

# Temporal Resolution of Uncertainty and Recursive Non-Expected Utility Models\*

## Abstract

If an agent (weakly) prefers early resolution of uncertainty then the recursive forms of both the most commonly used non-expected utility models, betweenness and rank dependence, almost reduce to Kreps & Porteus's (1978) recursive expected utility.

Simon Grant  
Department of Economics, Faculties  
Australian National University

Atsushi Kajii  
Institute of Policy and Planning Sciences  
University of Tsukuba

Ben Polak  
Economics Department  
Yale University

---

\*We thank Eddie Dekel and Boaz Moselle.

# 1 Introduction

Kreps & Porteus (1978) recursive expected utility model allows an agent to care about the timing of the resolution of uncertainty. For example, an anxious agent may prefer early resolution while a hopeful agent may prefer late. Recursive expected utility achieves this flexibility by relaxing the reduction of compound lottery axiom for temporal lotteries. The model remains tractable thanks to recursivity: preferences today are built up from preferences tomorrow that do not themselves depend on unrealized contingencies.<sup>1</sup> Kreps & Porteus were careful to separate recursivity (“temporal consistency”) from the standard independence or “substitution” axiom for choice over the lotteries in each stage, though they assumed both. In the past two decades, there has been considerable interest in violations of the independence axiom.<sup>2</sup> Recursive non-expected utility models keep the tractability of Kreps & Porteus’s analysis while allowing both for preferences about the timing of resolution of uncertainty and for violations of independence. That is, in evaluating lotteries at each stage, recursive non-expected utility models replace independence by some weaker axiom. Kreps & Porteus’s approach has been extended and then applied to finance and macroeconomics by, for example, Epstein & Zin (1989, 1991).

The two most studied forms of non-expected utility model are the betweenness model (Chew 1989, and Dekel 1986) and the rank dependent model (Quiggin 1982). Both Chew’s (1983) weighted expected utility and Gul’s (1991) disappointment aversion are examples of betweenness models. Yaari’s (1987) dual model is an example of rank dependence. Wakker (1990) has shown that (given probabilistic sophistication and stochastic dominance), Choquet expected utility reduces to rank dependence. Loosely speaking, however, we show that if an agent’s preferences are recursive and the agent always (weakly) prefers early resolution then both the betweenness and rank dependent models almost collapse back to Kreps & Porteus’s model. More precisely, in the case of betweenness, violations of expected utility are only possible in the first stage. For the case of rank dependence, such violations are only possible in the final stage.

In the formal statement of the results, we invoke (weak) risk aversion in the agent’s preferences over the first-stage of multi-stage lotteries. This assumption simplifies the analysis but is inessential. We could replace it by assuming that these preferences are smooth.<sup>3</sup> The results in the current paper do not require smoothness. Also, we state the results for the case of (weak) preference for early resolution but similar results apply if the agent always prefers late resolution.<sup>4</sup> More generally, whether an agent prefers early or late resolution

---

<sup>1</sup> This is sometimes referred to as the fold-back procedure.

<sup>2</sup> For surveys see, for example, Machina (1987) and Sugden (1995).

<sup>3</sup> Gateaux differentiable is sufficient. Strictly speaking, the rank dependent result would now permit deviations from independence in only the first and last stages. See Grant, Kajii & Polak (1997).

<sup>4</sup> Strictly speaking, for the case of betweenness, risk aversion in the first stage would be replaced by either

might depend on the outcomes at stake. At some notational cost, however, the results could be adapted to ‘local’ statements; for example, we could state the results for a particular set of outcomes over which the agent prefers early resolution. At still greater notational cost, we could even localize the results to subsets of simplexes.

Our result could be interpreted as providing support for Kreps & Porteus’s model. Loosely speaking, to escape the recursive expected utility model, either (a) the agent’s within-stage preferences must fail to conform to either the betweenness or rank-dependence models; (b) she must be quite inconsistent in her preferences for early or late resolution; or (c) she must violate recursivity. Certainly, there are non-expected utility models other than betweenness and rank-dependence though they lack some of these models’ attractive features.<sup>5</sup>

We can think of no normative argument in favor of consistently preferring early or late resolution. It seems a desirable property of a model about attitudes toward the timing of resolution, however, that it be flexible enough at least to allow for the possibility that an agent might consistently prefer early or late resolution.

The results could be taken as an argument against restricting attention to recursive models. Machina (1989), for example, has argued against assuming the degree of consequentialism implicit in recursivity. Epstein (1992), however, strongly defends the recursive approach on practical grounds. To our knowledge, all existing papers that drop recursivity (while maintaining dynamic consistency) implicitly assume that the agent does not care about the timing of resolution; that is, they assume reduction even for temporal lotteries.<sup>6</sup> Starmer & Sugden (1991) and Bernasconi (1994) present evidence consistent with recursivity. Conlisk (1989) shows that violations of independence are less frequent and no longer systematic when questions are rephrased in a compound lottery form.

Sarin & Wakker (1996) examine a restriction on recursive preferences over multi-stage decision trees that they call “sequential consistency”. It requires that the agent use a member of the same family of preference not only to evaluate each stage of a temporal lottery but also to evaluate the lotteries induced by each (normal form) strategy for the whole tree. They find that recursivity and structural consistency together with betweenness (respectively, rank dependence) imply that deviations from expected utility are also only possible in the first (respectively, last) stage. Sequential consistency, however, does not restrict an agent to prefer early resolution of uncertainty, and preference for early resolution does not imply sequential consistency. It is an open question why these seemingly quite different restrictions have such similar consequences.

Chew & Epstein (1989) define a premium to measure the degree of preference for early risk loving or smoothness. For rank dependence, now, only the first stage preferences can violate expected utility.

<sup>5</sup> For example, Grant, Kajii & Polak (1997) show that quadratic utility can accommodate preference for early resolution.

<sup>6</sup> See, for example, Machina (1989) and Gul & Lantto (1990). Segal (1990) argues that recursivity is more natural than reduction, even in atemporal lotteries.

resolution in recursive betweenness models. They show that a constant premium implies independence. A constant premium, however, might seem a strong assumption: think of the analogy to a risk premium. Our result shows that it suffices for the premium not to change sign.

## 2 Framework and Result

First, let us first establish some generic notation. For any non-empty closed subset,  $\mathcal{Z}$ , of a metric space, let  $\mathcal{L}_0(\mathcal{Z})$  denote the set of (Borel) probability measures on  $\mathcal{Z}$  with finite support. Notice that the set  $\mathcal{L}_0(\mathcal{Z})$  has a natural linear structure. That is, if  $\mu$  and  $\nu$  are elements of  $\mathcal{L}_0(\mathcal{Z})$  then for any Borel subset  $B$  of  $\mathcal{Z}$  and any  $\alpha$  in  $[0, 1]$ ,  $\alpha\mu + (1 - \alpha)\nu$  is the element of  $\mathcal{L}_0(\mathcal{Z})$  defined by the rule  $(\alpha\mu + (1 - \alpha)\nu)(B) = \alpha\mu(B) + (1 - \alpha)\nu(B)$ . In particular, if  $B$  is the singleton set  $\{\zeta\}$  for some  $\zeta$  in  $\mathcal{Z}$ , then  $\alpha\mu + (1 - \alpha)\nu$  assigns to  $\zeta$  the weighted sum of the probabilities assigned to it by  $\mu$  and  $\nu$ . For each  $\zeta$  in  $\mathcal{Z}$ , let  $\delta_\zeta$  in  $\mathcal{L}_0(\mathcal{Z})$  denote the (degenerate) probability measure that assigns probability one to  $\zeta$ . Using the above rule, for any finite list  $(\zeta_1, \dots, \zeta_M)$  where each  $\zeta_j$  is in  $\mathcal{Z}$ , the convex combination  $\sum_{j=1}^M p_j \delta_{\zeta_j}$  is the element of  $\mathcal{L}_0(\mathcal{Z})$  that assigns to each  $\zeta$  in  $\mathcal{Z}$ , the probability  $\sum_{j:\zeta_j=\zeta} p_j$ . With slight abuse of notation, however, we shall use  $[(\zeta_j, p_j)_{j=1}^M]$  to denote  $\sum_{j=1}^M p_j \delta_{\zeta_j}$ , even though the  $\zeta_j$ 's need not be distinct.

Let  $\mathcal{X}$  be a set of outcomes with generic element  $x$ . In particular, unless otherwise stated, we take  $\mathcal{X}$  be the interval  $[0, 1]$ .<sup>7</sup> With the weak topology,  $\mathcal{L}_0(\mathcal{X})$  is a dense subset of a compact metric space, and so is  $\mathcal{L}_0^2(\mathcal{X}) := \mathcal{L}_0(\mathcal{L}_0(\mathcal{X}))$ . For any integer  $t > 0$ , define  $\mathcal{L}_0^t(\mathcal{X})$  similarly. We will adopt the convention,  $\mathcal{L}_0^0(\mathcal{X}) := \mathcal{X}$ . With slight abuse of terminology, we will refer to elements of  $\mathcal{L}_0(\mathcal{X})$  as one-stage lotteries and, more generally, to elements of  $\mathcal{L}_0^t(\mathcal{X})$  as  $t$ -stage lotteries. Let  $X^t$  denote a generic element of  $\mathcal{L}_0^t(\mathcal{X})$ , so that, in particular, each  $X^0$  is identified with an outcome  $x$  in  $\mathcal{X}$ , and  $X^1$  is identified with a one-stage lottery  $X$  in  $\mathcal{L}_0(\mathcal{X})$ . Following our generic notation, for any integer  $t > 0$ ,  $X^t = [(X_i^{t-1}, q_i)_{i=1}^N]$  is the  $t$ -stage lottery in  $\mathcal{L}_0^t(\mathcal{X})$  that assigns to each  $(t - 1)$ -stage lottery  $X^{t-1}$  in  $\mathcal{L}_0^{t-1}(\mathcal{X})$ , the probability  $\sum_{i: X_i^{t-1} = X^{t-1}} q_i$ . For any integer  $t > 0$ , we shall identify each outcome  $x$  in  $\mathcal{X}$  with the  $t$ -stage lottery that yields  $x$  with probability 1, a lottery that is degenerate in all stages. With slight abuse of notation we will refer to this lottery as  $\delta_x$ .

Suppose that,  $T$  periods from now, an agent will consume a risky outcome from the set  $\mathcal{X}$ . We are interested in the agent's preferences about when the uncertainty concerning this outcome is resolved. One extreme case is that all uncertainty is resolved in the first stage; that is, by the end of the first period. The other extreme is that the agent learns nothing until she gets the outcome at the end of the  $T$ th period. In general, uncertainty about the

---

<sup>7</sup> The analysis can readily be extended to outcome sets that are general compact metric spaces if we assume that all welfare-relevant risk can be characterized as risk over the ranks of outcomes; see Grant, Kajii & Polak (1992a).

outcome could be resolved, partially or completely, in each of the  $T$  periods. In more general problems, there could also be uncertainty about consumption in the intervening periods. It is enough for our purposes, however, to consider the case where there is no such uncertainty and, indeed, it is sufficient to look at the special case where there is no consumption at all before the end. Thus, the objects of choice for our agent are isomorphic to the set of  $T$ -stage lotteries, where each stage denotes a time period when uncertainty could be resolved. For simplicity, we will only consider risk with finite support and finite realizations of uncertainty so we can confine attention to simple lotteries. The extension to general lotteries, however, is straightforward.

Let  $\succeq_T$  be an agent's preference relation over  $T$ -stage lotteries,  $\mathcal{L}_0^T(\mathcal{X})$ . These preferences pertain to an agent at time 0, looking forward at a  $T$ -stage process during which uncertainty about the final  $T$ -period-away outcome will be resolved; that is, these preferences refer to a moment in time before any uncertainty has yet had any chance to be resolved.<sup>8</sup> We will assume throughout that, for all  $T$ ,  $\succeq_T$  is continuous as a relation on  $\mathcal{L}_0^T(\mathcal{X})$ .

To define what it means for preferences to be recursive and what it means for them to exhibit preference for early resolution of uncertainty, it is useful to identify special subsets of  $T$ -stage lotteries where no uncertainty is resolved in early stages. For any  $t$  in  $\{1, \dots, T\}$ , let  $\mathcal{G}_{T,t} \subset \mathcal{L}_0^T(\mathcal{X})$  be the subset of  $\mathcal{L}_0^T(\mathcal{X})$  whose elements are degenerate in the first  $(t-1)$  stages.<sup>9</sup> Notice that  $\mathcal{G}_{T,1} = \mathcal{L}_0^T(\mathcal{X})$ . For any  $t$  in  $\{1, \dots, T\}$ , all elements of  $\mathcal{G}_{T,t}$  are  $T$ -stage lotteries but, since they are degenerate in the first  $(t-1)$  stages, the set  $\mathcal{G}_{T,t}$  is isomorphic to the set  $\mathcal{L}_0^{T+1-t}(\mathcal{X})$ . That is, each  $(T+1-t)$ -stage lottery  $X^{T+1-t} = [(X_i^{T-t}, q_i)_{i=1}^N]$  in  $\mathcal{L}_0^{T+1-t}(\mathcal{X})$  is naturally associated with the  $T$ -stage lottery in  $\mathcal{G}_{T,t}$  for which the total probability of getting each  $(T-t)$ -stage lottery  $X^{T-t}$  in  $\mathcal{L}_0^{T-t}(\mathcal{X})$  is given by  $\sum_{i: X_i^{T-t} = X^{T-t}} q_i$ . We shall denote this isomorphism from  $\mathcal{L}_0^{T+1-t}(\mathcal{X})$  to  $\mathcal{G}_{T,t}$  by  $\Gamma_{T,t}$ .

**Definition** For any finite length of lotteries  $T \geq 2$ , we say that an agent's preference relation over  $T$ -stage lotteries,  $\succeq_T$ , satisfies recursivity if for all  $t$  in  $\{1, \dots, T-1\}$ , and all pairs of  $(T+1-t)$ -stage lotteries of the form  $X^{T+1-t} = [(X_i^{T-t}, q_i)_{i=1}^N]$  and  $Y^{T+1-t} = [(X_1^{T-t}, q_1; \dots; X_{j-1}^{T-t}, q_{j-1}; Y_j^{T-t}, q_j; X_{j+1}^{T-t}, q_{j+1}; \dots; X_N^{T-t}, q_N)]$  in  $\mathcal{L}_0^{T+1-t}(\mathcal{X})$  with  $q_1 > 0$ :  $\Gamma_{T,t}(X^{T+1-t}) \succeq_T \Gamma_{T,t}(Y^{T+1-t})$  if and only if  $\Gamma_{T,t+1}(X_j^{T-t}) \succeq_T \Gamma_{T,t+1}(Y_j^{T-t})$ .

Recursivity is a substitution property that involves replacing one  $(t+1)$ th stage  $(T-t)$ -stage lottery by another within a  $T$ -stage lottery in the set,  $\mathcal{G}_{T,t}$ , that are degenerate until the  $t$ th stage. The power of this property is that, by using it recursively, we can apply it directly

---

<sup>8</sup> This preference relation is analogous to that defined in Kreps & Porteus (1978) section 4, p.195, specialized to the case where no consumption takes place before the last period.

<sup>9</sup> The sets  $\mathcal{G}_{T,t}$  are analogous to Kreps & Porteus's (1978) sets  $P_t(y_t)$ , except again we restrict consumption to the end.

to substitutions of  $(t + 1)$ th stage  $(T - t)$ -stage lotteries within a general  $T$ -stage lottery in  $\mathcal{L}_0^T(\mathcal{X})$ .

In their original formulation of recursive expected utility, Kreps & Porteus assume a similar recursivity axiom that they call “temporal consistency”. Loosely speaking, at each stage  $t$ , Kreps & Porteus’s agent has a preference relation over the remaining sub-lotteries of length  $(T + 1 - t)$ , the number of stages still to go. They assume that these later preference relations are independent of any unrealized contingencies; that is, of the parent lotteries in which the sub-lotteries reside. Still loosely speaking, temporal consistency then requires that preferences over parent lotteries respect these later orderings. Thus, temporal consistency implies and is implied by our recursivity property if we make an identification between the agent’s stage  $t$  preferences over sub-lotteries of length  $(T + 1 - t)$  and the agent’s stage 1 preference relation  $\succeq_T$  restricted to lotteries that are degenerate in the first  $t - 1$  stages. Given recursivity, this is a natural identification to make (although not strictly necessary for the analysis below).

Notice that recursivity (or temporal consistency) does *not* imply the following stronger substitution axiom: for all  $X^{T+1-t}$ ,  $Y^{T+1-t}$ , and  $Z^{T+1-t}$  in  $\mathcal{L}_0^{T+1-t}(\mathcal{X})$  and all  $\alpha$  in  $(0, 1]$ ,  $\Gamma_{T,t}(X^{T+1-t}) \succeq_{T,t} \Gamma_{T,t}(Y^{T+1-t})$  if and only if  $\Gamma_{T,t}(\alpha X^{T+1-t} + (1 - \alpha)Z^{T+1-t}) \succeq_{T,t} \Gamma_{T,t}(\alpha Y^{T+1-t} + (1 - \alpha)Z^{T+1-t})$ . This assumption would be analogous to Kreps & Porteus’s “temporal substitution” axiom. By retaining recursivity but dropping “temporal substitution”, we move from recursive expected utility to more general recursive forms of preference.

The following substitution property captures the idea that, for any stage  $t < T$ , the agent prefers uncertainty to be resolved in stage  $t$  rather than in stage  $t + 1$ .

**Definition** *For any finite length of lotteries  $T \geq 2$ , we say that an agent’s preference relation over  $T$ -stage lotteries,  $\succeq_T$ , exhibits preference for early-resolution of uncertainty if for all  $t$  in  $\{1, \dots, T - 1\}$ , and all pairs of  $(T + 1 - t)$ -stage lotteries of the form  $X^{T+1-t} = [(X_i^{T-t}, q_i)_{i=1}^N]$  and  $Y^{T+1-t} = [(X_1^{T-t}, q_1; \dots; X_{j-1}^{T-t}, q_{j-1}; Y_1^{T-t}, \beta q_j; Y_2^{T-t}, (1 - \beta)q_j; X_{j+1}^{T-t}, q_{j+1}; \dots; X_N^{T-t}, q_N)]$  in  $\mathcal{L}_0^{T+1-t}(\mathcal{X})$  with  $\beta$  in  $[0, 1]$ : if  $X_j^{T-t} = \beta Y_1^{T-t} + (1 - \beta)Y_2^{T-t}$  then  $\Gamma_{T,t}(Y^{T+1-t}) \succeq_{T,t} \Gamma_{T,t}(X^{T+1-t})$ .*

Whereas, loosely speaking, recursivity involves the replacement of one ‘branch’ of a lottery by another branch, preference for early resolution involves ‘splitting’ one branch into two.

For any  $t$  in  $\{1, \dots, T\}$ , let  $\mathcal{F}_{T,t} \subset \mathcal{G}_{T,t}$  be the subset of  $\mathcal{L}_0^T(\mathcal{X})$  whose elements are degenerate in all but the  $t$ th stage. Notice that  $\mathcal{F}_{T,T} = \mathcal{G}_{T,T}$ . Let  $\succeq_{T,t}$  denote the restriction of  $\succeq_T$  to the set  $\mathcal{F}_{T,t}$ . For each  $t$ , the preference relation  $\succeq_{T,t}$  inherits continuity from  $\succeq_T$ . Since each element of  $\mathcal{F}_{T,t}$  is degenerate in all but one-stage, the set  $\mathcal{F}_{T,t}$  is isomorphic to the set of one-stage lotteries. That is, for each  $t$  in  $\{1, \dots, T\}$ , each one-stage lottery  $X = [(x_i, p_i)_{i=1}^N]$  in  $\mathcal{L}_0(\mathcal{X})$  is naturally associated with the  $T$ -stage lottery in  $\mathcal{F}_{T,t}$  for which the total probability of getting each outcome  $x$  in  $\mathcal{X}$  is given by  $\sum_{i: x_i=x} p_i$ . We denote this isomorphism from  $\mathcal{L}_0(\mathcal{X})$  to  $\mathcal{F}_{T,t}$  by  $\Phi_{T,t}$ . Since we want to allow the timing of resolution

of uncertainty to matter for the agent, we do *not* require that  $\Phi_{T,t}(X) \sim_T \Phi_{T,t'}(X)$  for all one-stage lotteries  $X$  in  $\mathcal{L}_0(\mathcal{X})$  and all stages  $t, t'$  in  $\{1, \dots, T\}$ . Notice that for all outcomes  $x$  in  $\mathcal{X}$ ,  $\delta_x$  is an element of  $\mathcal{F}_{T,t}$  for all  $t$ ; indeed,  $\delta_x = \Phi_{T,t}(\delta_x)$ .

Given these isomorphisms, for each  $t$  in  $\{1, \dots, T\}$ , the preference relation  $\succeq_{T,t}$  can be endowed with properties of preferences over one-stage lotteries. For example, we will assume throughout that for each  $t$ , the preference relation  $\succeq_{T,t}$  respects strict first order stochastic dominance. We will also assume that the preference relation  $\succeq_{T,1}$  over lotteries that are degenerate in all but the first stage exhibits risk aversion.

The following properties of preferences over (sets isomorphic to) one-stage lotteries can be axiomatized but since they are well-known, we define them in terms of their representations. First, exploiting the isomorphism again, we define a representation of a preference relation  $\succeq_{T,1}$ , as a function with domain  $\mathcal{L}_0(\mathcal{X})$ , (not  $\mathcal{F}_{T,t}$ ).

**Definition** *We say that the preference relation  $\succeq_{T,t}$  is represented by a functional  $V_{T,t} : \mathcal{L}_0(\mathcal{X}) \rightarrow \mathbb{R}$  if for all pairs of lotteries  $X$  and  $Y$  in  $\mathcal{L}_0(\mathcal{X})$ ,  $V_{T,t}(X) \geq V_{T,t}(Y)$  if and only if  $\Phi_{T,t}(X) \succeq_{T,t} \Phi_{T,t}(Y)$ . We say that the preference relation  $\succeq_{T,t}$  satisfies:*

- (a) *expected utility if, there exists a function  $u_{T,t} : \mathcal{X} \rightarrow \mathbb{R}$ , and a representation  $V_{T,t}$  given by  $V_{T,t}(X) = \int u_{T,t}(x) X(dx)$  for each  $X$  in  $\mathcal{L}_0(\mathcal{X})$ ;*
- (b) *betweenness<sup>10</sup> if, there exists a function  $v_{T,t} : \mathcal{X} \times \mathbb{R} \rightarrow \mathbb{R}$ , and a representation  $V_{T,t}$  implicitly defined by  $V_{T,t}(X) = \int v_{T,t}(x, V_{T,t}(X)) X(dx)$  for each  $X$  in  $\mathcal{L}_0(\mathcal{X})$ ;*
- (c) *rank dependence<sup>11</sup> if, there exists a function  $u_{T,t} : \mathcal{X} \rightarrow \mathbb{R}$ , a strictly increasing function  $g_{T,t} : [0, 1] \rightarrow [0, 1]$ , with  $g_{T,t}(0) = 0$  and  $g_{T,t}(1) = 1$ , and a representation  $V_{T,t}$  given by  $V_{T,t}(X) = \int u_{T,t}(x) d[g_{T,t} \circ G_X(x)]$ , where  $G_X$  is the decumulative function of  $X$ , for each  $X$  in  $\mathcal{L}_0(\mathcal{X})$ .*

As the name suggests, Kreps & Porteus's recursive expected utility model combines recursivity ("temporal consistency") with expected utility ("substitution") for all  $t$ . We can now state our result.

**Proposition 1** *Fix a  $T \geq 2$ . Suppose that an agent's preference relation,  $\succeq_T$ , over the set of  $T$ -stage lotteries satisfies recursivity and preference for early-resolution. Suppose further that the restricted preference relation,  $\succeq_{T,t}$ , respects first order stochastic dominance for all  $t$  in  $\{1, \dots, T\}$  and that the first-stage restricted preference relation,  $\succeq_{T,1}$ , satisfies risk aversion. Then:*

---

<sup>10</sup> See Chew (1983, 1989) and Dekel (1986).

<sup>11</sup> See Quiggin (1982), Yaari (1987).

(i) if the preference relation  $\succeq_{T,t}$  satisfies betweenness for all  $t$  in  $\{1, \dots, T\}$ , then  $\succeq_{T,t}$  satisfies expected utility for all  $t$  in  $\{2, \dots, T\}$ ;

(ii) if the preference relation  $\succeq_{T,t}$  satisfies rank dependence for all  $t$  in  $\{1, \dots, T\}$ , then  $\succeq_{T,t}$  satisfies expected utility for all  $t$  in  $\{1, \dots, T-1\}$ .

### 3 Proof

Notice that, for any  $X$  and  $Y$  in  $\mathcal{L}_0(\mathcal{X})$ , any  $t$  in  $\{1, \dots, T\}$  and any  $\alpha$  in  $[0, 1]$ ,  $\Phi_{T+1-t,1}(X)$  in  $\mathcal{F}_{T+1-t,1} \subset \mathcal{L}_0^{T+1-t}(\mathcal{X})$  is a  $(T+1-t)$ -stage lottery in which all uncertainty is resolved in the first stage. We use the following four facts: (i)  $\Gamma_{T,t}(\Phi_{T+1-t,1}(X)) = \Phi_{T,t}(X)$  (the corresponding  $T$ -stage lottery in which all uncertainty is resolved at the  $t$ th stage); (ii)  $\Phi_{T+1-t,1}(\alpha X + (1-\alpha)Y) = \alpha\Phi_{T+1-t,1}(X) + (1-\alpha)\Phi_{T+1-t,1}(Y)$ . Moreover, for any  $X^{T+1-t}$  and  $Y^{T+1-t}$  in  $\mathcal{L}_0^{T+1-t}(\mathcal{X})$  with  $t \neq 1$ , (iii)  $\Gamma_{T,t}(X^{T+1-t}) = \Gamma_{T,t-1}([X^{T+1-t}, 1])$ ; and (iv)  $\Gamma_{T,t-1}([X^{T+1-t}, \alpha; Y^{T+1-t}, (1-\alpha)]) = \Gamma_{T,t-1}(\alpha[X^{T+1-t}, 1] + (1-\alpha)[Y^{T+1-t}, 1])$ .

The following lemma uses the definition of preference for early resolution recursively.

**Lemma 1** *Suppose that an agent's preference relation,  $\succeq_T$ , over the set of  $T$ -stage lotteries satisfies preference for early-resolution. Then, for any  $t$  in  $\{1, \dots, T\}$ , pair of one stage lotteries  $X$  and  $Y$  in  $\mathcal{L}_0(\mathcal{X})$  and  $\alpha$  in  $[0, 1]$ ,  $\alpha\Phi_{T,t}(X) + (1-\alpha)\Phi_{T,t}(Y) \succeq_T \Phi_{T,t}(\alpha X + (1-\alpha)Y)$ .*

**Proof.** Fix any  $X, Y, t$  and  $\alpha$ . If  $t = 1$ , the claim follows immediately from fact (ii). So assume  $t \neq 1$ . Applying facts (i), (ii) and (iii) in turn, we get  $\Phi_{T,t}(\alpha X + (1-\alpha)Y) = \Gamma_{T,t}(\Phi_{T+1-t,1}(\alpha X + (1-\alpha)Y)) = \Gamma_{T,t}(\alpha\Phi_{T+1-t,1}(X) + (1-\alpha)\Phi_{T+1-t,1}(Y)) = \Gamma_{T,t-1}([\alpha\Phi_{T+1-t,1}(X) + (1-\alpha)\Phi_{T+1-t,1}(Y), 1])$ . Applying preference for early resolution we get  $\Gamma_{T,t-1}([\Phi_{T+1-t,1}(X), \alpha; \Phi_{T+1-t,1}(Y), (1-\alpha)]) \succeq_T \Gamma_{T,t-1}([\alpha\Phi_{T+1-t,1}(X) + (1-\alpha)\Phi_{T+1-t,1}(Y), 1])$ . And, applying fact (iv), we get  $\Gamma_{T,t-1}([\Phi_{T+1-t,1}(X), \alpha; \Phi_{T+1-t,1}(Y), (1-\alpha)]) = \Gamma_{T,t-1}(\alpha[\Phi_{T+1-t,1}(X), 1] + (1-\alpha)[\Phi_{T+1-t,1}(Y), 1])$ . If  $t = 2$ , this last expression becomes  $\Gamma_{T,1}(\alpha[\Phi_{T-2,1}(X), 1] + (1-\alpha)[\Phi_{T-2,1}(Y), 1]) = \alpha[\Phi_{T-2,1}(X), 1] + (1-\alpha)[\Phi_{T-2,1}(Y), 1] = \alpha\Phi_{T,t}(X) + (1-\alpha)\Phi_{T,t}(Y)$  and we are done. If not, repeating the use of fact (iii), preference for early resolution and fact (iv), we get  $\Gamma_{T,t-2}(\alpha[[\Phi_{T+1-t,1}(X), 1], 1] + (1-\alpha)[[\Phi_{T+1-t,1}(Y), 1], 1])$ . If  $t = 3$ , as before, we are done. In general, if we carry out these steps  $(t-1)$  times, we are done. ■

Let the certainty equivalent function  $\text{CE}_{T,t}(\cdot) : \mathcal{L}_0(\mathcal{X}) \rightarrow \mathcal{X}$  be the function that maps each  $X$  in  $\mathcal{L}_0(\mathcal{X})$  to the outcome in  $\mathcal{X}$  such that  $\Phi_{T,t}(X) \sim_T \delta_{\text{CE}_{T,t}(\Phi_{T,t}(X))}$ . Since we take  $\mathcal{X}$  to be  $[0, 1]$  and assume that the preference relation  $\succeq_{T,t}$  respect first order stochastic dominance for all  $t$  in  $\{1, \dots, T\}$ , the function  $\text{CE}_{T,t}(\cdot)$  is a representation of  $\succeq_{T,t}$ .

**Lemma 2** *Suppose that an agent's preference relation,  $\succeq_T$ , over the set of  $T$ -stage lotteries satisfies recursivity and preference for early-resolution. Suppose further that the restricted*

preference relation for the first stage  $\succeq_{T,1}$  respects first order stochastic dominance and satisfies risk aversion. Then for all  $t$  in  $\{2, \dots, T\}$ , the certainty equivalent function  $\text{CE}_{T,t}$  is convex in the probabilities.

**Proof.** From Lemma 1, we know that  $\alpha\Phi_{T,t}(X) + (1-\alpha)\Phi_{T,t}(Y) \succeq_T \Phi_{T,t}(\alpha X + (1-\alpha)Y)$ . But, by the definition of certainty equivalent,  $\Phi_{T,t}(\alpha X + (1-\alpha)Y) \sim_T \delta_{\text{CE}_{T,t}(\alpha X + (1-\alpha)Y)}$ , which is an element of  $\mathcal{F}_{T,1}$ . And, by the recursivity and the definition of certainty equivalent,  $\alpha\Phi_{T,t}(X) + (1-\alpha)\Phi_{T,t}(Y) \sim_T [\delta_{\text{CE}_{T,t}(X)}, \alpha; \delta_{\text{CE}_{T,t}(Y)}, (1-\alpha)]$ , which is also an element of  $\mathcal{F}_{T,1}$ . And by the risk aversion of  $\succeq_{T,1}$ , we have  $\delta_{\alpha \text{CE}_{T,t}(X) + (1-\alpha) \text{CE}_{T,t}(Y)} \succeq_{T,1} [\delta_{\text{CE}_{T,t}(X)}, \alpha; \delta_{\text{CE}_{T,t}(Y)}, (1-\alpha)]$ . Combining, we get  $\delta_{\alpha \text{CE}_{T,t}(X) + (1-\alpha) \text{CE}_{T,t}(Y)} \succeq_{T,1} \delta_{\text{CE}_{T,t}(\Phi_{T,t}(\alpha X + (1-\alpha)Y))}$ , which, by first order stochastic dominance (indeed just monotonicity) implies  $\alpha \text{CE}_{T,t}(X) + (1-\alpha) \text{CE}_{T,t}(Y) \geq \text{CE}_{T,t}(\alpha X + (1-\alpha)Y)$ . ■

We are now ready to prove the first part of the proposition.

**Proof of Part (i) (Betweenness).** Suppose that the function  $V_{T,t} : \mathcal{L}_0(\mathcal{X}) \rightarrow \mathbb{R}$  represents the preference relation,  $\succeq_{T,t}$ . If  $\succeq_{T,t}$  satisfies betweenness then, for all pairs of lotteries  $X$  and  $Y$  in  $\mathcal{L}_0(\mathcal{X})$  and all  $\alpha$  in  $[0, 1]$ , if  $V_{T,t}(X) = V_{T,t}(Y)$  then  $V_{T,t}(\alpha X + (1-\alpha)Y) = V_{T,t}(X)$ . Recall that the certainty equivalent function  $\text{CE}_{T,t} : \mathcal{L}_0(\mathcal{X}) \rightarrow \mathbb{R}$  is a representation of  $\succeq_{T,t}$ . By Lemma 2, for any  $t$  in  $\{2, \dots, T\}$ , this representation is convex in the probabilities. Therefore, it is enough to prove the following lemma which is an application of Kannai's (1976) general result.

**Lemma 3** *Suppose that  $V_{T,t}$  represents a preference relation  $\succeq_{T,t}$  that satisfies betweenness. If there exist a function  $h$  from  $\mathbb{R}$  to  $\mathbb{R}$  such that  $h \circ V_{T,t}$  is convex, then the preference relation satisfies expected utility.*

Roughly speaking, if indifference curves are planar, they cannot be represented by a convex utility function unless they are also parallel. So, from the convexity of  $\text{CE}_{T,t}$ , the preference relation  $\succeq_{T,t}$  must satisfy expected utility.

**Proof.** The argument is an application of Theorem 3.4 of Kannai (1976), which deals with concavity instead of convexity. It is clear that either will do for Lemma 3. To be consistent with Kannai's notation, we shall show that if there is an  $h$  such that  $h \circ V_{T,t}$  is *concave*, then independence (hence expected utility) must be satisfied.

Suppose that the preference relation  $\succeq_{T,t}$  does not satisfy independence. Thus, there exist  $X_1, X_2, Y$  in  $\mathcal{L}_0(\mathcal{X})$  such that  $V_{T,t}(X_1) = V_{T,t}(X_2)$  but  $V_{T,t}(\alpha X_1 + (1-\alpha)Y) > V_{T,t}(\alpha X_2 + (1-\alpha)Y)$  for some  $\alpha$  in  $(0, 1)$ . Let  $\Delta$  be the convex hull of  $\{X_1, X_2, Y\}$ , which is isomorphic to the set  $\{(p_1, p_2) \in R_+^2 : p_1 + p_2 \leq 1\}$ . The restriction of  $V_{T,t}$  on  $\Delta$  satisfies betweenness, so by construction, the indifference curves restricted on  $\Delta$  are (isomorphic to) straight lines, but they are not parallel. Abusing notation, we shall identify  $\Delta$  with the 2-dimensional set above, and we shall write  $V(p)$  for  $V_{T,t}(p_1 X_1 + p_2 X_2 + (1-p_1-p_2)Y)$  for each  $p = (p_1, p_2)$

in  $\Delta$ . The function  $h \circ V$  is concave on  $\Delta$ . So we are done if we can show that a continuous function on  $\mathbb{R}^2$  whose level curves are linear but not parallel on a convex set with non-empty interior cannot be concavified.

We shall construct a sequence of points which violates Kannai's necessary condition. Consider some rectangle  $R$  contained in the interior of  $\Delta$  on which some indifference curves are not parallel. By an appropriate choice of coordinates, we can identify the rectangle with  $[0, 1] \times [0, 1]$ . For each  $s$  in  $[0, 1]$ , let  $g(s)$  be the slope of the indifference curve going through point  $(0, s)$ . That is, point  $(x, y)$  on this curve is expressed by the linear equation  $y = s + g(s)x$ . Since preferences are continuous,  $g$  is a continuous function. Without loss of generality, set  $g(0) = 0$ ; that is, choose the rectangle so that its base lies along an indifference curve. Since indifference curves are not all parallel,  $g(s)$  is not identical to 0. So, without loss of generality, assume  $g(\bar{s}) > 0$  for some  $\bar{s} \in (0, 1]$ , and let  $s^* = \sup\{s \in [0, \bar{s}] : g(s) \leq 0\}$ . Since  $g$  is continuous,  $s^* < \bar{s}$ . Thus we can adjust our rectangle so that  $s^* = 0$ . To sum up, we have constructed a rectangle  $R$ , and a continuous function  $g$  on  $[0, \bar{s}]$  with  $g(0) = 0$  and  $g(s) > 0$  for every  $s$  in  $(0, \bar{s}]$ . Notice that we can make the rectangle  $R$  arbitrarily small. Also note that in our new coordinates, if  $g(s) \neq g(s')$ , the indifference curves corresponding to  $s$  and  $s'$  intersect at  $\left(-\frac{s-s'}{g(s)-g(s')}, \frac{s'g(s)-sg(s')}{g(s)-g(s')}\right)$ , which must be outside  $\Delta$ .

Now Kannai's condition in our context can be stated as follows. Let  $v(p) = (v_1(p), v_2(p))$  in  $\mathbb{R}^2$  be the normalized gradient of  $V$  at  $p$  in  $R$ . That is,  $|v(p)| = 1$  and  $V(p) = V(p')$  if and only if  $v(p) \cdot (p - p') = 0$ . Pick  $p, q$  in  $Z$  with  $V(p) > V(q)$  and let  $\lambda(p, q) = \sup\left\{\frac{v(p) \cdot (p' - q')}{v(q) \cdot (p' - q')} : V(p) = V(p'), V(q) = V(q'), p' \in \Delta, q' \in \Delta\right\}$ . Kannai's theorem shows that concavification is possible only if  $\prod_{i=0}^{n-1} \lambda(p_i, p_{i+1})$  is bounded for any points  $p_0, p_1, \dots, p_n$  with  $V(p_0) > V(p_1) > \dots > V(p_n)$ .

The term  $\lambda(p, q)$  can be expressed more conveniently as follows. Suppose  $v(p) \neq v(q)$ , thus the two corresponding indifference curves intersect. Let  $r(p, q)$  denote the intersection point. Let  $\theta(p, q)$  be the angle of the two indifference curves at  $r(p, q)$ . If  $V(p') = V(p)$  and  $V(q') = V(q)$ , then  $v(p) \cdot (p' - r(p, q)) = 0$  and  $v(q) \cdot (q' - r(p, q)) = 0$ , and  $v(p) \cdot (q' - r(p, q)) = |q' - r(p, q)| \cos(\frac{\pi}{2} + \theta) = -|q' - r(p, q)| \sin \theta$  and  $v(q) \cdot (p' - r(p, q)) = |p' - r(p, q)| \cos(\frac{\pi}{2} - \theta) = |p' - r(p, q)| \sin \theta$ . From  $\frac{v(p) \cdot (p' - q')}{v(q) \cdot (p' - q')} = \frac{v(p) \cdot ((p' - r(p, q)) - (q' - r(p, q)))}{v(q) \cdot ((p' - r(p, q)) - (q' - r(p, q)))}$ , we get  $\lambda(p, q) = \sup\left\{\frac{|q' - r(p, q)|}{|p' - r(p, q)|} : V(p) = V(p'), V(q) = V(q'), p' \in \Delta, q' \in \Delta\right\}$ .

We claim that, if the rectangle  $R$  is chosen small enough, then for all  $p, q$  in  $Z$  such that  $r(p, q)$  is bounded,  $\lambda(p, q)$  is bounded above 1. To see this, let  $d$  be the length of the diagonal of  $R$ , and let  $b$  be the (shortest) distance from  $R$  to the border of  $\Delta$ . Choose  $R$  small enough such that  $d < b$ . Then for any  $p, q \in R$ , with  $V(p) > V(q)$ ,  $|r(p, q) - p| - d \leq |r(p, q) - q|$ . So, if  $r(p, q)$  is bounded,  $\lambda(p, q) \geq \frac{|r(p, q) - q|}{|r(p, q) - p| - b} \geq \frac{|r(p, q) - p| - d}{|r(p, q) - p| - b} > 1$ . So choosing  $R$  small enough, we are done if there is a sequence of points  $\{p_i : i = 0, 1, \dots\}$  in  $R$  with  $V(p_{i-1}) > V(p_i)$  for each  $i = 1, 2, \dots$  and  $v(p_{i-1}) \neq v(p_i)$  for  $i = 1, 2, \dots$ , such that  $\{r(p_{i-1}, p_i) : i = 1, \dots\}$  is bounded. We shall construct such a sequence in the following.

To simplify our notation, set  $\rho_1(s, s') = -\frac{s-s'}{g(s)-g(s')}$ ; that is,  $\rho_1(s, s')$  is the horizontal

coordinate of  $r((0, s), (0, s'))$ , and  $\rho_2(s, s') = \frac{s'g(s) - sg(s')}{g(s) - g(s')}$ . By construction,  $\rho_1(0, s) = -\frac{s}{g(s)}$ , and since the intersection point lies outside the simplex, there is a positive constant  $k$  such that  $\rho_1(0, s) < -k$ , or equivalently,  $s > kg(s)$ , for any  $s$  in  $(0, \bar{s}]$ . Since  $g$  is continuous and  $g(s) > 0$  for all  $s$  in  $(0, \bar{s}]$ , the horizontal intercept  $\rho_1(0, s)$  is continuous in  $s$  on that domain. There are two possible cases.

Case 1. There exists a pair  $s, s'$  in  $(0, \bar{s}]$  with  $s < s'$  and  $\rho_1(0, s) < \rho_1(0, s')$  and, hence,  $sg(s') > s'g(s)$ . Together,  $s < s'$  and  $sg(s') > s'g(s)$  imply  $g(s) < g(s')$ , hence  $\rho_1(s, s') < 0$ . So we have  $\rho_1(s, s') - \rho_1(0, s) = -\frac{s-s'}{g(s)-g(s')} + \frac{s}{g(s)} = \frac{s'g(s) - sg(s')}{g(s)(g(s)-g(s'))} > 0$ . Hence  $0 > \rho_1(s, s') > \rho_1(0, s)$ . Also,  $0 < \rho_2(s, s') = \frac{sg(s') - s'g(s)}{g(s') - g(s)} < \frac{s'g(s') - s'g(s)}{g(s') - g(s)} = s'$ .

In this case, let  $s_1 := s'$ . By the continuity of  $\rho(0, s)$ , we can find an  $s_2$  in  $(s, s_1)$  such that  $\rho(0, s) < \rho(0, s_2) < \rho(0, s_1)$ , (hence,  $\frac{s}{g(s)} > \frac{s_2}{g(s_2)} > \frac{s_1}{g(s_1)}$ ). Notice that since  $s < s_2 < s_1$ , we have  $g(s) < g(s_2) < g(s_1)$ , so applying a similar calculation to that above, we conclude  $0 > \rho_1(s_2, s_1) > \rho_1(0, s)$  and  $0 < \rho_2(s_2, s_1) < s_1$ . For  $n = 3, \dots$ , we can iteratively find an  $s_n$  in  $(s, s_{n-1})$  such that  $\rho_1(0, s) < \rho_1(0, s_n) < \rho_1(0, s_{n-1})$ . Then we have a sequence  $\{s_n\}$  such that  $0 > \rho_1(s_n, s_{n-1}) > \rho_1(0, s)$  and  $0 < \rho_2(s_n, s_{n-1}) < s_1$  for all  $n$ , as we wanted.

Case 2. For any pair  $s, s'$  in  $(0, \bar{s}]$ , if  $s < s'$ , then  $\rho_1(0, s) \geq \rho_1(0, s')$ ; that is,  $\rho_1(0, \cdot)$  is an weakly decreasing function on  $(0, \bar{s}]$ . There are two subcases.

Case 2.1. There exist a pair  $s, s'$  in  $(0, \bar{s}]$  with  $s < s'$  and  $\rho_1(0, s) = \rho_1(0, s')$ . Then, for all  $s'', s'''$  in  $[s, s']$ ,  $\rho_1(0, s'') = \rho_1(0, s''') = \rho_1(0, s) = \rho_1(s'', s''')$ . Hence there exists a strictly decreasing sequence  $\{s_n\}$  in  $[s, s']$  such that  $\rho_1(s_n, s_{n+1})$  is constant and  $\rho_2(s_n, s_{n+1}) = 0$ .

Case 2.2. For any pair  $s, s'$  in  $(0, \bar{s}]$ , if  $s < s'$ , then  $\rho_1(0, s) > \rho_1(0, s')$ . Fix some  $s_1$  in  $(0, \bar{s}]$ . Let  $s_2$  be given by  $s_2 := kg(s_1)$ . That is, the line through  $(0, s_2)$  and  $(-k, 0)$  is parallel to the indifference curve through  $(0, s_1)$ . By construction, we have  $0 < s_2 < s_1$ . Since  $s_2$  is in  $(0, \bar{s}]$ , the indifference curve through  $(0, s_2)$  passes above the point  $(-k, 0)$ . So,  $\rho_1(0, s_2) < -k$ , and  $g(s_2) < g(s_1)$ . Therefore,  $-\infty < \rho_1(s_2, s_1) < 0$ . By our case definition, we also know that  $\rho_1(0, s_1) < \rho_1(0, s_2)$ , so  $-\infty < \rho_2(s_2, s_1) < 0$ .

Next let  $f(s_1, s_2) := \frac{\rho_2(s_2, s_1)}{(\rho_1(s_2, s_1) + k)}$  be the slope of the line through the intersection point  $r((0, s_2), (0, s_1))$  and  $(-k, 0)$ . And let  $s_3 := f(s_1, s_2)k$  be the vertical intercept of this line. Recall that the indifference curve through  $(0, s_2)$  passes through the point  $r((0, s_2), (0, s_1))$  and above the point  $(-k, 0)$  so, by construction,  $g(s_2) > f(s_1, s_2) > 0$ . Hence  $s_2 > s_3 > 0$ . Since  $s_3$  is in  $(0, \bar{s}]$ , the indifference curve through  $(0, s_3)$  passes above the point  $(-k, 0)$ . So,  $\rho_1(0, s_3) < -k$ , and  $g(s_3) < f(s_1, s_2) < g(s_2)$ . Thus  $\rho_1(s_3, s_2) < 0$  and  $\rho_1(s_3, s_2) = -\frac{s_2 - s_3}{g(s_2) - g(s_3)} > -\frac{s_2 - s_3}{g(s_2) - f(s_1, s_2)} = \rho_1(s_2, s_1)$ . Also, by our case definition, we also know that  $\rho_1(0, s_2) < \rho_1(0, s_3)$ , so  $\rho_2(s_2, s_1) < 0$ . Moreover,  $\rho_2(s_3, s_2) = s_3 + f(s_1, s_2)\rho_1(s_3, s_2) > s_3 + f(s_1, s_2)\rho_1(s_2, s_1) = \rho_2(s_2, s_1)$ . We can iteratively construct  $s_n$  from  $s_{n-1}$  and  $s_{n-2}$  as we constructed  $s_3$  from  $s_2$  and  $s_1$ . Then we have a sequence  $\{s_n\}$   $0 > \rho_1(s_n, s_{n-1}) > \rho_1(s_2, s_1)$  and  $0 > \rho_2(s_n, s_{n-1}) > \rho_2(s_2, s_1)$  for all  $n$ , as we wanted. ■

**Proof of Part (ii) (Rank-Dependence).** Recall that the certainty equivalent function

$CE_{T,t} : \mathcal{L}_0(\mathcal{X}) \rightarrow \mathbb{R}$  is a representation of  $\succeq_{T,t}$ . By Lemma 2, for any  $t$  in  $\{2, \dots, T\}$ , this representation is convex in the probabilities. So in particular, it is quasi-convex in the probabilities; that is, for all pairs of lotteries  $X$  and  $Y$  in  $\mathcal{L}_0(\mathcal{X})$ , and all  $\alpha$  in  $[0, 1]$ , if  $CE_{T,t}(X) \geq CE_{T,t}(Y)$  then  $CE_{T,t}(X) \geq CE_{T,t}(\alpha X + (1 - \alpha)Y)$ . Since quasi-convexity is an ordinal property, all representations of  $\succeq_{T,t}$  must be quasi-convex. Suppose that  $V_{T,t} : \mathcal{L}_0(\mathcal{X}) \rightarrow \mathbb{R}$  is a representation given by  $V_{T,t}(X) = \int u_{T,t}(x)d[g_{T,t} \circ G_X(x)]$ , where  $G_X$  is the decumulative function of  $X$ , and  $u_{T,t}$  and  $g_{T,t}$  are functions satisfying the definition of rank dependence. Wakker's (1994) Theorem 24 shows that such rank-dependent functionals are quasi-convex in the probabilities if and only if the function  $g_{T,t}$  is convex. Therefore  $g_{T,t}$  is convex for all  $t$  in  $\{2, \dots, T\}$

The following lemma extends a result of Chew, Karni and Safra (1987) eschewing the need for Gateaux differentiability.

**Lemma 4** *Let  $V : \mathcal{L}_0(\mathcal{X}) \rightarrow \mathbb{R}$  be given by  $V(X) = \int u(x)d[g \circ G_X(x)]$  where  $G_X$  is the decumulative function of  $X$ , for each  $X$  in  $\mathcal{L}_0(\mathcal{X})$ , and  $u$  and  $g$  are functions satisfying the definitions of rank dependence. If the preference relation represented by  $V$  is risk averse (loving) then  $g$  is convex (concave).*

**Proof.** Suppose  $g$  is not convex. Then there exists a  $\hat{q}$  in  $(0, 1)$  and an  $\eta$  in  $(0, \min\{\hat{q}, 1 - \hat{q}\})$  such that  $\frac{1}{2}g(\hat{q} + \eta) + \frac{1}{2}g(\hat{q} - \eta) < g(\hat{q})$ ; that is,  $[g(\hat{q}) - g(\hat{q} - \eta)] - [g(\hat{q} + \eta) - g(\hat{q})] > 0$ . Recall that we take  $\mathcal{X}$  to be the interval  $[0, 1]$ . Consider the pair of one-stage lotteries,  $Y = [0, (1 - \hat{q} - \eta); x, 2\eta; 1, (\hat{q} - \eta)]$  and  $Z = [0, (1 - \hat{q} - \eta); x - \varepsilon, \eta; x + \varepsilon, \eta; 1, (\hat{q} - \eta)]$  in  $\mathcal{L}_0(\mathcal{X})$  where both  $x$  is in  $(0, 1)$  and  $\varepsilon$  is in  $(0, \min\{x, 1 - x\})$ . Risk aversion implies that  $V(Y) \geq V(Z)$ ; that is,

$$[g(\hat{q} + \eta) - g(\hat{q})][u(x) - u(x - \varepsilon)] \geq [g(\hat{q}) - g(\hat{q} - \eta)][u(x + \varepsilon) - u(x)] \quad (1)$$

If  $u$  is not concave, then we can choose our  $x$  and  $\varepsilon$  such that  $[u(x) - u(x - \varepsilon)] < [u(x + \varepsilon) - u(x)]$ , yielding a contradiction. Therefore assume  $u$  is concave. Since preferences respect strict first order stochastic dominance, the function  $u$  is strictly increasing. Therefore  $u$  is differentiable almost everywhere. Moreover, since it is concave, the right and left hand derivatives exist everywhere and are increasing.<sup>12</sup> Therefore the derivative is strictly positive almost everywhere. So we can find an  $x$  such that  $u'(x)$  exists and is strictly positive. Therefore, dividing both sides of the expression 1 by  $\varepsilon$  and taking the limit as  $\varepsilon$  go to zero, we get  $[g(\hat{q} + \eta) - g(\hat{q})]u'(x) \geq [g(\hat{q}) - g(\hat{q} - \eta)]u'(x)$ , a contradiction. The proof for risk loving is similar. ■

An immediate implication of Lemma 4 is that, for the first stage,  $g_{T,1}$  is convex. We next show that  $g_{T,t}$  is concave for all  $t$  in  $\{1, \dots, T - 1\}$ .

Fix a  $t$  in  $\{1, \dots, T - 1\}$  and suppose that  $g_{T,t}$  is not concave. Then there exists a  $\hat{q}$  in  $(0, 1)$  and an  $\eta$  in  $(0, \min\{\hat{q}, 1 - \hat{q}\})$  such that  $[g(\hat{q}) - g(\hat{q} - \eta)] - [g(\hat{q} + \eta) - g(\hat{q})]$

<sup>12</sup> See, for example, Royden (1988) p. 113.

$< 0$ . For any  $r$  in  $[0, 1]$ , let  $X_r := [1, r; 0, (1 - r)]$  in  $\mathcal{L}_0(\mathcal{X})$  be the one-stage lottery that places weight  $r$  on the best outcome and  $(1 - r)$  on the worst. Consider the pair of  $T$ -stage lotteries,  $Y^T$  and  $Z^T$ , in  $\mathcal{L}_0^T(\mathcal{X})$  given by  $Y^T := \Gamma_{T,t}[[\Phi_{T-t,1}(X_0), (1 - \hat{q} - \eta); \Phi_{T-t,1}(X_r), 2\eta; \Phi_{T-t,1}(X_1), (\hat{q} - \eta)]]$  and  $Z^T := \Gamma_{T,t}[[\Phi_{T-t,1}(X_0), (1 - \hat{q} - \eta); \Phi_{T-t,1}(X_{r-\varepsilon}), \eta; \Phi_{T-t,1}(X_{r+\varepsilon}), \eta; \Phi_{T-t,1}(X_1), (\hat{q} - \eta)]]$  where  $\hat{r}$  is in  $(0, 1)$  and  $\varepsilon$  is in  $(0, \min\{\hat{r}, 1 - \hat{r}\})$ . Preference for early resolution implies  $Z^T \succeq_T Y^T$ .

Without loss of generality we normalize  $u_{T,t+1}(1) = 1$  and  $u_{T,t+1}(0) = 0$ . Notice that, for all  $r$  in  $[0, 1]$ ,  $\text{CE}_{T,t+1}(X_r) = u_{T,t+1}^{-1} \circ g_{T,t+1}(r)$ . Hence, using recursivity,  $Z^T \succeq_T Y^T$  implies

$$\begin{aligned} & [g_{T,t}(\hat{q} + \eta) - g_{T,t}(\hat{q})][u_{T,t} \circ u_{T,t+1}^{-1} \circ g_{T,t+1}(\hat{r}) - u_{T,t} \circ u_{T,t+1}^{-1} \circ g_{T,t+1}(\hat{r} - \varepsilon)] \quad (2) \\ & \leq [g_{T,t}(\hat{q}) - g_{T,t}(\hat{q} - \eta)][u_{T,t} \circ u_{T,t+1}^{-1} \circ g_{T,t+1}(\hat{r} + \varepsilon) - u_{T,t} \circ u_{T,t+1}^{-1} \circ g_{T,t+1}(\hat{r})]. \end{aligned}$$

Notice that expression 2 is analogous to expression 1 with the inequality reversed, the compound function  $u_{T,t} \circ u_{T,t+1}^{-1} \circ g_{T,t+1}$  taking the place of  $u$  and the function  $g_{T,t}$  taking the place of  $g$ . Since our choice of  $\hat{r}$  was arbitrary, an argument exactly analogous to that in lemma 4 implies  $g_{T,t}$  is concave.

Since  $g_{T,t}$  is both concave and convex for all  $t$  in  $\{1, \dots, T - 1\}$ , these stages must be expected utility. ■

## References

- Bernasconi, Michele (1994): “Non-Linear Preferences and Two-Stage Lotteries: Theories and Evidence”, *Economic Journal*, 104, 54-70.
- 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.
- Chew, Soo Hong (1989): “Axiomatic Utility Theories with the Betweenness Property”, *Annals of Operations Research*, 19, 273-298.
- Chew, Soo Hong and Larry Epstein (1989): “The Structure of Preferences and Attitudes Towards the Timing of the Resolution of Uncertainty”, *International Economic Review*, 30: 103-117.
- Chew, Soo Hong, Edi Karni and Zvi Safra (1987): “Risk Aversion in the Theory of Expected Utility with Rank-Dependent Preferences”, *Journal of Economic Theory*, 42, 370-381.
- Conlisk, John (1989): “Three Variants on the Allais Example”, *American Economic Review*, 79, 392-407.
- Dekel, Eddie (1986): “An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom”, *Journal of Economic Theory*, 40, 304-318.

- Epstein, Larry (1992): "Behavior under risk: recent developments in theory and applications", in *Advances in Economic Theory: Sixth World Congress Vol. II*, ed. Jean-Jacques Laffont, Cambridge University Press.
- Epstein, Larry and Stanley Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework", *Econometrica*, 57, 937-969.
- Epstein, Larry and Stanley Zin (1991): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis", *Journal of Political Economy*, 99, 263-286.
- Grant, Simon, Atsushi Kajii and Ben Polak (1992): "Many Good Choice Axioms: When Can Many be Treated As One?", *Journal of Economic Theory*, 56, 313-337.
- Grant, Simon, Atsushi Kajii and Ben Polak (1997): "Intrinsic Preference for Information", Australian National University, Department of Economics working paper no. 323.
- Gul, Faruk (1991): "A Theory of Disappointment Aversion", *Econometrica*, 59, 667-686.
- Gul, Faruk and Outi Lantto (1990): "Betweenness Satisfying Preferences and Dynamic Choice", *Journal of Economic Theory*, 52, 162-177.
- Kannai, Yakar, (1976): "Concavifiability and Constructions of Concave Utility Functions", *Journal of Mathematical Economics*, 4, 1-56.
- Kreps, David and Evan Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory", *Econometrica*, 46, 185-200.
- Machina, Mark (1987): "Choice Under Uncertainty: Problems Solved and Unsolved", *Journal of Economic Perspectives*, 1, 121-154.
- Machina, Mark (1989): "Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty", *Journal of Economic Literature*, 27, 1622-1668.
- Quiggin, John (1982): "A Theory of Anticipated Utility", *Journal of Economic Behavior and Organization*, 3, 323-43.
- Royden, H.L. (1988): *Real Analysis*, Macmillan Publishing Company, New York.
- Sarin, Rakesh and Peter Wakker (1996): "Dynamic Choice and Non-Expected Utility", working paper, University of Leiden.
- Segal, Uzi (1990): "Two-Stage Lotteries without the Reduction Axiom", *Econometrica*, 58, 349-377.
- Starmer Chris and Robert Sugden (1991): "Does the Random Lottery Incentive System Elicit True Preferences? An Experimental Investigation", *American Economic Review*, 81, 971-978.
- Sugden, Robert (1995): "Alternatives to Expected Utility", in *Handbook of Utility Theory*, ed. Peter Hammond.

Wakker, Peter (1994): “Separating Marginal Utility and Probabilistic Risk Aversion”, *Theory and Decision*, 36, 1-44.

Yaari, Menahem (1987): “The Dual Theory of Choice under Risk”, *Econometrica*, 55, 95-115.