ &= Q(q^2)\delta_{c(q^2)} + (1 - Q(q^2))\left(\frac{Q(q^1)}{(1 - Q(q^2))}q^1\right. \\ &\quad \left.+ \sum_{i \neq 1,2} \frac{Q(q^i)}{(1 - Q(q^2))} \delta_{c(q^i)}\right) \\ &\stackrel{\text{(NCI)}}{\succeq_1} Q(q^1)q^1 + Q(q^2)q^2 + \sum_{i \neq 1,2} Q(q^i)\delta_{c(q^i)} = \dots \end{aligned}$$

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

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 \preceq_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 $p \preceq_1 \delta_y$ 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

$$\begin{aligned}
 Q &\sim \lambda D_{\delta_{c(p)}} + (1 - \lambda)\sum_x q(x)D_{\delta_x} \\
 &> \sum_x [\lambda p(x) + (1 - \lambda)q(x)]D_{\delta_x} \sim D_{\rho(Q)}
 \end{aligned}$$

which violates PORU.

Q.E.D.

The idea behind Proposition 1 is simple: the second step of the folding-back procedure involves mixing all certainty equivalents of the corresponding second-stage lotteries. Applying NCI repeatedly implies that each certainty equivalent loses relatively more (or gains relatively less) from the mixture than the original lottery that it replaces would.

Proposition 1 ties together two notions that are defined on different domains. The equivalence of PORU and NCI suggests that being averse to the gradual resolution of uncertainty and being prone to Allais-type behavior 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 50 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 Axioms A0–A2. An immediate consequence of Blackwell's (1953) seminal result is that in the standard expected utility class, the 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) showed that if \succeq_1 is of the rank-dependent utility class (Quiggin (1982)), then the value of perfect information will always be nonnegative. 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 the 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 $p(s|j) = \pi(s, j)p_s / \sum_{s' \in S} \pi(s', j)p_{s'}$. 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 $V(\pi) := V(\sum_{j \in J} (\sum_{s \in S} \pi(s, j)p_s) D_{p^{j*}})$ 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 $V(\phi) = \max_{a \in A} V(\sum_{s \in S} p_s \delta_{u(a,s)})$ and that $V(I) = V(\sum_{s \in S} p_s \delta_{u(a(s),s)})$, 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: The relation \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 Axioms A0–A2, then the two statements below are equivalent:*

- (i) *The relation \succeq displays PORU.*
- (ii) *The relation \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 Propositions 1 and 2 we get the following corollary:

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

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, $\Delta\{\delta_{x_1}, \delta_{x_2}, \delta_{x_3}\} := \{p = (p_1, p_3) \mid p_1, p_3 \geq 0, p_1 + p_3 \leq 1\}$, 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 , which corresponds to some triple $x_3 > x_2 > x_1$.

LEMMA 2—Quasiconcavity: *If \succeq_1 satisfies NCI, then V is quasiconcave, 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.*

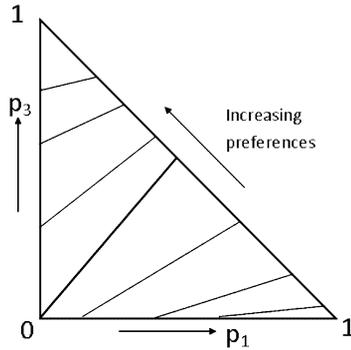


FIGURE 1.—The probability triangle (showing linear indifference curves). The bold indifference curve through the origin demonstrates the steepest middle slope property (Lemma 3).

Let $\mu(p)$ be the slope, relative to the (p_1, p_3) coordinates, of the indifference curve at lottery p . Slope $\mu(p)$ 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 toward risk: the greater the slope, the more (local) risk averse the DM is. Denote by $\mu^+(p)$ the right derivative of the indifference curve at p and denote by $\text{int}(\Delta)$ the interior of Δ . Let $\mathcal{I}_{(p)} := \{q \in \Delta \mid q \sim p\}$.

DEFINITION 3: The function V satisfies the *steepest middle slope property* if the following statements hold:

- (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.*

The applicability of the steepest middle slope property stems from its simplicity. To detect violation of PORU, one need not construct the (potentially complicated) exact choice problem. Rather, it is often sufficient to “examine” the slopes of one-dimensional indifference curves. This, in turn, is a relatively simple task, at least once a utility function is given. Proposition 3 below is based on this observation. The linearity of the indifference curve through the origin is implied by applying NCI twice: $p \succeq_1 \delta_x \Rightarrow p = \alpha p + (1 - \alpha)p \succeq_1 \alpha \delta_x + (1 - \alpha)p \succeq_1 \alpha \delta_x + (1 - \alpha)\delta_x = \delta_x$. Therefore, $p \sim_1 \delta_x \Rightarrow \alpha p + (1 - \alpha)\delta_x \sim_1 \delta_x$.

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

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 (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 rank-dependent utility is

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

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

The functional form for quadratic utility is

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

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

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

Confining his attention to smooth preferences, in the sense that the function V is Fréchet differentiable, Machina (1982) suggested 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

¹²Segal (1990, Section 5) used a different, but equivalent, way to write the functional form for rank-dependent utility, using the transformation $f(p) = 1 - g(1 - p)$. He showed that within this model, if f is convex and its elasticity is nondecreasing, 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 three prizes, 0, 1, and 2 and note that $V(\frac{1}{2}D_{\delta_1} + \frac{1}{2}D_{((\sqrt{2}-1)/\sqrt{2})\delta_0 + (1/\sqrt{2})\delta_2}) = 1 > 0.853 = V(\frac{\sqrt{2}-1}{2\sqrt{2}}\delta_0 + \frac{1}{2}\delta_1 + \frac{1}{2\sqrt{2}}\delta_2)$.

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 subtriangle. In this region, NCI is consistent with fanning out.¹³

The following example demonstrates that the steepest middle slope property is weaker than NCI.¹⁴ For a fixed $n \geq 4$, let $\underline{p} = \frac{1}{n} \sum_{j=1}^n \delta_{((j/n)w + ((n-j)/n)b)}$ and let $\overline{p} = \frac{1}{n} \sum_{j=0}^{n-1} \delta_{((j/n)w + ((n-j)/n)b)}$. Note that \overline{p} first-order stochastically dominates \underline{p} (denoted $\overline{p} \succ_1 \underline{p}$). Let $\mathcal{L}_{|j}^1 \subset \mathcal{L}^1$ be the set of lotteries with j possible outcomes, that is, $\mathcal{L}_{|j}^1 = \{p \in \mathcal{L}^1 : |S(p)| = j\}$. Define $\mathcal{L}^* := \{p \in \mathcal{L}^1 : \overline{p} \succ_1 p \succ_1 \underline{p}\}$. Observe that for $j \leq 3$, $q \in \mathcal{L}_{|j}^1 \Rightarrow q \notin \mathcal{L}^*$.

For any $p \in \mathcal{L}^1$, denote by p^* its cumulative distribution function. Let $d : \mathcal{L}^* \rightarrow [0, 1]$ be defined as $d(q) = (\frac{\|q^* - \overline{p}^*\|}{\|\underline{p}^* - \overline{p}^*\|})^2$, where $\|\cdot\|$ is the L_1 norm. Define $f : \mathcal{L}^* \rightarrow \mathcal{L}^*$ by $f(q) = d(q)\underline{p} + (1 - d(q))\overline{p}$. Note that $f(\underline{p}) = \underline{p}$ and that $f(\overline{p}) = \overline{p}$. Furthermore, if $q, r \in \mathcal{L}^*$ and $q \succ_1 r$, then $d(q) < d(r)$ and $f(q) \succ_1 f(r)$.

Denote by $e(p)$ the expectation of a lottery $p \in \mathcal{L}^1$, that is, $e(p) = \sum_x xp(x)$. Define

$$V(p) = \begin{cases} e(p), & \text{if } p \in \mathcal{L}^1 \setminus \mathcal{L}^*, \\ e(f(p)), & \text{if } p \in \mathcal{L}^*. \end{cases}$$

The function V is continuous, is monotone, and satisfies the steepest middle slope property, but does not satisfy NCI.¹⁵

4.2. Betweenness

For the rest of the section, assume that \succeq_1 is quasiconvex, that is, $\forall p, q \in \mathcal{L}^1$, $V(\alpha p + (1 - \alpha)q) \leq \max\{V(p), V(q)\}$. The conjunction of quasiconvexity with quasiconcavity (Lemma 2) yields the following axiom:

AXIOM A3—Single-Stage Betweenness: *For all $p, q \in \mathcal{L}^1$ and $\alpha \in [0, 1]$, $p \succeq_1 q$ implies $p \succeq_1 \alpha p + (1 - \alpha)q \succeq_1 q$.*

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

¹³The behavioral evidence that supports fanning out is generally weaker in the upper left subtriangle than in the lower right subtriangle (see Camerer (1995)).

¹⁴I thank Danielle Catambay for her help constructing this example.

¹⁵Suppose that $X = [0, 1]$. Let $p' = 0.5\underline{p} + 0.5\overline{p} \in \mathcal{L}^*$ and note that $V(p') = \frac{2n+1}{4n}$. But for γ sufficiently close to 1, $\gamma p' + (1 - \gamma)\delta_{(2n+1)/(4n)} \in \mathcal{L}^*$ and $V(\gamma p' + (1 - \gamma)\delta_{(2n+1)/(4n)}) \neq \frac{2n+1}{4n}$.

PROPOSITION—Chew (1989), Dekel (1986): *The relation \succeq_1 satisfies Axiom A3 if and only if there exists a local utility function $u: X \times [0, 1] \rightarrow [0, 1]$, which is continuous in both arguments, is 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*

$$\sum_x p(x)u(x, v) = v.$$

The next result gives the utility characterization of NCI within the betweenness class of preferences. Let $W(p, v) := \sum_x p(x)u(x, v)$ and denote by $\mathcal{L}^1_{|2}$ the set of all binary lotteries, that is, $\mathcal{L}^1_{|2} = \{p \in \mathcal{L}^1 : |S(p)| = 2\}$.

PROPOSITION 4: *If \succeq_1 satisfies Axiom A3, then the following three statements are equivalent:*

- (i) *The relation \succeq_1 satisfies NCI.*
- (ii) *For all $p \in \mathcal{L}^1$ and $\forall v \in [0, 1]$, $W(p, v) - W(\delta_{c(p)}, v) \geq 0$.*
- (iii) *For all $p \in \mathcal{L}^1_{|2}$ and $\forall v \in [0, 1]$, $W(p, v) - W(\delta_{c(p)}, v) \geq 0$.*

Dekel (1986) provided the following observation: If $W(p, v) = v$ and $W(q, v) = v'$, then $V(p) \geq V(q) \Leftrightarrow v \geq v'$. That is, to compare two lotteries p and q , it is enough to evaluate them at the same value v , which is between $V(p)$ and $V(q)$. The proof of Proposition 4 is based on Dekel’s observation.

The term $W(p, v)$ can be interpreted as the value (expected utility) of p relative to a reference utility level v . Roughly speaking, condition (ii) then implies that risk aversion is maximized at the true lottery value: by definition, $W(p, V(p)) = W(\delta_{c(p)}, V(p)) = V(p)$, whereas the value assigned to p relative to any other v is (weakly) greater than that of $\delta_{c(p)}$. Put differently, condition (ii) is the utility equivalent of the requirement that the conditional certainty equivalent of p (when p is a part of a mixture) is never less than its unconditional certainty equivalent (see Section 2). Condition (iii) is condition (ii) restricted to binary lotteries. It is equivalent to the following weaker version of NCI: $\forall q, \delta_x \in \mathcal{L}^1, p \in \mathcal{L}^1_{|2}$, and $\lambda \in [0, 1]$, $p \succeq_1 \delta_x$ implies $\lambda p + (1 - \lambda)q \succeq_1 \lambda \delta_x + (1 - \lambda)q$. We use the convexity of betweenness indifference sets to show that condition (iii) is also sufficient for condition (ii).¹⁶

4.2.1. Examples

In a dynamic context, expected utility preferences trivially satisfy PORU: a DM with such preferences is just indifferent to the way uncertainty is resolved.

¹⁶In terms of preferences, the steepest middle slope property is equivalent to NCI with the restrictions that $p \in \mathcal{L}^1_{|2}$ and $S(q) \subseteq x \cup S(p)$. Its analogous utility characterization is $\forall p \in \mathcal{L}^1_{|2}$ with two outcomes $\bar{x}_p > \underline{x}_p$, $W(p, v) - W(\delta_{c(p)}, v) \geq 0 \forall v \in (V(\delta_{\bar{x}_p}), V(\delta_{\underline{x}_p}))$. Note that this condition is weaker than condition (iii) in Proposition 4.

Gul (1991) proposed a theory of disappointment aversion. He derived 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 $\beta \in (-1, \infty)$ and $\phi : X \rightarrow \mathbb{R}$ increasing.

For Gul's preferences, the sign of β , the coefficient of disappointment aversion, unambiguously determines whether preferences satisfy PORU (if $\beta \geq 0$) or PGRU (if $-1 < \beta \leq 0$). (See Artstein-Avidan and Dillenberger (2010).)¹⁷

4.2.2. NCI and Differentiability

In most economic applications, it is assumed that individuals' preferences are "smooth." We confine our attention to the betweenness class and suppose that the local utility function $u : X \times [0, 1] \rightarrow [0, 1]$ is sufficiently differentiable with respect to both arguments. In this case, the function V is (continuously) Fréchet differentiable (Wang (1993)).¹⁸ 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.*

To prove Proposition 5, we use the fact that 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 $V(p) = v$,

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

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 $\bar{v} = u(x, \bar{v})$. Combining Lemma 3 with differentiability implies that for any $x > x(\bar{v}) > w$, the derivative with respect to v of

¹⁷The question of whether there is a continuous and monotone function $V : \mathcal{L}^1 \rightarrow \mathbb{R}$, which represents preferences that satisfy NCI but not betweenness, remains open.

¹⁸The 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).

$\mu(v|x, x(\bar{v}), w)$ 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 $\{(x, v) \mid v < u(x, v)\}$ 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, $\{(x, v) \mid v > u(x, v)\}$, $u(x, v) = h^2(v)g^2(x) + f^2(v)$ and $h^2(v) > 0$. Continuity and differentiability then imply that the functional form is equal in both regions, and, 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.¹⁹

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 extension of our setting to \mathcal{L}^n is as follows: equipped with a continuous and increasing function $V: \mathcal{L}^1 \rightarrow \mathbb{R}$, the 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 imply that the DM prefers to replace each compound sublottery 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 multistage lottery Q . We sometimes write Q^n to emphasize that we consider an n -stage lottery.

5.2. Definitions

As before, for any $p \in \mathcal{L}^1$ we denote by $e(p)$ the expectation of p . We say that p second-order stochastically dominates q if for every nondecreasing concave function u , $\sum_x u(x)p(x) \geq \sum_x u(x)q(x)$. The DM is risk averse if $\forall p, q \in \mathcal{L}^1$ with $e(p) = e(q)$, p second-order stochastically dominates 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 that satisfies $\delta_{e(p) - \text{rp}(p)} \sim_1 p$. The risk premium $\text{rp}(p)$ is the amount that the DM would pay to replace p with its expected value. By definition, $\text{rp}(p) \geq 0$ whenever the DM is risk averse.²⁰

¹⁹Neilson (1992) provided sufficient conditions for smooth (in the sense of Proposition 5) betweenness preferences to satisfy the mixed-fan hypothesis (that is, indifference curves fan out in the lower right subtriangle and fan in in the upper left subtriangle). 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.

²⁰Weak risk aversion is defined as 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.

DEFINITION 4: 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 that satisfies $\langle 1, \delta_{c(p) - \text{grp}(Q)} \rangle \sim Q$.

The gradual resolution premium $\text{grp}(Q)$ is the amount that the 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 that satisfies $\langle 1, \delta_{e(p) - \text{rp}(p) - \text{grp}(Q)} \rangle \sim Q$.

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 disappointment aversion preferences (see Section 4.2.1) are risk averse if and only if $\beta \geq 0$ and $\phi : X \rightarrow \mathbb{R}$ is concave (Gul (1991, Theorem 3)). However, for sufficiently small $\beta \geq 0$ and sufficiently convex 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 ,²¹ then the local utility function

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

has the property that $u(\cdot, v)$ is concave for all v . Therefore, the DM is globally risk averse (Dekel (1986, Property 2)), and hence $\text{rp}(p) \geq 0 \forall p \in \mathcal{L}^1$. However, these preferences do not satisfy NCI,²² meaning that there exists $Q \in \mathcal{P}(p)$ with $\text{grp}(Q) < 0$.

5.3. The Magnifying Effect

In the case where the 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 the 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 the DM; a particular splitting of a lottery drives down its value to the value of the worst prize in its support. Although the same result holds for more general preferences, for purposes of clarity, we state Proposition 6 below in terms of biseparable preferences.

DEFINITION 5: The relation \succeq_1 satisfies *biseparability* if there exist an increasing and continuous function π from $[0, 1]$ onto $[0, 1]$ and a mapping

²¹The condition that $\lambda(v)$ is nondecreasing is both necessary and sufficient for u to be a local utility function. See Nehring (2005).

²²Look 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 the steepest middle slope property.

$\phi: X \rightarrow \mathbb{R}$ (unique up to positive affine transformations), such that the restriction of \succeq_1 to $\{\alpha\delta_x + (1 - \alpha)\delta_y: \alpha \in [0, 1], x, y \in X \text{ and } x > y\}$ can be represented by the function

$$V(\alpha\delta_x + (1 - \alpha)\delta_y) = \pi(\alpha)\phi(x) + (1 - \pi(\alpha))\phi(y).$$

Examples of biseparable preferences include any rank-dependent utility (Section 4.1), as well as betweenness preferences that are represented by a local utility of the form

$$u(x, v) = \begin{cases} v + (\phi(x) - \phi(v))^\gamma, & \text{if } x > v, \\ v - \beta(\phi(v) - \phi(x))^\gamma, & \text{if } x \leq v, \end{cases}$$

with $\beta, \gamma > 0$ (Nehring (2005)). We consider biseparable preferences with $\pi(\alpha) < \alpha$.^{23,24}

PROPOSITION 6: *Suppose \succeq_1 satisfies biseparability and that $\pi(\alpha) < \alpha$. Then for any $\varepsilon > 0$ and for any lottery $p = \sum_{j=1}^m p(x_j)\delta_{x_j}$, there exist $T < \infty$ and a multistage lottery $Q^T \in \mathcal{P}(p)$ 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 biseparable preferences as in the premise of Proposition 6, is identical to the value

²³Note that these preferences need not satisfy NCI. For example, in rank-dependent utility, $\pi(p) = 1 - g(1 - p) < p$ if g is concave.

²⁴In the context of decision making under subjective uncertainty (with unknown probabilities), Ghirardato and Maccheroni (2001) argued that the biseparable preferences model is the most general model that achieves a separation between cardinal utility and a unique representation of beliefs.

calculated using recursive expected utility and probability $\pi(0.5) < 0.5$ for heads in each period. Applying the weak law of large numbers yields

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

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.

If most actual risks that individuals face are resolved gradually over time, then these risks cannot be compounded into a single lottery and, therefore, the gradual resolution premium should not be disregarded. The combination of risk aversion and PORU can help 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.²⁵ A formal analysis of this phenomenon will be developed in future work.

APPENDIX

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 that corresponds 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 to be 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, $V(Q) \leq V(\pi)$.

²⁵An example was 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."

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. *Q.E.D.*

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, contradicting PORU. *Q.E.D.*

PROOF OF LEMMA 3: (i) By monotonicity and continuity, there exists $q = (q, (1 - q)) \in \mathcal{I}_{((0,0))}$. By applying NCI twice, $q = \beta q + (1 - \beta)q \geq \beta q + (1 - \beta)(0, 0) \geq \beta(0, 0) + (1 - \beta)(0, 0) = (0, 0)$ 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 $\alpha q' + (1 - \alpha)(0, 0) \succ \alpha q' + (1 - \alpha)p$ a contradiction. *Q.E.D.*

PROOF OF PROPOSITION 3: (i) Suppose that \succeq_1 is of the rank-dependent utility class. Let $\mathcal{L}_{|2}^1$ be the set of all binary lotteries, that is, $\mathcal{L}_{|2}^1 = \{p \in \mathcal{L}^1 : |S(p)| = 2\}$. Consider the following axiom:

AXIOM A*: For all $q \in \mathcal{L}_{|2}^1$, $x \in X$ and $\alpha \in (0, 1)$, $q \sim_1 \delta_x$ implies $\alpha q + (1 - \alpha)\delta_x \sim_1 \delta_x$.

By Lemma 3, NCI implies Axiom A*. Bell and Fishburn (2003, Theorem 1) showed that if \succeq_1 is of the rank-dependent utility class and satisfies Axiom A*, then \succeq_1 is expected utility.

(ii) Suppose that \succeq_1 is of the quadratic utility class. Fix $x_3 > x_2 > x_1$. By the quadratic utility formula, $\mu(p)$ equals

$$\begin{aligned} &(p_1[\varphi(x_1, x_2) - \varphi(x_1, x_1)] + p_3[\varphi(x_2, x_3) - \varphi(x_1, x_3)]) \\ &\quad + (1 - p_1 - p_3)[\varphi(x_2, x_2) - \varphi(x_1, x_2)] \\ &/([p_1[\varphi(x_1, x_3) - \varphi(x_1, x_2)] + p_3[\varphi(x_3, x_3) - \varphi(x_2, x_3)] \\ &\quad + (1 - p_1 - p_3)[\varphi(x_2, x_3) - \varphi(x_2, x_2)]). \end{aligned}$$

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 subgradient to the indifference curves at p and q . Chew, Epstein, and Segal (1991) showed 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 $\mu^+(m, 0) = \mu(y, 1 - y)$. 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 2 and 3,

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

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

PROOF OF PROPOSITION 4: Let $W(q, v) := \sum_x q(x)u(x, v)$.

(i) \Rightarrow (ii) Suppose not. Then there exists a lottery p such that $W(p, v) - W(\delta_{c(p)}, v) < 0$ for some v . Pick $y \in X$ and $\alpha \in (0, 1)$ such that $V(\alpha p + (1 - \alpha)\delta_y) = v$. We have $v < \alpha u(c(p), v) + (1 - \alpha)u(y, v) = W(\alpha\delta_{c(p)} + (1 - \alpha)\delta_y, v)$ or $\alpha\delta_{c(p)} + (1 - \alpha)\delta_y \succ_1 \alpha p + (1 - \alpha)\delta_y$, contradicting NCI.

(ii) \Rightarrow (i) Assume $p \succeq_1 \delta_x$. Then $W(p, V(p)) \geq W(\delta_x, V(p))$. By (ii) and monotonicity, $W(p, v) \geq W(\delta_x, v)$ for all v and, in particular, for $v = V(\lambda p + (1 - \lambda)q)$.²⁶ Therefore, $W(\lambda p + (1 - \lambda)q, V(\lambda p + (1 - \lambda)q)) \geq W(\lambda\delta_x + (1 - \lambda)q, V(\lambda p + (1 - \lambda)q))$, which is equivalent to $\lambda p + (1 - \lambda)q \succeq_1 \lambda\delta_x + (1 - \lambda)q$.

(iii) \Rightarrow (ii) Take a lottery p with $|S(p)| = n - 1$ that belongs to an indifference set $I_v := \{p' : W(p', v) = v\}$ in an $(n - 1)$ -dimensional unit simplex. Assume further that for some $x_v \in (w, b)$ with $x_v \notin S(p)$, $(1, \delta_{x_v}) \in I_v$.²⁷ By monotonicity and continuity,²⁸ 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 convex combinations of two other lotteries with size of support equal to $n - 3$ and that belong to I_v . Continue in the same fashion to get an index set J and a collection of lotteries, $\{q^j\}_{j \in J}$, such that for all $j \in J$, $|S(q^j)| = 2$ and $q^j \in I_v$. Note that by monotonicity, if $y, z \in S(q^j)$ then either $z > x_v > y$ or $y > x_v > z$. By construc-

²⁶If $p \sim \delta_x$, the assertion is evident. Otherwise, we need to find p^* that is both first-order stochastically dominated by p and satisfies $p^* \sim \delta_x$, and to use the monotonicity of $u(\cdot, v)$ with respect to its first argument. By continuity, such p^* exists.

²⁷The analysis would be the same, although with messier notation, even if $|S(p)| = n$, that is, if $x \in S(p)$.

²⁸These two assumptions guarantee that no indifference set terminates in the relative interior of any $k \leq n - 1$ -dimensional unit simplex.

tion, for some $\alpha_1, \dots, \alpha_J$ with $\alpha_j > 0$ and $\sum_j \alpha_j = 1$, $\sum_j \alpha_j q^j = p$. By hypothesis, $W(q^j, v') \geq u(x_v, v')$ for all $j \in J$ and for all v' , and, therefore, also

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

(ii) \Rightarrow (iii) Obvious.

Q.E.D.

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 three 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 $\mu(v|x, x(\bar{v}), w) = [\frac{u(x(\bar{v}), v)}{u(x, v) - u(x(\bar{v}), v)}]$ is continuous and differentiable as a function of v on $[0, V(\delta_x)]$.

Since $\bar{v} \in (0, V(\delta_x))$, Lemma 3 implies that $\mu(v|x, x(\bar{v}), w)$ is maximized at $v = \bar{v}$. A necessary condition is

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

Alternatively,²⁹ using $\bar{v} = u(x(\bar{v}), \bar{v})$ and denoting by u_i the partial derivative of u with respect to its i th argument yields

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

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 again using Lemma 3, (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 .

²⁹Second-order conditions would be

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

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

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

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'Hôpital'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 $(\int_{s=0}^v \frac{u_2(x(s), s)}{s} ds)$ is finite as well.

To uncover $u(x, v)$ on the region $\{(x, v) | v > u(x, v)\}$, again fix 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 $\mu(v|b, x(\bar{v}), x) = [\frac{u(x(\bar{v}), v) - u(x, v)}{1 - u(x(\bar{v}), v)}]$ is continuous and differentiable as a function of v on $[V(\delta_x), b]$.

Since $\bar{v} \in (V(\delta_x), b)$, by using Lemma 3 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

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

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 $\{(x, v) | v > u(x, v)\}$:

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

Its solution is given by

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

which is again well defined since

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

and

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

is finite, and hence the whole integral is finite.

So far we have

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

We add the following restrictions:

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

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

Substituting into (3), we get

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

We add two further requirements:

(iii) Continuity at $x = x(v)$, which 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

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

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

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

and plug 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

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

Substituting (7) into (6), we have

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

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

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

and

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

or

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

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. *Q.E.D.*

PROOF OF PROPOSITION 6: We first show that the result holds for lotteries of the form $\alpha\delta_x + (1 - \alpha)\delta_y$, with $x > y$. There are three cases to consider:

CASE 1— $\alpha = 0.5$: 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

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

The terminal nodes are

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

We claim that $\lim_{n \rightarrow \infty} V(Q^n) = V(\delta_y) = \phi(y)$. To prove this claim, we use the fact that the value of the lottery using recursive biseparable preferences (with $\pi(0.5) < 0.5$) and probability 0.5 for success in each period is equal to the value of the lottery using recursive expected utility and probability $\pi(0.5)$ for success in each period. Since z_i 's are independent and identically distributed

(i.i.d.) random variables, the weak law of large numbers implies

$$\frac{\sum_{i=1}^n z_i}{n} \xrightarrow{p} \pi(0.5) < 0.5$$

or

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

Therefore,

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

CASE 2— $\alpha < 0.5$: Take $Q^{n+1} = \langle 2\alpha, Q^n; 1 - 2\alpha, \delta_y \rangle$, with Q^n as defined above.

CASE 3— $\alpha > 0.5$: Fix $\varepsilon > 0$. Using the construction in Case 1, obtain Q^{T_1} with $V(Q^{T_1}) \in (\phi(y), \phi(y) + \frac{\varepsilon}{2})$. Reconstruct a lottery as above, but replace δ_y with Q^{T_1} in the terminal node. By the same argument, there exist 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 $\alpha < 0.75$. Repeat in the same fashion to show that the assertion is true for $\alpha^k < \frac{3+4k}{4+4k}$, $k = 1, 2, \dots$, and note that $\alpha^k \rightarrow 1$.

Now take any finite lottery $\sum_{j=1}^m \alpha_j \delta_{x_j}$ 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 multistage 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. *Q.E.D.*

REFERENCES

- ALLAIS, M. (1953): "Le Comportement de l'Homme Rationnel Devant le Risque: Critique des Postulats et Axiomes de l'École Américaine," *Econometrica*, 21, 503–546. [1980]

- ARTSTEIN-AVIDAN, S., AND D. DILLENBERGER (2010): "Dynamic Disappointment Aversion: Don't Tell Me Anything Until You Know for Sure," Unpublished Manuscript, University of Pennsylvania; Working Paper 10-025, PIER. [1989]
- BARBERIS, N., AND M. HUANG (2007): "The Loss Aversion/Narrow Framing Approach to the Equity Premium Puzzle," in *Handbook of the Equity Risk Premium*, Vol. 1, ed. by R. Mehra. Amsterdam, The Netherlands: Elsevier. [1976]
- BARBERIS, N., M. HUANG, AND R. H. THALER (2006): "Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case for Narrow Framing," *American Economic Review*, 96, 1069–1090. [1976]
- BELL, D. E., AND P. C. FISHBURN (2003): "Probability Weights in Rank-Dependent Utility With Binary Even-Chance Independence," *Journal of Mathematical Psychology*, 47, 244–258. [1994]
- BELLEMARE, C., M. KRAUSE, S. KRÖGER, AND C. ZHANG (2005): "Myopic Loss Aversion: Information Feedback vs. Investment Flexibility," *Economics Letters*, 87, 319–324. [1973]
- BENARTZI, S., AND R. THALER (1995): "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 110, 73–92. [1976]
- BLACKWELL, D. (1953): "Equivalent Comparison of Experiments," *Annals of Mathematics and Statistics*, 24, 265–272. [1983]
- CAMERER, C. (1995): "Individual Decision Making," in *The Handbook of Experimental Economics*, ed. by J. H. Kagel and A. E. Roth. Princeton, NJ: Princeton University Press. [1980,1987]
- CHEW, S. H. (1989): "Axiomatic Utility Theories With the Betweenness Property," *Annals of Operations Research*, 19, 273–298. [1988]
- CHEW, S. H., AND L. G. EPSTEIN (1989): "The Structure of Preferences and Attitudes Towards the Timing of the Resolution of Uncertainty," *International Economic Review*, 30, 103–117. [1977]
- CHEW, S. H., L. G. EPSTEIN, AND U. SEGAL (1991): "Mixture Symmetry and Quadratic Utility," *Econometrica*, 59, 139–163. [1986,1995]
- CONLISK, J. (1989): "Three Variations on the Allais Example," *American Economic Review*, 79, 392–407. [1980,1981]
- DEKEL, E. (1986): "An Axiomatic Characterization of Preferences Under Uncertainty: Weakening the Independence Axiom," *Journal of Economic Theory*, 40, 304–318. [1988,1991]
- EPSTEIN, L. G., AND S. E. ZIN (1989): "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969. [1977]
- GHIRARDATO, P., AND F. MACCHERONI (2001): "Risk, Ambiguity and the Separation of Utility and Beliefs," *Mathematics of Operations Research*, 26, 864–890. [1992]
- GNEEZY, U., AND J. POTTERS (1997): "An Experiment on Risk Taking and Evaluation Periods," *Quarterly Journal of Economics*, 112, 632–645. [1973]
- GRANT, S., A. KAJI, AND B. POLAK (1992): "Many Good Choice Axioms: When Can Many-Good Lotteries Be Treated as Money Lotteries?" *Journal of Economic Theory*, 56, 313–337. [1978]
- (1998): "Intrinsic Preference for Information," *Journal of Economic Theory*, 83, 233–259. [1977]
- (2000): "Temporal Resolution of Uncertainty and Recursive Non-Expected Utility Models," *Econometrica*, 68, 425–434. [1977]
- GUL, F. (1991): "A Theory of Disappointment Aversion," *Econometrica*, 59, 667–686. [1975,1988, 1989,1991]
- HALEVY, Y. (2007): "Ellsberg Revisited: An Experimental Study," *Econometrica*, 75, 503–536. [1979]
- HAIGH, M. S., AND J. A. LIST (2005): "Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis," *Journal of Finance*, 60, 523–534. [1973]
- KAHNEMAN, D., AND D. LOVALLO (1993): "Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk-Taking," *Management Science*, 39, 17–31. [1976]
- KAHNEMAN, D., AND A. TVERSKY (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263–291. [1974,1976,1980]

- KNIGHT, F. (1921): *Risk, Uncertainty and Profit*. Boston, MA: Houghton Mifflin. [1974]
- KÖSZEGI, B., AND M. RABIN (2009): "Reference-Dependent Consumption Plans," *American Economic Review*, 99, 909–936. [1976]
- KREPS, D. M., AND E. L. PORTEUS (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46, 185–200. [1977]
- MACHINA, M. J. (1982): "Expected Utility' Analysis Without the Independence Axiom," *Econometrica*, 50, 277–323. [1985, 1986, 1989]
- (1987): "Choice Under Uncertainty: Problems Solved and Unsolved," *Journal of Economic Perspectives*, 1, 121–154. [1980]
- MEHRA, R., AND E. C. PRESCOTT (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15, 145–161. [1976]
- NEHRING, K. (2005): "Notes on Expected Utility With Psychological Consequences: The Logic of Betweenness-Based Risk Preferences," Paper presented at *Risk, Uncertainty and Decision 2005 Conference*, Heidelberg Akademie der Wissenschaften, Heidelberg. [1991, 1992]
- NEILSON, W. S. (1992): "A Mixed Fan Hypothesis and Its Implications for Behavior Towards Risk," *Journal of Economic Behavior and Organization*, 19, 197–211. [1989, 1990]
- PALACIOUS-HUERTA, I. (1999): "The Aversion to the Sequential Resolution of Uncertainty," *Journal of Risk and Uncertainty*, 18, 249–269. [1974, 1975]
- QUIGGIN, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3, 323–343. [1975, 1983, 1986]
- READ, D., G. LOEWENSTEIN, AND M. RABIN (1999): "Choice Bracketing," *Journal of Risk and Uncertainty*, 19, 171–197. [1976]
- ROCKAFELLAR, T. R. (1970): *Convex Analysis*. Princeton, NJ: Princeton University Press. [1985]
- SAFRA, Z., AND U. SEGAL (2002): "On the Economic Meaning of Machina's Fréchet Differentiability Assumption," *Journal of Economic Theory*, 104, 450–461. [1989]
- SAFRA, Z., AND E. SULGANIK (1995): "On the Nonexistence of Blackwell's Theorem-Type Results With General Preference Relations," *Journal of Risk and Uncertainty*, 10, 187–201. [1975, 1983]
- SCHLEE, E. (1990): "The Value of Information in Anticipated Utility Theory," *Journal of Risk and Uncertainty*, 3, 83–92. [1983]
- SCHMIDT, U. (1998): "A Measurement of the Certainty Effect," *Journal of Mathematical Psychology*, 42, 32–47. [1975]
- SEGAL, U. (1987): "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach," *International Economic Review*, 28, 175–202. [1975, 1986]
- (1990): "Two-Stage Lotteries Without the Reduction Axiom," *Econometrica*, 58, 349–377. [1974, 1975, 1978, 1986]
- TVERSKY, A., AND D. KAHNEMAN (1981): "The Framing of Decisions and the Psychology of Choice," *Science*, 211, 453–458. [1976]
- WANG, T. (1993): " L_p -Fréchet Differentiable Preference and 'Local Utility' Analysis," *Journal of Economic Theory*, 61, 139–159. [1989]

Dept. of Economics, University of Pennsylvania, 160 McNeil Building, 3718 Locust Walk, Philadelphia, PA 19104-6297, U.S.A.; ddill@sas.upenn.edu.

Manuscript received October, 2008; final revision received April, 2010.