COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 1114

PREFERENCE FOR INFORMATION

Simon Grant, Atsushi Kajii and Ben Polak

January 1996

Current Version: Jan 1996
(First Version: Jan 1994)

Preference for Information.*

Abstract

What is the relationship between an agent's attitude towards information, and her attitude towards risk? If an agent always prefers more information, does this imply that she obeys the independence axiom? We provide a substitution property on preferences that is equivalent to the agent (intrinsically) liking information in the absence of contingent choices. We use this property to explore both questions, first in general, then for recursive smooth preferences, and then in specific recursive non-expected utility models. Given smoothness, for both the rank dependence and betweenness models, if an agent is information-loving then her preferences can depart from Kreps & Porteus's (1978) temporal expected utility model in at most one stage. This result does not extend to quadratic utility. Finally, we give several conditions such that, provided the agent intrinsically likes information, Blackwell's (1953) result holds; that is, she will always prefer more informative signals, whether or not she can condition her subsequent behavior on the signal.

Simon Grant
Economics Programme, RSSS
Australian National University

Atsushi Kajii
Economics Department
University of Pennsylvania

Ben Polak
Economics Department
Yale University

* The geometric interpretation of information used here was suggested to us in some classes of Jerry Green. We thank Eddie Dekel, Al Klevorick, Stephen Morris, Boaz Moselle, David Pearce, Tim Van Zandt, and Peter Wakker. Parts of this research were undertaken when Grant was visiting the Cowles Foundation at Yale, Kajii was visiting CORE, and Polak was visiting RSSS at the ANU. We thank these institutions for their hospitality. Atsushi also gratefully acknowledges support from the U. Penn Institute for Economics Research Fund. We are responsible for any errors.

1 Introduction

This paper is about preference for information and preference for the early resolution of uncertainty. In the standard model of individual choice, all agents at least weakly prefer more information to less. This is part of Blackwell's (1953) theorem. But in Blackwell's treatment, the agent's preference for information is only instrumental. She likes information only because it lets her design better strategies: if she does not or cannot condition her actions on what she learns, information is of no value to her. Introspection suggests, however, that we sometimes intrinsically prefer more information to less, even in the absence of any instrumental purpose. For example, when on the job market, Ben would have paid good money to learn sooner rather than later whether or not he was destined for the dole queue. Similarly, we sometimes choose not to be informed. Simon might pay good money not to be told the sex of his expected child were it discernible on the ultrasound.

Consider the decision whether to be tested for an incurable genetic disorder. A director of a genetic counseling program recently told the New York Times that

> "there are basically two types of people. There are 'want-to-knowers' and there are 'avoiders'. There are some people who, even in the absence of being able to alter outcomes, find information of this sort beneficial. The more they know, the more their anxiety level goes down. But there are others who cope by avoiding, who would rather stay hopeful and optimistic and not have the unanswered questions answered."[1]

If our models do not allow agents to care intrinsically about information, we will miscalculate the welfare costs and benefits of expensive public goods like the human genome project. Concern for the timing of the resolution of uncertainty may also affect the relative demand for different financial assets. It may influence the design of state lotteries. It may even change our judgment of two otherwise similar risky macroeconomic policies.[2]

Intrinsic preference for information is no more irrational than is risk aversion. Risk can be thought of as probability distributions over payoff-relevant states. An informative signal is a probability distribution over signal realizations each of which induces a posterior belief

[1] Quoted in Charles Siebert, "Living with Toxic Knowledge: The DNA We've Been Dealt", New York Times Magazine, Sept 17, 1995, p.52. We thank David Pearce for this reference. David Kelsey (1989) discusses the example of incurable diseases in a comment on a paper by Peter Wakker on aversion to information.

[2] Chew and Ho (1994) discuss how both hope and anxiety affect well-being. They suggest that asymmetries in agents' preferences for information may affect the design of news programs, of punishments, of contracts, and of financial instruments.

(itself a probability distribution) over payoff-relevant states. Thus, it can be thought of as a probability distribution over probability distributions. Faced with a probability distribution over payoff-relevant states, an agent may be concerned with more than just the expected (average) consequence: other moments of the distribution may affect her current well-being. If so, she is not risk-neutral. Faced with a distribution over distributions over payoff-relevant states, an agent may be concerned with more than just the expected (reduced) distribution: the moments at which uncertainty is resolved may affect her current well-being. If so, she is not information-neutral.

This paper explores this and other connections between risk and information. Introspection suggests that attitudes towards risk and attitudes towards information are closely related. For example, Atsushi may dislike risk in part because he is uncomfortable living with uncertainty. If so, he has reason to seek information that resolves this uncertainty and thus relieves his anxiety. Blackwell's original setup, however, is quite restrictive in the way it models attitudes towards both risk and information. For example, if Blackwell's agent always prefers the complete elimination of a risk (the mean of the original distribution) then she also always prefers any partial removal of risk (any mean-preserving contraction). This is a consequence of the expected utility hypothesis. Non-expected utility models can accommodate a richer range of both introspective and experimental evidence on attitudes towards risk; see, for example, Machina (1982, 1987). By so doing, they might also allow richer descriptions of attitudes towards information.

We therefore consider a framework with, at least initially, few restrictions on preferences. We show that the way in which we model an agent's attitude towards risk has implications for her attitude towards information and vice versa. For example, if we wish to maintain Blackwell's result that an agent always prefers more information to less, we have to restrict the overall shape of the agent's preferences over risky prospects. For some forms of preferences, these restrictions severely limit the degree to which we can relax the assumptions of the expected utility model. In fact, if preference for information is assumed, then some popular generalizations of the expected utility model allow virtually no generalization at all.

A formal statement of the results requires some investment in notation, so let us informally preview the main results here. The central part of the paper considers attitudes towards information in the absence of contingent choices; that is, where the agent has only one action available. This case allows us to focus on intrinsic attitudes toward information and serves as a benchmark for the many-action case discussed later. It is also of interest in itself as a theory of dynamic, non-expected utility preferences. We show that intrinsic preference for information is equivalent to a simple substitution property of preferences over two-stage lotteries. We call this property *single-action information loving* or *SAIL.* Since this substitution property is defined directly on preferences, not on any particular functional representation, it is applicable to all non-expected utility models. Single-action information loving is related to attitudes toward risk in two quite different ways. First, preference for information restricts how an agent's attitude towards risk in lotteries that resolve early compare to her attitude towards risk in lotteries that resolve late. For example, single-action information loving implies that the certainty equivalent of a late-resolution lottery is no greater than that of the corresponding early-resolution lottery. Second, single-action information loving is analogous to 'risk loving' with respect to the distribution of posteriors. Intuitively, information causes posteriors to be more widely dispersed. Loosely speaking, this analogy allows us to translate known results about attitudes towards risk into new results about attitudes towards information. For example, it follows immediately that if preferences are recursive then preference for information implies that preferences over late-resolution lotteries are quasi-convex in the probabilities. Intuitively, liking information is related to disliking averages, or mixtures, of posteriors.

As we add more structure, the implications of preference for information get more precise. For example, if preferences are both recursive and smooth (Gateaux differentiable), then single-action information loving can be expressed as conditions on the local utility functions. This is analogous to Machina's (1982) analysis of risk aversion for generalized expected utility preferences. The conditions are local versions of Kreps & Porteus's (1978) global condition for their temporal expected utility model.

Next, we consider recursive preferences where the agent's preferences over early- and late-resolution lotteries both come from the same family of preferences. We review the temporal expected utility model, and then consider the recursive extensions of the two most common weakenings of expected utility: rank-dependence and betweenness. For each, we provide necessary and sufficient conditions for single-action information loving. In the Kreps-Porteus model, preference for information is equivalent to preferences over early-resolution lotteries being no more risk-averse than preferences over late-resolution lotteries. The recursive rank-dependent model shows that this equivalence does not generalize. The reason is connected to the distinction mentioned above between complete and partial removal of risk. Both the rank-dependence and betweenness models, however, yield a more startling result: given smooth preferences, if an agent always prefers more information, then her preferences over multi-stage lotteries must satisfy independence in all but one stage.[3] In both cases the result comes from conflict between the restrictions put on the shape of preferences over

[3] Strictly speaking, for rank-dependence, we need risk aversion in the first stage.

stage-lotteries by the specific form of preference and those put by single-action information loving. Fortunately for non-expected utility theory, the result is not general. We provide a counterexample based on Chew, Epstein & Segal's (1991) quadratic utility. This model allows a looser connection between attitudes toward risk and the shape of preferences in the simplex.

Finally, we return to Blackwell's setting where the agent can choose among many actions contingent on the realization of a signal. We provide conditions that, combined with the substitution property SAIL, are sufficient for Blackwell's (1953) result to hold; that is, for the agent always to prefer more information, whether or not that information can affect her subsequent behavior. Intuitively, given some notion of dynamic consistency, instrumental uses of information should always reinforce any intrinsic preference for information. The only complication arises if a less informative signal allows more opportunity to construct 'randomized strategies'.

This paper builds on several earlier studies of preference for the early resolution of uncertainty and of preference for information. Kreps & Porteus's (1978) recursive or temporal expected utility model maintains the independence axiom within each set of risks that pertain to consumption in a given period and are resolved in a given (no later) period. By allowing the von Neumann-Morgenstern utility indices to vary according to how many periods separate the resolution and the consumption, the model can incorporate preference for early or for late resolution. Epstein & Zin (1989, 1991) develop and extend a homothetic version of the Kreps-Porteus model. They show that agents' risk aversion and intertemporal substitution can be estimated separately using the extended model and that this extra degree of freedom helps explain both consumption behavior and asset returns over time. Chew & Epstein (1989) weaken independence to betweenness in the Kreps-Porteus model. They provide conditions under which either a weak form of information neutrality or the existence of a particular form of information premium implies betweenness. Chew & Ho (1994) show, among other things, that experimental evidence favors a recursive betweenness model against either a standard or a temporal expected utility model.

Epstein & Zin (1989) found that attitudes toward the timing of the resolution of uncertainty were intertwined with attitudes toward risk and time substitution:

> "we suspect that the lack of separation ... reflects the inherent inseparability of these three aspects of preference rather than a deficiency of our theoretical framework. Further study of this issue is required."[4]

This paper addresses risk and information. Since we restrict consumption to take place only

[4] Epstein & Zin (1989) p. 953.

at the end of the tree, however, we abstract from questions of intertemporal substitution. But, this assumption is used just to simplify our analysis. Other than introducing notational complexity, it would be relatively easy to extend the model to allow for consumption in earlier periods. Epstein and Zin (1989) examined versions of the recursive betweenness and temporal expected utility models, but left as an open question how to characterize preference for early resolution of uncertainty in more general recursive utility functions. For the recursive betweenness model, our results strengthen those of Chew & Epstein's and Epstein & Zin's. In a sense, however, our result is negative: with preference for early resolution, the betweenness model rapidly collapses into temporal expected utility. More generally, we extend the analysis of preference for early resolution and its relation to risk aversion beyond the forms of preference previously considered in this context.

Previous analyses of preference for information and non-expected utility theory get results that, at first, appear to contradict ours. Wakker (1988) argues that if an agent always prefers more information to less, then her preferences must satisfy independence. Wakker's assumptions, however, implicitly exclude agents who have intrinsic strict preference for or against information. Schlee (1991) argues that all agents who respect first-order stochastic dominance must prefer perfect information to no information. Schlee's assumptions, however, implicitly exclude agents who have intrinsic strict preference between complete information and no information. Schlee's (1990) study of rank-dependent expected utility makes similar assumptions. We show that both Wakker's and Schlee's results depend on these assumptions. If, however, we confine attention to recursive betweenness or rank-dependent preferences, then intrinsic preference for information almost implies Wakker's conclusion of full independence.

Like Wakker, Machina (1989) is concerned with instrumental rather than intrinsic preference for information. Under Machina's assumptions, however, preferring more information to less does not imply independence. Wakker's agents anticipate that their future behavior will not be (dynamically) consistent with their current preferences. They forgo information precisely to restrict the choices available to their future selves. In contrast, Machina's agents are dynamically consistent or 'resolute'. Moreover, their preferences do not necessarily have a recursive or fold-back form; indeed, Machina argues against this degree of consequentialism. In this paper, we confine attention to agents who are either dynamically consistent in Machina's sense or who fail to anticipate their future inconsistencies. Many of our results depend on preferences having a recursive form but, cognizant of Machina's objections, we sometimes forgo this assumption. In particular, the substitution property SAIL does not rely on the fold-back property. Moreover, both Wakker and Machina (absent intrinsic concerns) favor preference for information as an axiom of individual choice. We show that, among resolute agents, the fold-back property is not necessary for single-action information loving to imply a general preference for information, regardless of the number of available actions.

For most of the analysis below, we have in mind that the agents' choices and the resolution of uncertainty take place in real time. That is, we think of all decision trees presented below as temporal decision trees and all multi-stage lotteries as temporal multi-stage lotteries. This is the main motivation for our concern with intrinsic preference for information or for early resolution of uncertainty. However, an agent might also be intrinsically concerned about the sequencing of information per se. She might prefer to learn some information before rather than after some (possibly degenerate) decision node, regardless of whether any significant time passes in between. Our analysis applies equally to this case, though the interpretation and plausibility of each assumed property of preferences will change. Some might prefer to use the term 'preference for early resolution' for the real time case, reserving 'preference for information' for the latter case. We do not feel strongly about this, and since both cases are reduced to the same analysis under our maintained assumptions, we use both terminologies.

Segal's (1990) analysis of preferences over two-stage lotteries is mostly concerned with the atemporal case. Among many other things, Segal considers preferences over two-stage lotteries that respect different notions of first-order stochastic dominance. Preference for information can be seen as respect for one notion of second-order stochastic dominance. Thus, at a technical level, part of our work can be seen as an extension of Segal's study.

Section 2 establishes notation and some definitions. The set-up is a simplified version of Kreps & Porteus's (1978) framework. We provide conditions on preferences over simple decision trees such that they can be rationalized by preferences over two-stage lotteries. These are similar to the dynamic consistency definitions of Karni & Schmeidler (1991) and others. We then define several possible restrictions on preferences over two-stage lotteries. Many of these are due to Segal (1990). Finally, we review Blackwell's theorem on preference for information and present a related lemma that, while it must be well-known, we could not find explicitly stated anywhere.

Section 3 discusses intrinsic preference for information in the absence of any possible instrumental purpose. Section 4 discusses the combination of intrinsic and instrumental attitudes about information, assuming some notion of dynamic consistency. Section 5 briefly considers how intrinsic preference for information might itself lead an agent to be dynamically inconsistent. An appendix contains all proofs not given in the text.

2 Notation and Groundwork

We first consider behavior within and preference among a class of simple decision trees. Let a set of payoff-relevant states be represented by the interval $[0,1]$, and let $\mathcal{L}[0,1]$ be the set of probability measures on $[0,1]$. With slight abuse of terminology, we will sometimes refer to these as distributions or lotteries. Let $\mathcal{X}$ be a set of consequences and $\mathcal{A}$ be a set of actions. We take them both to be compact convex subsets of $\mathbb{R}$: in particular, let $\mathcal{X}$ be the interval $[\underline{x}, \overline{x}]$, $\underline{x} < \overline{x}$. Let $\mathcal{C}$ be the set of measurable functions, $c : [0,1] \times \mathcal{A} \to \mathcal{X}$, that assign a consequence to each state-action pair. For any probability measure P in $\mathcal{L}[0,1]$, any action a in $\mathcal{A}$, and any consequence function c in $\mathcal{C}$, let $\widetilde{c}(P,a) \in \mathcal{L}(\mathcal{X})$ be the probability measure over consequences given by $\Pr[\widetilde{c}(P,a) \leq x] = P(\{\omega \in [0,1] : c(\omega, a) \leq x\})$ for all x in $\mathcal{X}$.[5]

We want to describe decision trees in which an agent chooses an action from a closed set after the resolution of a signal. Let $S = (s_1, \ldots, s_N)$ be the list of possible realizations of a signal, and let q_i in $(0,1]$ be the probability that the signal's realization is s_i. Let P_i in $\mathcal{L}[0,1]$ be the posterior belief over the payoff-relevant states induced by the signal realization s_i for the prior belief $\sum_i q_i P_i$. Associated with each signal realization, s_i in S, is a closed action set, $A_i \subset \mathcal{A}$, from which the agent chooses. A simple decision tree can then be denoted $\left\langle (s_i, q_i, P_i, A_i)_{i=1}^N, c \right\rangle$, where c in $\mathcal{C}$ maps the states and chosen actions to consequences. Let $\mathcal{T}$ be the set of all decision trees of the simple form $\left\langle (s_i, q_i, P_i, A_i)_{i=1}^N, c \right\rangle$.

With the weak topology, the set of lotteries $\mathcal{L}(\mathcal{X})$ is a compact metric space. So, $\mathcal{L}(\mathcal{L}(\mathcal{X}))$ is a compact metric space with the weak topology. We consider the natural linear structure of $\mathcal{L}(\mathcal{X})$. That is, if F_1 and F_2 are elements of $\mathcal{L}(\mathcal{X})$ then for any Borel subset B of $\mathcal{X}$ and any α in $[0,1]$, $\alpha F_1 + (1-\alpha)F_2$ is defined by the rule $(\alpha F_1 + (1-\alpha)F_2)(B) = \alpha F_1(B) + (1-\alpha)F_2(B)$. So for instance, if a function U on $\mathcal{L}(\mathcal{X})$ has the expected utility representation, we call it a linear function. Similarly, $\mathcal{L}(\mathcal{L}(\mathcal{X}))$ is endowed with the natural linear structure as follows: if μ_1 and μ_2 are probability measures on $\mathcal{L}(\mathcal{X})$ then for any Borel subset $\mathcal{B}$ of $\mathcal{L}(\mathcal{X})$ and any α in $[0,1]$, $\alpha\mu_1 + (1-\alpha)\mu_2$ is the measure in $\mathcal{L}(\mathcal{L}(\mathcal{X}))$ defined by the rule $(\alpha\mu_1 + (1-\alpha)\mu_2)(\mathcal{B}) = \alpha\mu_1(\mathcal{B}) + (1-\alpha)\mu_2(\mathcal{B})$. Notice, in particular that if $\mathcal{B}$ is the singleton set $\{F\}$ for some F in $\mathcal{L}(\mathcal{X})$, then $\alpha\mu_1 + (1-\alpha)\mu_2$ assigns to F the weighted sum of the probabilities assigned to it by μ_1 and μ_2.

Let $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$ denote the set of Borel measures on $\mathcal{L}(\mathcal{X})$ with finite support. We will interpret each $[(F_j, q_j)_{j=1}^M]$ in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$ as a two-stage lottery that is discrete in the first stage, where each F_j is a one-stage lottery in $\mathcal{L}(\mathcal{X})$ that can occur at the second stage and

[5] The analysis can readily be extended to outcome and action 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).

each q_j is the first-stage probability of obtaining that second-stage lottery. The (contingent) behavior of an agent in a tree T in $\mathcal{T}$, can be described by a function, b_T, that assigns an action from the available set, A_i, to each signal realization s_i in S. In any tree T, each possible behavior b_T induces a two-stage lottery over consequences, $[(\widetilde{c}(P_i, b_T(s_i)),\ q_i)_{i=1}^N]$ in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$, defined by $[(F_j,\ \sum_{i:\widetilde{c}(P_i,b_T(s_i))=F_j} q_i)_{j=1}^M]$. Let $(b_T)_{T\in\mathcal{T}}$ be a description of an agent's (contingent) behavior in all trees in $\mathcal{T}$.

Let $\succeq_{\mathcal{T}}$ denote an agent's (complete and transitive) preference relation over $\mathcal{T}$,[6] and let $\succeq_2$ denote the restriction of $\succeq_{\mathcal{T}}$ to the subset of $\mathcal{T}$ for which each set A_i is a singleton. We can view $\succeq_2$ as the agent's preference among those trees in which the agent has *committed* to her contingent actions before the signal is realized. Since the agent's behavior within each "commitment tree" is degenerate, each such tree, $\left\langle (s_i, q_i, P_i, \{a_i\}_i)_{i=1}^N, c\right\rangle$, maps directly to a two-stage lottery over consequences, $[(\widetilde{c}(P_i, a_i),\ q_i)_{i=1}^N]$. Hence, following Kreps & Porteus (1978), we interpret $\succeq_2$ as representing the agent's preferences defined on two-stage lotteries that are discrete in the first-stage. In order for these preferences to be well-defined, we need to make the assumption that if two commitment trees map to the same two-stage lottery, the agent is indifferent between them. That is, we assume throughout:

Assumption A *For all pairs of commitment trees* $T = \left\langle (s_i, q_i, P_i, \{a_i\}_i)_{i=1}^N, c\right\rangle$ *and* $\widehat{T} = \left\langle (\widehat{s}_j, \widehat{q}_j, \widehat{P}_j, \{\widehat{a}_j\}_j)_{j=1}^{\widehat{N}}, \widehat{c}\right\rangle$, *if* $[(\widetilde{c}(P_i, a_i),\ q_i)_{i=1}^N] = [(\widetilde{c}(\widehat{P}_j, \widehat{a}_j),\ \widehat{q}_j)_{j=1}^{\widehat{N}}]$ *then* $T \sim_{\mathcal{T}} \widehat{T}$.[7]

We also assume throughout that $\succeq_2$ is continuous as a relation on $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$.

We want an agent's preferences over simple decision trees to be determined by her preferences over two-stage lotteries. In particular, we assume that the agent judges each tree by the best two-stage lottery that behavior within the tree can induce.

Definition *An agent's preference relation over simple decision trees,* $\succeq_{\mathcal{T}}$, *is* freedom independent (FI) *if for all pairs of trees* $T = \left\langle (s_i, q_i, P_i, A_i)_{i=1}^N, c\right\rangle$ *and* $T' = \left\langle (s_i', q_i', P_i', A_i')_{i=1}^{N'}, c'\right\rangle$ *in* $\mathcal{T}$, $T \succeq_{\mathcal{T}} T'$ *if and only if there exists an action profile* $(a_i^*)_{i=1}^N$ *in* $A_1\times \ldots \times A_N$ *such that* $[(\widetilde{c}(P_i, a_i^*)),\ q_i)_{i=1}^N] \succeq_2 [(\widetilde{c}'(P_i', a_i'),\ q_i')_{i=1}^{N'}]$ *for all action profiles* $(a_i')_{i=1}^{N'}$ *in* $A_1'\times \ldots \times A_N'$.

Freedom independence implies that the agent does not care intrinsically about the number of choices available in a tree. If an agent's behavior within trees sometimes fails to select

[6] Implicitly, this allows the agent to have preferences over her priors. This is reasonable if we think that the "prior" is generated by some underlying objective process.

[7] Assumption A implies that the agent does not care whether a particular second stage sub-lottery occurs at signal realization s or s' provided s and s' occur with the same probability. That is, it excludes preferences that depend intrinsically on the *signal* state; contrast Johnsen & Donaldson (1984) and Sarin & Wakker (1993,1994).

the best available two-stage lottery then her preferences among trees may fail to satisfy freedom independence. If the agent anticipates that her future behavior will not accord with her current preferences, she may prefer, like Ulysses approaching the sirens, to bind herself. Following Machina (1989) and McClennen (1990), we say an agent is *resolute* if her behavior when she cannot commit is the same as that she would choose were she able to commit. That is, resolute choice imposes (dynamic) consistency between the agent's behavior within decision trees and her preferences over two-stage lotteries.

Definition *An agent with preferences over two-stage lotteries, $\succeq_2$, is* resolute *if in each tree $T = \left\langle (s_i, q_i, P_i, (A_i)_{i=1}^N, c \right\rangle$ in $\mathcal{T}$, her behavior, b_T, satisfies $[(\tilde{c}(P_i, b_T(s_i)),\ q_i)_{i=1}^N] \succeq_2 [(\tilde{c}(P_i, a_i'),\ q_i)_{i=1}^N]$ for all action profiles $(a_i')_{i=1}^N$ in $A_1 \times \ldots \times A_N$.*[8]

On its own, resolute choice is weaker than what are sometimes called fold-back or backward induction properties. Nevertheless, if we assume resolute choice, we are able to ignore problems of "weakness of will". If we assume freedom independence but not resolute choice, it is as if the agent herself is ignoring problems of "weakness of will". In either case, the agent's attitude towards information can be characterized by properties of the agent's preference relation over two-stage lotteries, $\succeq_2$.

One such property is the reduction axiom. For any two-stage lottery, $X = [(F_i, q_j)_{j=1}^M]$ in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$, let $\rho(X)$ be the reduced one-stage lottery given by $\Pr[\rho(X) \le x] = \sum_{j=1}^M q_j F_j(\{\omega \in [0,1] : c(\omega, a) \le x\})$ for all x in $\mathcal{X}$.

Definition *An agent's preference relation over two-stage lotteries, $\succeq_2$, satisfies the* reduction of compound lottery axiom (RCLA) *if for all pairs of two-stage lotteries X and X' in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$: $\rho(X) = \rho(X')$ implies $X \sim_2 X'$.*[9]

We can identify two special subclasses of two-stage lotteries: *early-resolution lotteries* in which all uncertainty is resolved in the first stage; and *late-resolution lotteries* in which no uncertainty is resolved in the first stage. Let $\succeq_{er}$ denote the restriction of two-stage lottery preferences, $\succeq_2$, to early-resolution lotteries, and let $\succeq_{lr}$ denote the restriction of $\succeq_2$ to late-resolution lotteries. Since both the set of early- and the set of late-resolution lotteries are isomorphic to the set of one-stage lotteries, $\mathcal{L}(\mathcal{X})$, the preference relations $\succeq_{er}$ and $\succeq_{lr}$ can be endowed with properties of one-stage lottery preferences such as risk-aversion, independence, betweenness, or rank-dependence. With slight abuse of terminology, we say that

[8] Resolute choice is related to Karni & Schmeidler's (1991) "dynamic consistency" and to Johnsen & Donaldson's (1984) "time consistent preference".

[9] Some authors use the term RCLA to refer only to *atemporal* two-stage lotteries.

the preferences over late-resolution lotteries are quasi-convex (respectively, quasi-concave) if $[F_1, 1] \succeq_{lr} [F_2, 1]$ implies $[F_1, 1] \succeq_{lr} [\alpha F_1 + (1-\alpha)F_2, 1]$ (respectively, $[\alpha F_1 + (1-\alpha)F_2, 1] \succeq_{lr} [F_2, 1]$) for all F_1 and F_2 in $\mathcal{L}(\mathcal{X})$ and all α in $[0, 1]$. Similarly, define quasi-convexity and quasi-concavity of $\succeq_{er}$. Both $\succeq_{er}$ and $\succeq_{lr}$ inherit continuity from $\succeq_2$. In addition, we assume throughout that both $\succeq_{er}$ and $\succeq_{lr}$ respect first-order stochastic dominance. We do not, however, identify either early- or late-resolution lotteries with their one-stage counterparts.

Following Segal (1990), we say an agent is *time neutral* if the preference relations $\succeq_{er}$ and $\succeq_{lr}$ coincide. Formally, let δ_x in $\mathcal{L}(\mathcal{X})$ denote the one stage lottery that assigns probability one to the consequence x. Then,

Definition *An agent's preference relation over two-stage lotteries satisfies* time neutrality (TN) *if for all pairs of* discrete *one-stage lotteries,* $(x_i, r_i)_{i=1}^N$, $(x'_i, r'_i)_{i=1}^{N'}$ *in* $\mathcal{L}(\mathcal{X})$*:* $[(x_i, r_i)_{i=1}^N,\ 1] \succeq_{lr} [(x'_i, r'_i)_{i=1}^{N'},\ 1]$ *if and only if* $[(\delta_{x_i},\ r_i)_{i=1}^N] \succeq_{er} [(\delta_{x'_i},\ r'_i)_{i=1}^{N'}]$. *Or, equivalently, for all discrete one-stage lotteries,* $(x_i, r_i)_{i=1}^N$ *in* $\mathcal{L}(\mathcal{X})$*:* $[(x_i, r_i)_{i=1}^N,\ 1] \sim_2 [(\delta_{x_i},\ r_i)_{i=1}^N]$.

The reduction of compound lottery axiom implies but is not implied by time neutrality.

Segal (1990) also suggests the following substitution property for preferences over two-stage lotteries.

Definition *An agent's preference relation over two-stage lotteries,* $\succeq_2$*, satisfies* compound independence (CI) *if for all pairs of two-stage lotteries of the form* $X = [(F_1, q_1;\ \ldots;\ F_j, q_j;\ \ldots; F_N, q_N)]$ *and* $Y = [(F_1, q_1;\ \ldots;\ F_{j-1}, q_{j-1};\ F'_j, q_j;\ F_{j+1}, q_{j+1};\ \ldots;\ F_N, q_N)]$ *in* $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$*, with* $q_j > 0$*:* $X \succeq_2 Y$ *if and only if* $[F_j, 1] \succeq_{lr} [F'_j, 1]$.

Notice that $(F_i)_{i=1}^N$, and F'_j in the above definition are not necessarily distinct. Thus, in particular, $[F_1, 1] \succeq_{lr} [F_1, q_1;\ F_2, q_2]$ if and only if $[F_1, 1] \succeq_{lr} [F_2, 1]$.

If the agent satisfies the reduction axiom then compound independence implies that she is an atemporal expected utility maximizer. But, without RCLA, compound independence does not imply that the agent's preferences over either early- or late-resolution lotteries satisfy independence. Compound independence is nevertheless quite strong. For example, with compound independence, resolute choice is equivalent to a fold-back or backward induction property. Formally, if an agent's preferences over two-stage lotteries, $\succeq_2$, satisfy compound independence then the agent is resolute if and only if, in each tree $T = \left\langle (s_i, q_i, P_i, A_i)_{i=1}^N, c \right\rangle$ in $\mathcal{T}$, her behavior, b_T, satisfies $[\widetilde{c}(P_i, b_T(s_i)), 1] \succeq_{lr} [\widetilde{c}(P_i, a'_i), 1]$ for all actions a' in A_i, at all s_i in S.[10] Some find this fold-back property, and hence compound independence, intuitively

[10] This fold back property is related to Kreps & Porteus's (1978, axiom 3.1) "temporal consistency",

appealing. Others, however, find it too restrictive (see Machina, 1989). We will therefore sometimes make do with the following less restrictive monotonicity properties.

Definition *An agent's preference relation over two-stage lotteries, $\succeq_2$, satisfies* conditional quasi-convexity (CQV) *if for all pairs of two-stage lotteries of the form* $X = [(F_1, q_1; \ldots; F_j, q_j; \ldots; F_N, q_N)]$ *and* $Y = [(F_1, q_1; \ldots; F_{j-1}, q_{j-1}; F'_j, q_j; F_{j+1}, q_{j+1}; \ldots; F_N, q_N)]$ *in* $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$, *with* $q_j > 0$: *if* $X \succeq_2 Y$ *then, for all* α *in (0,1),* $X \succeq_2 [(F_1, q_1; \ldots; F_{j-1}, q_{j-1}; F_j, \alpha q_j; F'_j, (1-\alpha)q_j; F_{j+1}, q_{j+1}; \ldots; F_N, q_N)]$. *Similarly, if* $\succeq_2$ *satisfies* conditional quasi-concavity (CQC) *then* $[(F_1, q_1; \ldots; F_{j-1}, q_{j-1}; F_j, \alpha q_j; F'_j, (1-\alpha)q_j; F_{j+1}, q_{j+1}; \ldots; F_N, q_N)] \succeq_2 Y$.

Although it is strictly weaker, conditional quasi-convexity does not meet all of Machina's (1989) objections to compound independence. In particular, an agent who satisfies CQV never strictly prefers to randomize. CQV is also strictly weaker than quasi-convexity. For example, compound independence implies both CQV and CQC but it does not imply (nor is it implied by) quasi-convexity or quasi-concavity of the preference relation $\succeq_2$.[11]

Blackwell's theorem applies to agents who satisfy freedom independence, reduction and compound independence. Consider preferences over trees in the subset of $\mathcal{T}$ such that the set of available actions is the same at each signal realization; that is, $A_i = A$ for $i = 1, \ldots, N$. We will call these Blackwell decision trees. Following Blackwell, we can define a signal by its (finite) list of possible realisations, S, and its likelihood function, $\lambda : S \times [0,1] \to [0,1]$, where for any ω in $[0,1]$, $\sum_{s \in S} \lambda(s|\omega) = 1$. Information is then defined as follows.

Definition *The signal* (S, λ) *is* more informative than *the signal* (S', λ') *with respect to the prior belief* π *in* $\mathcal{L}([0,1])$ *if there exists a function* $\alpha : S' \times S \to [0,1]$ *such that* $\sum_{s'} \alpha(s', s) = 1$ *for all* s *in* S, *and* $\lambda'(s'|\omega) = \sum_s \alpha(s', s)\lambda(s|\omega)$ *for* π*-almost all* ω *and all* s' *in* S'.

Given a prior belief π, a consequence function c in $\mathcal{C}$, and a fixed action set $A \subset \mathcal{A}$, the signal, (S, λ) induces the Blackwell decision tree $\left\langle (s_i, q_i, P_i, A)_{i=1}^N, c \right\rangle$ where $q_i = \int \lambda(s_i|\omega)\, \pi(d\omega)$ and P_i is the posterior induced by the signal realization s_i.[12]

LaValle & Wapman's (1985) "recursive analysis" or "rolling-back" property, and Chew & Epstein's (1989) "consistency". It is stronger than Johnsen & Donaldson's (1984) weak separability or Sarin & Wakker's (1993, 1994) folding back procedure since they allow for signal state-dependent preferences.

[11] Again with slight abuse of terminology, by quasi-convexity (respectively, quasi-concavity) of $\succeq_2$ we mean that $X \succeq_2 Y$ implies that $X \succeq_2 [\alpha X + (1-\alpha)Y]$ (respectively, $[\alpha X + (1-\alpha)Y] \succeq_2 Y$) for any measures X and Y in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$ and any α in $[0,1]$.

[12] If $q_i = 0$, simply omit s_i from the list of possible signal realizations, S.

Blackwell showed that, regardless of their prior beliefs, available action sets and consequence functions, (loosely speaking) all atemporal expected utility maximizers prefer more informative signals. Moreover, if a signal is not more informative than a second signal then there exists an expected utility maximizer who strictly prefers the second signal. Formally,

Blackwell's Theorem *For all priors π in $\mathcal{L}([0,1])$ the following two statements are equivalent.*

(i) For all consequence functions c in $\mathcal{C}$, and all action sets $A \subset \mathcal{A}$, the Blackwell decision tree induced by the signal (S, λ) is (weakly) preferred to that induced by the signal (S', λ') by all agents whose preferences over trees, $\succeq_T$, satisfy freedom independence, whose two-stage lottery preferences satisfy RCLA and whose induced preferences over reduced one-stage lotteries satisfy independence.

(ii) The signal (S, λ) is more informative than the signal (S', λ') with respect to the prior belief π.[13]

We want to extend Blackwell's analysis beyond atemporal expected utility theory. When we relax assumptions such as independence, however, we enlarge the set of preference relations under consideration. This makes it easier to find two agents who disagree which is the better of any two signals not ranked by information. Therefore, we limit attention to extending the simpler direction of Blackwell's theorem; that is, finding conditions beyond atemporal expected utility under which more informative signals are always preferred. To do this, we will make extensive use of the following fact: an increase in the informational content of a signal can be represented by a sequence of elementary mean preserving spreads, or linear bifurcations, in the distribution of induced posteriors.

Definition *A* linear bifurcation *of the distribution of posteriors $[(P_i, q_i)_{i=1}^N]$ is a distribution of posteriors of the form $[(\widehat{P}_i, \widehat{q}_i)_{i=0}^N] = [\widehat{P}_0, \beta q_1;\ \widehat{P}_1, (1-\beta)q_1;\ P_2, q_2;\ \ldots;\ P_N, q_N]$, where β is in $[0,1]$ and $P_1 = \beta\widehat{P}_0 + (1-\beta)\widehat{P}_1$.*

Lemma 2.1 *Suppose that $[(P_i, q_i)_{i=1}^N]$ and $[(P'_j, q'_j)_{j=1}^{N'}]$ are the distributions of posteriors on $[0,1]$ induced by the signals (S, λ) and (S', λ') respectively. Suppose that both distributions of posteriors have the same prior, that is, $\sum_i q_i P_i = \sum_j q'_j P'_j = \pi$. Then the following three statements are equivalent:*

(i) The signal (S, λ) is more informative than the signal (S', λ') with respect to the prior belief π.

[13] Strictly speaking, Blackwell's theorem does not require the agent to be resolute, but it is usually implicitly assumed. The theorem is often expressed without a consequence function, but with the agents allowed to have different preferences over the consequences. This version is equivalent.

(ii) *There exist weights* $\beta_{ij} \geq 0$, $i = 1, \ldots, N$, $j = 1 \ldots, N'$, *such that* $\sum_i \beta_{ij} = 1$ *for all* j, $\sum_j \beta_{ij} q'_j = q_i$ *for all* i, *and* $\sum_i \beta_{ij} P_i = P'_j$ *for* π*-almost all* ω *and all* j.

(iii) *There exists of sequence of posterior lotteries* $([(P_i^k,\, q_i^k)_{i=1}^{N^k}])_{k=1}^{K}$, *with* $[(P_i^1,\, q_i^1)_{i=1}^{N^1}] = [(P'_j,\, q'_j)_{j=1}^{N'}]$ *and* $[(P_i^K,\, q_i^K)_{i=1}^{N^K}] = [(P_i, q_i)_{i=1}^{N}]$, *such that* $[(P_i^{k+1},\, q_i^{k+1})_{i=1}^{N^{k+1}}]$ *is a linear bifurcation of* $[(P_i^k,\, q_i^k)_{i=1}^{N^k}]$ *for* $k = 1, \ldots, K-1$.

3 Preference for Information in Single-Action Trees.

In this section, we consider the agent's preference over decision trees in which she has only one action available, and the action is the same at every signal realization. We do this to focus on intrinsic attitudes towards information in the absence of any possible instrumental purpose for information. We will return to the many-action case in the next section. Let the set of Blackwell commitment trees be the set of trees of the form $\left\langle (s_i, q_i, P_i, \{a_i\}_i)_{i=1}^{N}, c \right\rangle$ where $a_i = a$ at all realizations s_i. Clearly, dynamic consistency is not an issue when considering these trees.

General Preferences.

We first consider intrinsic attitudes toward information in general, without restricting the preferences over two-stage lotteries to be of any particular expected or non-expected utility form. Lemma 2.1 allows us to interpret information in terms of the distributions of posterior beliefs induced by different signals. Using this fact, we can reduce intrinsic preference for information to a simple substitution property of an agent's preferences over two-stage lotteries.

Definition *An agent's preference relation over two-stage lotteries,* $\succeq_2$, *satisfies* single-action information loving (SAIL) *if* $Y \succeq_2 X$ *for all pairs of two-stage lotteries of the form* $X = [(F_1, q_1;\ \ldots\ ;\ F_j, q_j;\ \ldots\ ; F_N, q_N)]$ *and* $Y = [(F_1, q_1;\ \ldots\ ;\ F_{j-1}, q_{j-1};\ F'_j, \beta q_j;\ F''_j, (1-\beta) q_j;\ F_{j+1}, q_{j+1};\ \ldots\ ; F_N, q_N)]$ *in* $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$, *with* $q_j > 0$, β *in* $[0,1]$, *and* $F_j = \beta F'_j + (1-\beta) F''_j$. *Similarly, if her preference relation,* $\succeq_2$, *satisfies* single-action information aversion (SAIA) *then* $X \succeq_2 Y$, *and if it satisfies* single-action information neutrality (SAIN) *then* $X \sim_2 Y$.[14]

[14] Single-action information neutrality corresponds to Kreps & Porteus's (1979) "resolution neutrality", and Chew & Epstein's (1989) "time indifference". Notice that the lotteries $(F_i)_{i=1}^{N}$, F'_j and F''_j need not be distinct.

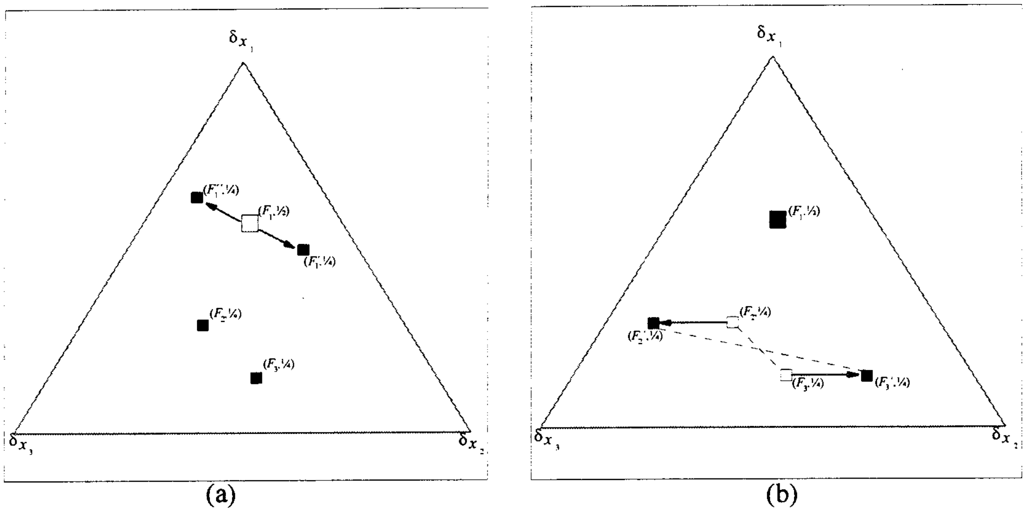

Figure 1

This substitution property has a natural geometric interpretation. Figure 1a represents a pair of two-stage lotteries, $(F_1, \frac{1}{2};\ F_2, \frac{1}{4};\ F_3, \frac{1}{4})$ and $(F_1', \frac{1}{4};\ F_1'', \frac{1}{4};\ F_2, \frac{1}{4};\ F_3, \frac{1}{4})$, by the distributions they induce on the relevant simplex of late-resolution lotteries with outcome support $(x_i)_{i=1}^3$. The two distributions differ by a simple linear bifurcation and hence are ranked by SAIL. The pair of two-stage lotteries in Figure 1b, $(F_1, \frac{1}{2};\ F_2, \frac{1}{4};\ F_3, \frac{1}{4})$ and $(F_1, \frac{1}{2};\ F_2', \frac{1}{4};\ F_3', \frac{1}{4})$, have the same mean and the second looks more spread out than the first, but it can not be reached from the first by a sequence of simple linear bifurcations. In particular, the second stage lotteries F_2, F_3, F_2' and F_3' are not colinear. Hence, the two-stage lotteries in Figure 1b are not ranked by SAIL.

Although SAIL is thus a discontinuous property, it is exactly what we need to describe intrinsic preference for information.

Proposition 3.1 *The following are equivalent:*

(i) For all priors π in $\mathcal{L}([0,1])$, all c in C, and all single actions a in $\mathcal{A}$, if the signal (S, λ) is more informative than the signal (S', λ') with respect to the prior π, then the Blackwell commitment tree induced by (S, λ) is weakly preferred (respectively, not preferred / indifferent) to that induced by (S', λ').

(ii) The agent's preference relation over two-stage lotteries, $\succeq_2$, satisfies SAIL (respectively, SAIA / SAIN).

This proposition has some immediate consequences. Single-action information neutrality is just the reduction axiom. Similarly, time neutrality can be thought of as intrinsic indifference between complete early resolution of uncertainty and no early resolution. Hence, if the agent always (weakly) prefers more information but is time neutral then she is intrinsically indifferent toward all information revelation, partial as well as complete. That is,

Corollary 3.1.1 *The following properties of a preference relation over two-stage lotteries are equivalent: (i) RCLA; (ii) SAIN; and (iii) TN and either SAIL or/and SAIA.*

The equivalence of information neutrality and the reduction axiom was discussed by Kreps & Porteus (1978) and Chew & Epstein (1989). The second equivalence, while simple, is new as far as we know. Corollary 3.1.1 helps to clarify how some previous work on preference for information contrasts with ours. Both Wakker (1987) and Schlee (1991) choose to focus on instrumental preference for information. They both implicitly assume away the possibility that information might have intrinsic value. Wakker's analysis does this by assuming RCLA, which by Corollary 3.1.1, implies that the agent is information-neutral in the absence of future choices. Schlee's analysis assumes time neutrality, which by Corollary 3.1.1, implies that the agent cannot satisfy SAIL or SAIA without being information-neutral in the absence of future choices.

Proposition 3.1 also provides a new axiomatization for the standard model of individual choice. Some might argue that, regardless of the number of actions available to her, an agent should always prefer more to less information. Thus in particular, the agent should satisfy SAIL. They might also argue that preferences over early- and late-resolution lotteries should coincide. If they also argue for compound independence then they have a normative justification for the atemporal expected utility model.

Corollary 3.1.2 *A preference relation over two-stage lotteries satisfies CI, TN and either SAIL or/and SAIA if and only if it satisfies RCLA and the induced preferences over reduced one-stage lotteries satisfy independence.*

This follows from combining Corollary 3.1.1 with Kreps & Porteus's (1978) corollary 3, or Chew & Epstein's (1989) Theorem 1, or Segal's (1990) Theorem 3(a).

The pair of two-stage lotteries in Figure 1a ranked by SAIL look a lot like two multivariate distributions ranked by increasing risk. In an earlier paper (Grant, Kajii & Polak, 1992b), we defined a substitution property that we called many-good bifurcation risk aversion. Loosely speaking, single-action information aversion is analogous to this notion of many-good risk aversion, where the multi-dimensional commodity space is replaced by a multi-dimensional simplex representing second-stage distributions over consequences. If we identify the payoff-relevant states with the set of possible consequences, we can think of the multi-dimensional simplex as representing the set of possible posteriors on those states. That is, we can think of single-action information loving as risk loving over posteriors; information aversion as risk aversion over posteriors; and information neutrality as risk neutrality over posteriors. Similarly, compound independence is analogous to the substitution property that we called degenerate independence (Grant, Kajii & Polak, 1992a).

This analogy suggests that some results about attitudes towards risk can be translated into results about attitudes towards information. For example, just as many-good bifurcation risk aversion and degenerate independence imply that the ordinal preferences on commodity bundles are quasi-concave, so single-action information loving and compound independence imply that the preferences over late-resolution lotteries are quasi-convex. Indeed, we do not need compound independence for this result: conditional quasi-convexity is enough.

Proposition 3.2 *If an agent's preference relation over two-stage lotteries satisfies* CQV *and* SAIL *(respectively,* CQC *and* SAIA*) then her preferences for late-resolution lotteries,* $\succeq_{lr}$*, are quasi-convex (respectively, quasi-concave) in the probabilities.*

Proof. For all pairs of discrete one-stage lotteries F, F' in $\mathcal{L}(\mathcal{X})$ such that $[F,1] \succeq_{lr} [F',1]$, CQV implies that $[F,1] \succeq_2 [F,(1-\alpha);\ F',\alpha]$ for all α in $(0,1)$. SAIL implies $[F,(1-\beta);\ F',\beta] \succeq_2 [(1-\beta)F + \beta F',\ 1]$. Hence, $[F,1] \succeq_{lr} [(1-\beta)F + \beta F',\ 1]$. The proof for CQC and SAIA is similar. ■

Corollary 3.2.1 *If an agent's preference relation over two-stage lotteries satisfies* CI *and* SAIL *(respectively,* SAIA*) then her preferences for late-resolution lotteries are quasi-convex (respectively, quasi-concave) in the probabilities.*

Corollary 3.2.1 concerns single-action preference for information and does not depend on assuming any other property of preferences over either early- or late-resolution lotteries. Kreps & Porteus (1979) and Machina (1984), however, show that if an agent has more than one action available and is a temporal expected utility maximizer then the *induced* preferences over lotteries are quasi-convex. Thus, loosely speaking, both intrinsic and instrumental preference for information imply quasi-convexity.

The analogy with our notion of many-good bifurcation risk aversion is not the only connection between attitudes towards risk and attitudes towards information. One measure of an agent's attitude towards a particular risk is her certainty equivalent for that risk. Formally, for any one-stage lottery F in $\mathcal{L}(\mathcal{X})$, its late-resolution certainty equivalent,

$\mathrm{CE}_{lr}(F)$, is given by $[(\mathrm{CE}_{lr}(F), 1),\ 1] \sim_{lr} [F,\ 1]$. For any discrete one-stage lottery $F = (x_i, r_i)_{i=1}^{N}$ in $\mathcal{L}_0(\mathcal{X})$, its early-resolution certainty equivalent, $\mathrm{CE}_{er}(F)$, is given by $[(\delta_{\mathrm{CE}_{er}(F)}, 1)] \sim_{er} [(\delta_{x_i}, r_i)_{i=1}^{N}]$; and for any discrete one-stage lottery $F' = (x'_i, r'_i)_{i=1}^{N'}$ in $\mathcal{L}_0(\mathcal{X})$, and any α in $(0, 1]$, the early-resolution conditional certainty equivalent of F given F' and α, $\mathrm{CCE}_{er}(F; F', \alpha)$, is given by $[(\delta_{\mathrm{CCE}_{er}(F)}, \alpha; \delta_{x'_1}, (1-\alpha)r'_1; \ldots; \delta_{x'_{N'}}, (1-\alpha)r'_{N'})] \sim_{er} [\delta_{x_1}, \alpha r_1; \ldots; \delta_{x_N}, \alpha r_N; \delta_{x'_1}, (1-\alpha)r'_1; \ldots; \delta_{x'_{N'}}, (1-\alpha)r'_{N'})]$.

Since both the preferences over early- and over late-resolution lotteries satisfy continuity and first-order stochastic dominance, we can find both the early- and late-resolution certainty equivalents for any one-stage lottery. Preference for information places restrictions on how early- and late-resolution certainty equivalents compare. That is, it places at least some restriction on how the preferences over early- and late-resolution lotteries assess similar risks.

Proposition 3.3 *If a preference relation over two-stage lotteries, $\succeq_2$, satisfies SAIL (respectively, SAIA) then:*

(i) for all discrete one-stage lotteries F in $\mathcal{L}_0(\mathcal{X})$, $\mathrm{CE}_{lr}(F) \leq (\geq)\ \mathrm{CE}_{er}(F)$;

(ii) if, in addition, $\succeq_2$ satisfies CI then, for all pairs of discrete one-stage lotteries F and F' in $\mathcal{L}_0(\mathcal{X})$, and all α in $(0, 1]$, $\mathrm{CE}_{lr}(F) \leq (\geq)\ \mathrm{CCE}_{er}(F; F', \alpha)$;

(iii) if, in addition, $\succeq_2$ satisfies CI and $\succeq_{er}$ is weakly risk-averse (respectively, risk-loving) then, $\mathrm{CE}_{lr}(\cdot)$ is convex (respectively, concave).

Proposition 3.3(i) implies that, if an agent always prefers more information to less, then preferences over early-resolution lotteries cannot be more risk-averse than her preferences over late-resolution lotteries. This result does not depend on any other property of the agent's preferences, not even compound independence. Adding compound independence in part (ii) of the Proposition, allows us to strengthen the conclusion. If we also add that the agent's preferences over early-resolution lotteries are risk-averse then her late-resolution certainty-equivalent function must be convex.

Smooth, Recursive Preferences.

Preferences that satisfy compound independence are sometimes referred to as having a recursive form. Given compound independence, we can use certainty equivalents to break the evaluation of two-stage lotteries into two parts, where the first part only uses $\succeq_{lr}$ and the second only uses $\succeq_{er}$. First, we replace the second-stage lotteries in a two-stage lottery by their late-resolution certainty equivalents. That is, given CI, any two-stage lottery $X = [(F_i, q_i)_{i=1}^{N}]$ is indifferent to the early-resolution lottery $[(\delta_{\mathrm{CE}_{lr}(F_i)}, q_i)_{i=1}^{N}]$. We then evaluate this lottery using $\succeq_{er}$. This well-known recursive method shows that CI is compatible with any specific forms of preference we choose for $\succeq_{er}$ and $\succeq_{lr}$. Indeed the proofs of parts (ii) and (iii) of Proposition 3.3 already used this method.

We next consider recursive preferences in which the preferences over early-resolution lotteries are smooth. The smoothness assumption is in the spirit of Machina's (1982) generalized expected utility theory, except that where he assumed Frechet differentiability, we only need assume Gateaux differentiability. With slight abuse of notation, let G denote both a lottery in $\mathcal{L}(\mathcal{X})$ and its associated cumulative distribution function on $\mathcal{X}$. Suppose that an agent's preferences over early-resolution lotteries are represented by the function $V_{er} : \mathcal{L}_0(\mathcal{X}) \rightarrow \mathbb{R}$. We say that V_{er} is Gateaux differentiable if there exists a function $u_{er} : \mathcal{X} \times \mathcal{L}_0(\mathcal{X}) \rightarrow \mathbb{R}$ such that for any pair of lotteries G and G' in $\mathcal{L}_0(\mathcal{X})$, $V(\alpha G' + (1-\alpha)G) - V(G) = \int u(x;G)(\alpha G'(dx) - \alpha G(dx)) + o(\alpha)$. We will sometimes refer to $u_{er}(\cdot, F)$ as the local utility function.

Machina (1982) exploits the fact that Frechet differentiable preferences over one-stage lotteries can be approximately represented, at least locally, by expected utility functionals. Chew, Karni & Safra (1987) extend this idea to Gateaux differentiable preferences. We use this method to provide a necessary and sufficient condition for an agent with recursive smooth preferences always to exhibit intrinsic preference for information.

Proposition 3.4 *Suppose that an agent satisfies* CI *and that her preference relation over early-resolution lotteries, $\succeq_{er}$ can be represented by a Gateaux differentiable utility function with local utility functions $u_{er}(\cdot, G)$. Then the agent satisfies* SAIL *(respectively,* SAIA*) if and only if the compound function $u_{er}(\mathrm{CE}_{lr}(\cdot), G)$ from the set of late-resolution lotteries to the reals is convex (respectively, concave) for all lotteries G in $\mathcal{L}(\mathcal{X})$.*

Corollary 3.4.1 *Suppose that an agent satisfies* CI *and that her preference relation over early-resolution lotteries, $\succeq_{er}$ can be represented by a Gateaux differentiable utility function. Then* SAIL *implies that her preferences over late-resolution lotteries are convexifiable; that is, $\succeq_{lr}$ has a convex utility representation.*

The conclusion of Corollary 3.4.1 is stronger than that of Corollary 3.2.1 since not all quasi-convex functions are convexifiable. This observation will be relevant below. Moreover, Corollary 3.4.1 differs from Proposition 3.3 part (iii) in that it replaces the assumption that preferences over early-resolution lotteries are risk-averse with the assumption that they are smooth.

The idea of the proof of Proposition 3.4 further exploits the analogy between SAIL and multi-dimensional risk aversion. Recall that we can think of SAIL as risk loving where the commodity space is replaced by the simplex of posteriors. Just as Machina's (1982) local expected utility functions map from commodities to the reals, so the compound functions $u_{er}(\mathrm{CE}_{lr}(\cdot), G)$ map from posteriors to the reals. Machina showed that an agent is globally risk averse if and only if all her local expected utility functions are concave; that is, the agent is everywhere locally risk averse. Similarly, if a compound function $u_{er}(\mathrm{CE}_{lr}(\cdot), G)$ is convex, we can think of the agent as locally information loving.

Particular Recursive Forms of Preference.

Temporal Expected Utility. We next explore the implications of intrinsic attitudes towards information, where the agent's preferences over early- and late-resolution lotteries both satisfy the same property of one-stage lottery preferences. Given compound independence, if we assume that preferences for early- and late-resolution lotteries both satisfy (one-stage lottery) independence then the agent's preferences over two-stage lotteries satisfy Kreps & Porteus's (1978) temporal (von Neumann-Morgenstern) expected utility. Proposition 3.4 already resembles a 'local version' of Kreps & Porteus's Theorem 3. The following result can be obtained as a corollary of their Theorem or, if we assume Gateaux differentiability, as a corollary of our Proposition. Since it is instructive, however, we provide a direct proof. Given temporal expected utility, the agent satisfies SAIL if and only if her preferences for late-resolution lotteries are more risk-averse than those for early-resolution lotteries. Formally,

Proposition 3.5 (Kreps & Porteus 1978) *Suppose that an agent's preference relation over two-stage lotteries, $\succeq_2$, satisfies* CI, *and that her preference relations over early- and late-resolution lotteries, $\succeq_{er}$ and $\succeq_{lr}$, satisfy independence. Let u_{er} and u_{lr} be von Neumann-Morgenstern utility indexes for $\succeq_{er}$ and $\succeq_{lr}$ respectively. Then, the agent satisfies SAIL (respectively, SAIA) if and only if $u_{er} \circ u_{lr}^{-1}$ is convex (respectively, concave).*

Proof. With slight abuse of notation, for all F in $\mathcal{L}(\mathcal{X})$, let $U_{lr}(F) = \int_{x\in\mathcal{X}} u_{lr}(x)F(dx)$. Notice that, for all F in $\mathcal{L}(\mathcal{X})$, $\mathrm{CE}_{lr}(F) = u_{lr}^{-1}(U_{lr}(F))$. Let $X = [(F_1, q_1;\ \ldots\ ; F_j, q_j;\ \ldots\ ; F_N, q_N)]$ and $Y = [(F_1, q_1;\ \ldots\ ; F_{j-1}, q_{j-1};\ F'_j, \beta q_j;\ F''_j, (1-\beta)q_j;\ F_{j+1}, q_{j+1};\ \ldots\ ; F_N, q_N)]$ in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$, where $q_j > 0$ and $F_j = \beta F'_j + (1-\beta)F''_j$. Since $\succeq_{lr}$ satisfies independence, $\beta U_{lr}(F'_j) + (1-\beta)U_{lr}(F''_j) = U_{lr}(F_j)$. Given CI and that $\succeq_{er}$ satisfies independence, $X \succeq_2 Y$ if and only if $\beta u_{er}(u_{lr}^{-1}(U_{lr}(F'_j))) + (1-\beta)u_{er}(u_{lr}^{-1}(U_{lr}(F''_j))) \geq u_{er}(u_{lr}^{-1}(U_{lr}(F_j))) = u_{er} \circ u_{lr}^{-1}[\beta U_{lr}(F'_j) + (1-\beta)U_{lr}(F''_j)]$. Hence, SAIL if and only if $u_{er} \circ u_{lr}^{-1}$ is convex. The proof for SAIA is similar. ■

If both u_{er} and u_{lr} are twice differentiable, then $u_{er} \circ u_{lr}^{-1}$ is convex if and only if the Arrow-Pratt coefficient of absolute risk aversion for $\succeq_{lr}$ is greater than that for $\succeq_{er}$. Thus, an intuition for Proposition 3.5 is as follows. Suppose that an agent is more risk-averse for late- than for early-resolution lotteries. For the case of temporal expected utility, this is the same as saying that she dislikes risk more the longer she has to wait for its resolution. Hence, she will prefer early resolution; that is, she is SAIL.

This intuition seems so natural that it suggests the result might extend beyond extended utility. That is, one might conjecture that, regardless of their particular form, preferences over early-resolution lotteries are less risk-averse than those for late-resolution lotteries if and only if the agent satisfies SAIL. This conjecture, however, turns out to be false. It follows from Proposition 3.3 that $\mathrm{CE}_{lr}(F) \leq \mathrm{CE}_{er}(F)$ for all discrete one-stage lotteries F if and only if the agent intrinsically prefers full information to no information. Thus, in particular, SAIL implies that $\succeq_{er}$ is not more risk-averse than $\succeq_{lr}$. But, this ranking of certainty equivalents does not imply that the agent is more risk-averse for late- than for early-resolution lotteries. Loosely speaking, this would require a ranking of all conditional certainty equivalents. Given independence, certainty equivalents and conditional certainty equivalents are equal, but this does not generalize beyond expected utility.

To summarize: SAIL implies that the agent would pay more for the complete elimination of a late-resolving risk than for the complete elimination of the equivalent early-resolving risk. But, SAIL is not sufficient to rank the preferences over early- and late-resolution lotteries by their attitudes to all decreases in risk, whether partial or complete. Similarly, if preferences over late-resolution lotteries are more risk-averse than those over early-resolution lotteries, then the agent prefers complete information to no information. But, this is not sufficient for the agent to prefer any increase in information, whether partial or complete.

Temporal Rank-Dependence. An example that illustrates both directions of the failure of the conjecture is Quiggin's (1982) and Yaari's (1987) rank-dependent expected utility model. With slight abuse of notation, let G and G' represent the *decumulative* distributions functions of any pair of one-stage lotteries in $\mathcal{L}(\mathcal{X})$ and the lotteries themselves. Let u_{lr} and g_{lr} be strictly increasing functions $u_{lr} : \mathcal{X} \to \mathbb{R}$, and $g_{lr} : [0,1] \to [0,1]$ such that $g_{lr}(0) = 0$ and $g_{lr}(1) = 1$. For any such functions, u_{lr} and g_{lr}, define the associated RDEU functional by $V_{lr}(F) := -\int u_{lr}(x)(g_{lr} \circ G(dx))$. The preference relation over late-resolution lotteries satisfies rank-dependent expected utility (RDEU) if it can be represented by an RDEU functional V_{lr}. By analogy, define the RDEU functional $V_{er}(G)$ for similarly defined functions u_{er} and g_{er} where G is a discrete one-stage lottery in $\mathcal{L}(\mathcal{X})$. That is, for $G = (x_i, r_i)_{i=1}^{N}$, where without loss of generality $x_i \leq x_{i-1}$ for $i = 2, \ldots, N$, let $V_{er}(F) = \sum_{i=1}^{N}(u_{er}(x_i)\ [g_{er}(\sum_{j=0}^{i} r_i)$ $-\ g_{er}(\sum_{j=0}^{i-1} r_i)])$ where we set $r_0 := 0$. The preference relation over early-resolution lotteries satisfies RDEU if it can be represented by an RDEU functional V_{er}.

The RDEU representation has been popular in applications because, loosely speaking, it maintains the separability of outcomes and probabilities when assessing lotteries, while also allowing for some observed violations of the independence axiom. For example, RDEU does not require indifference sets in the simplex to be linear. We can think of the function g as a "distortion" of decumulative probabilities just as the function u distorts outcomes. If, for example, $g(G(x)) < G(x)$ for all x in $\mathcal{X}$ (for example, if g is convex) then the agent acts as if she is a pessimist; that is, she underweights the probability of high outcomes.

Given compound independence, if the agent's preference for early- and late-resolution lotteries both satisfy RDEU then we say that the agent's preferences over two-stage lotteries satisfy temporal RDEU. The following proposition provides both necessary and sufficient conditions for a temporal RDEU agent always to exhibit preference for information. These conditions are very restrictive.

Proposition 3.6 *Suppose that an agent's preference relation over two-stage lotteries, $\succeq_2$, satisfies* CI, *and that her preferences for early- and late-resolution lotteries, $\succeq_{er}$ and $\succeq_{lr}$, both satisfy* RDEU, *with functions u_{er}, g_{er}, u_{lr} and g_{lr} respectively. Suppose further that the corresponding functions, V_{er} and V_{lr}*[15] *, are both Gateaux differentiable on $\mathcal{L}(\mathcal{X})$. Then,*

(i) if the agent satisfies SAIL *then g_{lr} is convex, g_{er} is concave, and $u_{er} \circ u_{lr}^{-1} \circ g_{lr}$ is convex;*

(ii) if g_{lr} is convex, g_{er} is concave, and $u_{er} \circ u_{lr}^{-1}$ is convex then the agent satisfies SAIL*;*

(iii) if g_{lr}'''' is continuous, g_{lr} is convex and g_{lr}'/g_{lr}'' is concave, and g_{er} is concave and $u_{er} \circ u_{lr}^{-1} \circ g_{lr}$ is convex, then the agent satisfies SAIL.

Necessary and sufficient conditions for SAIA *are similar, mutatis mutandis.*

The idea of the proof of Proposition 3.6 is to apply Corollary 3.2.1 and Proposition 3.4 to the special case of temporal rank-dependent preferences. For example, rank-dependent preferences are quasi-convex if and only if the distortion function g is convex. Notice that, given Gateaux differentiability, Parts (i) and (ii) of Proposition 3.6 imply Proposition 3.5 as a special case.

Proposition 3.6 confirms that our earlier conjecture about information loving and the relative attitude toward risk of an agent's preferences over early- and late-resolution lotteries is false. Applying Chew, Karni and Safra's (1987) Theorem 1, we get that the preference over early-resolution lotteries are less risk-averse than those for late-resolution lotteries if and only if $u_{er} \circ u_{lr}^{-1}$ is convex and $g_{er} \circ g_{lr}^{-1}$ is concave. Proposition 3.6(i) shows that this

[15] Strictly speaking, we need only assume that V_{er} is Gateaux differentiable and that u_{lr} and g_{lr} are continuous.

is not sufficient for the agent to be information-loving. In particular, g_{er} itself has to be concave; that is, the preference relation $\succeq_{er}$ has to be quasi-concave. Proposition 3.6(iii) shows that less risk aversion for early-resolution lotteries is also not necessary for SAIL. In particular, $u_{er} \circ u_{lr}^{-1}$ does not have to be convex. For example, an agent with the temporal RDEU preferences given by $u_{er}(x) = x$, $u_{lr}(x) = x^2$, $g_{er}(p) = p$, and $g_{lr}(p) = p^3$ is a (strict) single-action information lover. In this example, $\succeq_{er}$ are risk neutral but $\succeq_{lr}$ are not risk-averse.

In his discussion of RDEU preferences and attitudes toward information, Schlee (1990) concludes that an agent with temporal RDEU preferences always (weakly) likes complete information revelation. Proposition 3.6 shows, however, that this result depends on time neutrality. Schlee provides an example of an agent who is temporal RDEU and who sometimes likes and sometimes dislikes information. However, if we want a temporal RDEU agent both to be risk averse and always to like any increase in information, partial or complete, then the agent's preferences over two-stage lotteries can only deviate from temporal expected utility in the second stage. To see this, notice that a consequence of Proposition 3.6(i) is that if an agent always likes information then her distortion function g_{er} must be concave and hence her preferences over early-resolution lotteries must quasi-concave. We know from Chew, Karni & Safra (1987) that if these preferences are risk-averse then they are quasi-convex. But, the only RDEU functional that is both quasi-convex and quasi-concave is expected utility. That is,

Corollary 3.6.1 *Suppose that an agent's preference relation over two-stage lotteries, $\succeq_2$, are* temporal RDEU *(with the corresponding functions, V_{er} and V_{lr}, both Gateaux differentiable on $\mathcal{L}(\mathcal{X})$), and that her preference relation on early-resolution lotteries, $\succeq_{er}$, is weakly risk-averse (respectively, loving). Then, if the agent satisfies* SAIL *(*SAIA*), the preference relation $\succeq_{er}$ satisfies independence.*

Preference for information also has implications for temporal RDEU preferences defined on more than two stage lotteries. These are discussed below.

Temporal Betweenness. A second popular alternative to expected utility is the betweenness model introduced by Dekel (1986) and Chew (1983, 1989). A betweenness preference relation conforms to expected utility within any given indifference set. That is, as with an expected utility preference relation, indifference sets are 'linear' (or more precisely, hyperplanes) in the simplex, but they need not be parallel. More formally, the preference relation over late-resolution lotteries satisfies the betweenness property if for all pairs of lotteries F, F' in $\mathcal{L}(\mathcal{X})$ and α in $(0,1)$, $[F, 1] \sim_{lr} [F', 1]$ implies $F \sim_{lr} [\alpha F + (1-\alpha)F', 1]$. Or, equivalently, if it can be represented by a functional $V_{lr} : \mathcal{L}(\mathcal{X}) \to \mathbb{R}$, implicitly defined as: $V_{lr}(F) = \int v_{lr}(x, V_{lr}(F))F(dx)$, where v_{lr} is a continuous function from $\mathcal{X} \times \mathbb{R}$ to $\mathbb{R}$, strictly increasing in its first argument. By analogy, the preference relation over early-resolution lotteries satisfies the betweenness property if there exists an appropriately defined function v_{er} such that the relation can be represented by the implicitly defined functional $V_{er}((x_i, r_i)_{i=1}^N) = \sum_{i=1}^N r_i \cdot v_{er}(x_i, V_{er}((x_i, r_i)_{i=1}^N)$ for all $(x_i, r_i)_{i=1}^N$ in $\mathcal{L}_0(\mathcal{X})$.

Given compound independence, if the agent's preference for early- and late-resolution lotteries both satisfy the betweenness property then we say that the agent's preferences over two-stage lotteries satisfy Chew & Epstein's (1989) temporal betweenness. Chew & Epstein (1989) show that, given compound independence, only temporal betweenness preferences satisfy a particular weakening of SAIN such that if $[F, 1] \sim_{lr} [F', 1]$ then $[F, \alpha; F', 1 - \alpha] \sim_2 [\alpha F + (1 - \alpha)F', 1]$. Temporal betweenness preferences have been employed by Epstein & Zin (1989) to analyse intertemporal consumption and asset pricing. They left as an open question, what preference for information implies for this class of preference. The following proposition provides necessary and sufficient conditions for a temporal betweenness agent always to exhibit preference for information. As was the case with temporal RDEU, these conditions are very restrictive.

Proposition 3.7 *Suppose that an agent's preference relation over temporal two-stage lotteries, $\succeq_2$, satisfies* CI*; and that her preference relations over early- and late-resolution lotteries, $\succeq_{er}$ and $\succeq_{lr}$, both satisfy the betweenness property, with functions v_{er} and v_{lr} respectively. Suppose further that the representation function V_{er} is Gateaux differentiable on $\mathcal{L}(\mathcal{X})$. Then the following two statements are equivalent.*

(i) The agent satisfies SAIL.

(ii) The preference relation for late-resolution lotteries, $\succeq_{lr}$, satisfies independence with a von Neumann-Morgenstern utility index u_{lr}, and $v_{er}(u_{lr}^{-1}(\cdot), \overline{v})$ is convex for all $\overline{v}$ in the range of V_{er}.

Necessary and sufficient conditions for SAIA are similar, mutatis mutandis.

The idea of the Proof of Proposition 3.7 is as follows. For (i) implies (ii), recall from the corollary of Proposition 3.4 that the preference relation $\succeq_{lr}$ must be convexifiable. But, the only betweenness preferences that are convexifiable also satisfy independence. The proof for (ii) implies (i) is analogous to Dekel's (1986) discussion of risk aversion for betweenness preferences.

The assumption of Gateaux differentiability seems very weak. For instance, it follows if v_{er} is differentiable with respect to the second argument, and the derivative $\frac{\partial}{\partial u} v_{er}(x, u)$ is integrable with respect to the first argument (see Chew & Nishimura (1992)). Nevertheless, it might be a restriction in some contexts. Gateaux differentiability can be dispensed with, however, if the early-resolution preferences are (weakly) risk averse, since the convexifiability of $\succeq_{lr}$ is then implied by Proposition 3.3(iii).

To put it differently, Proposition 3.7 shows that, if we require an agent to be risk-averse and always to like information, then temporal betweenness implies that the agent's preferences over two-stage lotteries can only deviate from temporal expected utility at the first stage. This confirms Epstein & Zin's (1989) conjecture that it might be difficult to separate attitudes toward risk from attitudes toward information. It also at least qualifies Chew & Ho's (1994) suggestion that relaxing temporal expected utility to temporal betweenness allows greater flexibility in combining attitudes toward risk and attitudes toward information. Chew & Epstein (1989) define a "timing probability premium" to measure the degree of preference for information (or early resolution) for temporal betweenness preferences. Their Theorem 4 shows that if this premium is constant then the preference for late-resolution lotteries satisfies independence. Proposition 3.7, however, shows that, given Gateaux differentiability, we do not need the premium to be constant. All we need is for the premium always to be non-negative. Preference for information also has implications for temporal betweenness preferences defined on more than two stage lotteries. These are discussed below.

A less restrictive example: Quadratic Utility. For each of the forms of recursive preferences we have examined so far, temporal expected utility, temporal rank dependence and temporal betweenness, imposing both risk aversion and preference for information implied that the preferences over either early- or late-resolution lotteries (or both) satisfy the independence axiom. Moreover, in each case, given preference for information, the preferences over early-resolution lotteries were quasi-concave. However, neither of these results are general. The following is an example of a form of non-expected utility preferences that satisfy risk aversion and preference for information but do not satisfy independence in either stage. In fact, preferences are strictly quasi-convex in both stages.

Example 1 *Assume compound independence and, for concreteness take* $X = [0,1]$. *Let preferences over early-resolution lotteries be represented by* $W_{er}(F) = \frac{1}{2}[\int x^{\frac{1}{2}} F\,(dx) + (\int x^{\frac{1}{4}} F(dx))^2]$ *and let preferences over late-resolution lotteries be represented by* $W_{lr}(F) = \frac{1}{2}[\int x^{\frac{1}{4}} F\,(dx) + (\int x^{\frac{1}{8}} F(dx))^2]$.

Given Proposition 3.4, to show that these preferences exhibit SAIL, it is enough to show that $u_{er}(\mathrm{CE}_{lr}(\cdot); G)$ is convex for all G in $\mathcal{L}(\mathcal{X})$. The local utility function of W_{er} is given by $u_{er}(x; G) = \frac{1}{2}x^{\frac{1}{2}} + [\int y^{\frac{1}{4}} G(dy)] \times x^{\frac{1}{4}}$. The certainty equivalent function derived from the preference for late-resolution lotteries is given by $\mathrm{CE}_{lr}(F) = h_{lr} \circ W_{lr}(F)$, where $h_{lr}(z) = z^4$. Hence the compound function $u_{er}(h_{lr}(\cdot); G)$ from $[0,1]$ to the reals is convex. Since $W_{lr}(\cdot)$ is convex for probability mixtures, the result follows.

The functions W_{er} and W_{lr} are examples of Chew, Epstein & Segal's (1991) quadratic utility, hence example 1 is a recursive or temporal quadratic utility model. Recall that increasing information is analogous to increasing risk in the space of posteriors. Recall also that preference for information is related to quasi-convexity of preferences over late resolution lotteries. One motivation for Chew, Epstein and Segal's model is that it allows attitudes toward risk to be separated from quasi-convexity and quasi-concavity in the probabilities; in particular, preferences can be both risk averse and quasi-concave. Both the rank-dependence and betweenness models (and indeed all ordinal independence models) lack this flexibility. It is this flexibility that allows the temporal quadratic utility model to incorporate both intrinsic preference for information and violations of the independence axiom at both stages. We know from Proposition 3.3, however, that there must still be some connection between attitudes towards risk and attitudes towards information. Indeed, in example 1, the preferences over early-resolution lotteries are less risk averse than those over late-resolution lotteries.

Preferences over Extended Lotteries.

So far we have focused on two-stage lotteries and hence implicitly on simple decision trees, but the analysis can readily be extended to preferences over longer finite period lotteries. Let $\succeq_L$ denote the agent's preferences over lotteries of length L that are simple in all but the last stage, $\mathcal{L}_0^{L-1}(\mathcal{L}(\mathcal{X}))$. As before, we can think of these as derived from preferences over commitment trees of the appropriate length. For each l in $\{1, \ldots, L\}$, let $\succeq_{L,l}$ denote the restriction of $\succeq_L$ to L-stage lotteries that are degenerate in all but the lth stage. As with preferences over early- and late-resolution lotteries above, these preferences can be endowed with properties of preferences over one-stage lotteries. We assume that, for all L and all l, $\succeq_{L,l}$ respects first-order stochastic dominance. For each l in $\{1, \ldots, L\}$, let $\mathcal{F}_l^L \subset \mathcal{L}_0^{L-1}(\mathcal{L}(\mathcal{X}))$ be the subset of L-stage lotteries that are degenerate in the first $l-1$ stages. Thus, for example, $\mathcal{F}_1^L = \mathcal{L}_0^{L-1}(\mathcal{L}(\mathcal{X}))$. Let $\succeq_{\mathcal{F}_l^L}$ denote the restriction of $\succeq_L$ to L-stage lotteries that are degenerate in the first $l-1$ stages. Thus, for example, $\succeq_{\mathcal{F}_L^L} = \succeq_{L,L}$. Abusing notation, for each l in $\{1, \ldots, L-1\}$, let $F^l = [(F_i^{l+1}, q_i)_{i=1}^N]$ denote a typical element of $\mathcal{F}_l^L$ where each F_i^{l+1} is an element of $\mathcal{F}_l^{L+1}$. We can then extend the definitions of compound independence and of single-action information loving to L-stage lotteries.

Definition *For any finite length of lotteries $L \geq 2$, we say that an agent's preference relation over L-stage lotteries, $\succeq_L$, satisfies:*

(i) extended compound independence (ECI) *if for all* l *in* $\{1, \ldots, L-1\}$*, all pairs of* L*-stage lotteries of the form* $F^l = [(F_1^{l+1}, q_1; \ldots ; F_j^{l+1}, q_j; \ldots ; F_N^{l+1}, q_N)]$ *and* $\widehat{F}^l = [(F_1^{l+1}, q_1; \ldots ; F_{j-1}^{l+1}, q_{j-1}; \widehat{F}_j^{l+1}, q_j; F_{j+1}^{l+1}, q_{j+1}; \ldots ; F_N^{l+1}, q_N)]$ *in* $\mathcal{F}_l^L$ *with* $q_j > 0$*:* $F^l \succeq_2 \widehat{F}^l$ *if and only if* $[F_j^{l+1}, 1] \succeq_{\mathcal{F}_l^L} [\widehat{F}_j^{l+1}, 1]$.

(ii) extended single-action information loving (ESAIL) *if for all* l *in* $\{1, \ldots, L-1\}$*, all pairs of* L*-stage lotteries of the form* $F^l = [(F_1^{l+1}, q_1; \ldots ; F_j^{l+1}, q_j; \ldots ; F_N^{l+1}, q_N)]$ *and* $\widehat{F}^l = [(F_1^{l+1}, q_1; \ldots ; F_{j-1}^{l+1}, q_{j-1}; \widehat{F}_j^{l+1}, \beta q_j; \widetilde{F}_j^{l+1}, (1-\beta) q_j; F_{j+1}^{l+1}, q_{j+1}; \ldots ; F_N^{l+1}, q_N)]$ *in* $\mathcal{F}_l^L$ *with* $q_j > 0$*,* β *in* $[0,1]$*, and* $F_j^{l+1} = \beta \widehat{F}_j^{l+1} + (1-\beta)\widetilde{F}_j^{l+1}$*:* $\widehat{F}^l \succeq_L F^l$*. Similarly, if her preference relation,* $\succeq_L$*, satisfies* extended single-action information aversion (ESAIA) *then* $F^l \succeq_L \widehat{F}^l$*; and if it satisfies* extended single-action information neutrality (ESAIN) *then* $F^l \sim_L \widehat{F}^l$.

Most of the results for two stage lotteries can now be rewritten for extended lotteries. For example by Corollary 3.2.1, if $\succeq_L$ satisfies ECI and ESAIL then for all l in $\{2, \ldots, L\}$, the preference relation $\succeq_{\mathcal{F}_l^L}$ is quasi-convex in the probabilities. If, in addition, for all l in $\{1, \ldots, L-1\}$, the preference relation $\succeq_{L,l}$ is risk-averse then by Proposition 3.3, for all l in $\{2, \ldots, L\}$, the corresponding certainty-equivalent function is convex. The most dramatic extensions are the implications of preference for information for extended temporal RDEU and temporal betweenness preferences. If we require agents to exhibit preference for information, then neither of these models allows for much generalization of the original Kreps & Porteus's (1978) temporal expected utility model. Given RDEU, deviations from expected utility are only possible for the preferences over lotteries that resolve in the last stage. Given betweenness, deviations from expected utility are only possible for the preferences over lotteries that resolve in the first stage. These are immediate consequences of Propositions 3.6(i) and 3.7 respectively. Formally,

Proposition 3.8 *Suppose that an agent's preference relation over lotteries of finite length* L *satisfies* ECI *and* ESAIL*. Suppose further that, for all* l *in* $\{1, \ldots, L\}$ *and the preference relations* $\succeq_{L,l}$ *can be represented by Gateaux differentiable utility functions. Then:*

(i) if for all l *in* $\{1, \ldots, L\}$*, the preference relation* $\succeq_{L,l}$ *satisfies* RDEU*, and if the preference relation* $\succeq_{L,1}$*is (weakly) risk-averse, then* $\succeq_{L,l}$ *satisfies independence for all* l *in* $\{1, \ldots, L-1\}$*;*

(ii) if for all l *in* $\{1, \ldots, L\}$*, the preference relation* $\succeq_{L,l}$ *satisfies betweenness then* $\succeq_{L,l}$ *satisfies independence for all* l *in* $\{2, \ldots, L\}$.

Similar results hold SAIA, mutatis mutandis.

In contrast to Corollary 3.6.1, the extended result for RDEU in Proposition 3.8 part (i) does not require risk aversion for the intermediate stages 2, ..., $L-1$. The reason is that Proposition 3.6 part (i) requires preferences over early-resolution lotteries to be quasi-concave and preferences over late-resolution lotteries to be quasi-concave. But, all immediate stage lotteries are, in a sense, both early and late. So for all intermediate stages, the preference relation $\succeq_{L,l}$ must satisfy betweenness. And, betweenness combined with RDEU imply independence. For the Betweenness result, as before, we could replace Gateaux differentiability by risk aversion throughout.

Sarin & Wakker (1993) examine a restriction on preferences over multi-stage decision trees that they call "structural consistency". Roughly speaking, in the context of folding back and of RDEU (respectively, betweenness), structural consistency imposes that not only does the agent use some RDEU (respectively, betweenness) preferences at each stage of the (folded back) tree but also that she use some (possibly different) RDEU (respectively, betweenness) preferences in a direct evaluation of the associated (normal form) strategies. Sarin & Wakker find that their folding-back and structural consistency conditions impose similar restrictions as we find are imposed by SAIL, CI and risk aversion. That is, given RDEU (respectively, betweenness), folding back and structural consistency imply that deviations from expected utility are only possible in the last (respectively, first) stage. Structural consistency, however, is not about the agent's attitude towards information, and SAIL is not about structural consistency. It is an open question why these seemingly quite different restrictions have such similar consequences.

To show that SAIL does not generally imply independence, we can extend our earlier temporal quadratic utility example as follows.

Example 2 *Fix $L \geq 2$ and as in Example 1 take $X = [0,1]$. Let $\succeq_L$ denote the agent's preferences over lotteries of length L and assume they satisfy extended compound independence (ECI). Further for each l in $\{1, \ldots, L\}$, let $\succeq_{L,l}$ be represented by the following quadratic functional on one-stage lotteries: $W_l(F) = \frac{1}{2}\int v_l(x)\,F(dx) + \left(\int u_l(x)\,F(dx)\right)^2$, where $v_l(x) = x^{2^{-l}}$, and $u_l(x) = x^{2^{-(l+1)}}$.*

Notice that $W_l(\delta_x) = x^{2^{-l}}$ and so $\mathrm{CE}_l(F) = h_l \circ W_l(F)$, where $h_l(z) = z^{2^l}$. Hence, as was the case for Example 1, it is straightforward to show that for each l in $\{1, \ldots, L-1\}$, the compound function $u_l(\mathrm{CE}_{l+1}(\cdot), G)$ is convex for all G in $\mathcal{L}(\mathcal{X})$. ESAIL thus follows from the natural extension of Proposition 3.4.

4 Preference for information in many-action trees.

We now return to the agent's preferences over general Blackwell decision trees. That is, we allow for both intrinsic and instrumental attitudes towards information. We will assume that the agent satisfies freedom independence to focus on the case where either the agent is dynamically consistent or she does not take dynamic inconsistency into account in her preference over trees. Strictly speaking we do not need to assume resolute choice but our analysis is probably of most relevance where the agent is resolute. The case where the agent is not dynamically consistent is covered elsewhere in the literature (see especially Wakker (1988)).

A preference for information in all Blackwell decision trees implies a preference for information in Blackwell commitment trees: hence, SAIL is a necessary condition for general preference for information. We are interested in conditions under which SAIL is also sufficient for the agent always to prefer more informative signals regardless of the number of actions available to her.

Definition *An agent is* many-action information-loving (MAIL) *if: for all priors π in $\mathcal{L}([0,1])$, all consequence functions c in C, and all closed action sets $A \subset \mathcal{A}$, if the signal (S, λ) is more informative than the signal (S', λ') with respect to the prior π, then the Blackwell tree induced by (S, λ) is weakly preferred to that induced by (S', λ').*

Consider the standard case covered by Blackwell's Theorem in which the agent satisfies RCLA (SAIN) and expected utility. Loosely speaking, an intuition for why an agent always prefers more informative signals in this case is that information enlarges the set of reduced lotteries available to her. For this argument to be enough, however, we need to allow the agent to use mixed (behavioral) strategies. In the standard case, it does not matter if we restrict the agent to use only pure strategies since, with expected utility, the agent never strictly prefers to randomize. Once we drop independence, however, the agent might strictly prefer mixed strategies if they were permitted. We will therefore consider preference for information with many actions first admitting and then excluding the use of mixed strategies.

Once we drop SAIN, if we allow randomization, the agent may care whether the uncertainty of the mixed strategy itself is resolved early or late; that is, when the "coin is flipped". With *early-resolution mixed strategies*, the resolution of the randomizing device is contemporaneous with the resolution of the signal: the agent can be thought of as selecting a pure strategy at the second stage, with the choice depending jointly on the resolutions of the randomizing device and of the signal. With *late-resolution mixed strategies*, the resolution of the randomizing device is contemporaneous with the final realization of the state: that is, the agent mixes at the second stage. Suppose that an early-resolution mixed strategy specifies that the agent chooses action a_1 with conditional probability α and action a_2 with $(1-\alpha)$ at the signal realization s which occurs with probability q. Then, we can think of this as if the agent faces a related signal in which there are two realizations s_1 and s_2 instead of s, both of which induce the same posterior, and which occur with probabilities αq and $(1-\alpha)q$ respectively. An example of an early-resolution mixed strategy is the randomization chosen for reasons of procedural fairness by Machina's "mom" between giving an indivisible treat to her children, Abigail and Benjie.[16]

To see why the ability to randomize may be an issue consider the following example. Suppose that the prior assigns positive probability to only two states. Posteriors on these two states are represented by points on a one-dimensional simplex, measuring the posterior probability that state 1 obtains. Consider the following two signals. The first signal has three possible realizations, each of which occurs with probability 1/3. They lead to the posteriors 1/4, 1/2 and 3/4, respectively. The second signal has only two realizations, each of which occurs with probability 1/2. They lead to the posteriors 1/4 and 3/4, respectively. The second distribution is a linear bifurcation of the first, hence the second signal is more informative. Given the first signal, the agent can choose actions that induce a two-stage lottery with three different second-stage sublotteries over outcomes. For example, she could choose a different action at each signal realization. Given the second signal, however, the agent cannot choose three different actions without using a mixed strategy. Thus, she can only induce a two-stage lottery with two second-stage sublotteries. But, if the agent does not satisfy conditional quasi-convexity, the agent might sometimes prefer the two-stage lottery with a first-stage mixture of three second-stage sublotteries to the two-stage lottery with a first-stage mixture of only two second-stage sublotteries. For example, Machina's "mom" might have such preferences if she had three children but only one indivisible treat. If the agent cannot randomize, she may wish to use signal as if they were randomization devices. A signal can be less informative but nevertheless have more realizations and hence offer more scope to randomize. Therefore, it is possible for an agent who likes to randomize but, by assumption, is denied a randomization device, to prefer a less informative signal.

Given freedom independence, if we either allow the agent to use early-resolution mixed strategies or restrict her preferences such that she never prefers randomizations, however, then single-action preference for information implies many-action preference for information. If the agent can use early-resolution mixed strategies, her choice of two-stage lotteries cannot

[16] The point of Machina's example is to illustrate resolute choice. If Machina's mom had used a late resolution mixed strategy, then no issue of dynamic consistency would arise. See Machina (1989).

be reduced by additional information. Even if the number of signal realizations is reduced, she can design early-resolution mixed strategies to act as if additional realizations were available. Alternatively, if the agent satisfies conditional quasi-convexity then she will never strictly prefer to randomize.

Proposition 4.1 *Given freedom independence,*

(i) if the agent can employ early-resolution mixed strategies with finite support then her preference relation over two-stage lotteries satisfies SAIL *if and only if she is* MAIL*;*

(ii) regardless of whether the agent can or cannot employ mixed strategies, if the agent's preference relation over two-stage lotteries satisfies CQV *then it satisfies* SAIL *if and only if she is* MAIL.

We did not assume in proposition 4.1 that the agent satisfy the fold-back property associated with compound independence. However, since compound independence implies conditional quasi-convexity, we obtain the following corollary.

Corollary 4.1.1 *Given freedom independence, regardless of whether the agent can or cannot employ mixed strategies, if the agent's preference relation over two-stage lotteries satisfies* CI *then it satisfies* SAIL *if and only if she is* MAIL.

The assumption of compound independence here does not imply full independence. This contrasts sharply with Wakker's (1988) result. The reason is that we do not have the reduction axiom; that is, we do not assume single-action information neutrality.

Even if we do not allow the agent to use mixed strategies and even if her preferences over two-stage lotteries do not satisfy CQV, single-action information loving still implies that she will almost always prefer more information in many-action problems. Loosely speaking, whenever a more informative signal offers fewer opportunities to randomize, we can always slightly perturb the signal to increase the number of possible realizations. Recall from Lemma 2.1 that we can always obtain the distribution of posteriors generated by a more informative signal from that of a less informative signal by a sequence of linear bifurcations. The agent will always prefer the more informative signal provided we restrict the sequence such that the two posteriors produced by each bifurcation do not lie in the support of the previous distribution of posteriors.

Proposition 4.2 *Suppose that the agent satisfies freedom independence and SAIL. Let* $[(P_i, q_i)_{i=1}^{N}]$ *and* $[(P'_j, q'_j)_{j=1}^{N'}]$ *be two distributions of posteriors with* $\sum_i q_i P_i = \sum_j q'_j P'_j = \pi$, *for which there exists of sequence of posterior lotteries* $([(P_i^k, q_i^k)_{i=1}^{N^k}])_{k=1}^{K}$, *with* $[(P_i^1, q_i^1)_{i=1}^{N^1}] =$

$[(P'_j, q'_j)_{j=1}^{N'}]$ *and* $[(P_i^K, q_i^K)_{i=1}^{N^K}] = [(P_i, q_i)_{i=1}^{N}]$, *such that* $[(P_i^{k+1}, q_i^{k+1})_{i=1}^{N^{k+1}}]$ *is a linear bifurcation of* $[(P_i^k, q_i^k)_{i=1}^{N^k}]$ *and such that there exists two posteriors* P_j^{k+1}, $P_{j'}^{k+1}$ *in* $(P_i^{k+1})_{i=1}^{N^{k+1}}$ *that are not in* $(P_i^k)_{i=1}^{N^k}$ *for* $k = 1, \ldots, K-1$. *Then, for all consequence functions* c *in* $\mathcal{C}$, *and all closed action sets* $A \subset \mathcal{A}$, *and all lists of signal realizations* $(s_i)_{i=1}^{N}$ *and* $(s'_i)_{i=1}^{N'}$*: the Blackwell decision tree* $\left\langle (s_i, q_i, P_i, A)_{i=1}^{N}, c \right\rangle$ *is weakly preferred to the Blackwell decision tree* $\left\langle (s'_i, q'_i, P'_i, A)_{i=1}^{N'}, c \right\rangle$.

5 Preference for Information and Dynamic Inconsistency

Although we have focussed throughout on dynamically consistent agents, we conclude by observing that intrinsic preference for information may itself sometimes lead agents to be dynamically inconsistent. Consider the dilemma of a prisoner, let's call her Jo, who is anxiously awaiting the outcome of her trial for murder. If she is found guilty she will spend the next ten years in jail. The verdict has been delivered, but Judge Oti has sealed it in an envelope overnight. Jo would much rather that the envelope be opened immediately; she has intrinsic preference for information. Meanwhile, the DA is considering offering Jo a plea bargain involving only four years in jail. Suppose that if the DA were to offer Jo a choice between opening the envelope now or taking the plea, Jo would be more or less indifferent. Perhaps she would choose to rip open the envelope. Suppose instead that the DA offered either the plea or simply waiting for the envelope to be opened in the morning. By transitivity, Jo would want to take the plea. But, now suppose that the plea must be declared in front of Judge Oti in the morning, and Jo cannot commit ahead of time. Come the morning, the choice between the plea or the envelope will look much the same as the choice looked the night before between the plea and immediately opening the envelope, not waiting overnight. Therefore, Jo may refuse the plea come morning. But, anticipating the inconsistency of her future self, Jo is destined for an uneasy night.

Although any resemblance between these characters and any real people is purely coincidental, the story is perhaps suggestive. In our analysis, as in Kreps & Porteus (1978), concern for early or late resolution can be thought of as concern about the time that separates the resolution of uncertainty from the impact of the event itself. In this framework, at least resolute choice if not the fold-back property, seems a natural assumption. But introspection and our prisoner story suggest that concern for early resolution should instead be thought of as concern about the time that separates the resolution of uncertainty from the decision maker herself; that is, from the present. In this case, the fold-back property and even resolute choice seem less natural. If uncertainties cause us more (or less) anxiety the further we are away from their resolution, then, over time, as their resolution gets closer, the choices we make to relieve that anxiety may change. The fold-back property excludes such inclinations, and resolute choice excludes such behavior. We could, for example, adapt the Kreps-Porteus model to allow an agent's degree of risk aversion to depend on the time interval that separates the agent from resolution rather than that between resolution and the impact of the event. Such a model may offer a better description of how intrinsic concern for information affects behavior. Not unlike hyperbolic discounting, however, such an approach would commit us to dynamically inconsistent agents. It remains an avenue for further research.

Appendix

Proof of Lemma 2.1(i) $\Rightarrow$ (ii): Without loss of generality, we can identify S with $(s_i)_{i=1}^N$ and S' with $(s'_j)_{j=1}^{N'}$ where any s in S or s' in S' that has zero probability under the prior π is omitted. Then, since for π-almost all ω, $\lambda'(s'_j|\omega) = \sum_i \alpha(s'_j, s_i)\lambda(s_i|\omega)$ for each j, Bayes Rule implies that $P'_j(\omega)q'_j = \sum_i P_i(\omega)\alpha(s'_j, s_i)q_i$ for π-almost all ω. Therefore, for π-almost all ω, $\sum_i \beta_{ij}P_i = P'_j$ for all j, where $\beta_{ij} = (\alpha(s'_j, s_i)q_i/q'_j)$. Next, observe that $q'_j = \int_\omega \lambda'(s'_j|\omega)\pi(d\omega)$ $= \sum_i \alpha(s'_j, s_i)\int_\omega \lambda(s_i|\omega)\pi(d\omega) = \sum_i \alpha(s'_j, s_i)q_i$, so that $\sum_i \beta_{ij} = (1/q'_j)\sum_i \alpha(s'_j, s_i)q_i = 1$. And, since $\sum_j \alpha(s'_j, s_i) = 1$, we have $\sum_j \beta_{ij}q'_j = q_i$. □

(ii)$\Rightarrow$(i): Let S_0 be those signal realizations that occur with zero probability under the prior π; that is, $\{s \in S | \int_\omega \lambda(s|\omega)\pi(d\omega) = 0\}$. Similarly define S'_0. We can then identify $S\backslash S_0$ with $(s_i)_{i=1}^N$, and $S'\backslash S'_0$ with $(s'_j)_{j=1}^{N'}$. Since for π-almost all ω, $\sum_i \beta_{ij}P_i = P'_j$ for all j, Bayes Rule implies that $\lambda'(s'_j|\omega)\pi(\omega) = q'_j\sum_i \lambda(s_i|\omega)\pi(\omega)\beta_{ij}/q_i$ for π-almost all ω. Therefore, for π-almost all ω, $\sum_i \alpha(s'_j, s_i)\lambda(s_i|\omega) = \lambda'(s'_j|\omega)$ where $\alpha(s'_j, s_i) = (\beta_{ij}q'_j/q_i)$ for all i and all j. Set $\alpha(s'_0, s) = 0$, for all s'_0 in S'_0 and all s in S. Set $\alpha(s', s_0)$ such that $\sum_{s'\in S'\backslash S'_0}\alpha(s', s_0) = 1$, for all s_0 in S_0. Then, since $\sum_j \beta_{ij}q'_j = q_i$ for all i, $\sum_j \alpha(s'_j, s) = (1/q_i)\sum_j \beta_{ij}q'_j = 1$. □

(ii)$\Rightarrow$(iii): We will sketch how to construct such a sequence. Let $[(P_i^2, q_i^2)_{i=1}^{N^2}] = [P_0^2, (1-\beta_{11})q'_1; P_1, \beta_{11}q'_1; P'_2, q'_2; \ldots; P'_{N'}, q'_{N'}]$ where $(1-\beta_{11})P_0^2 + \beta_{11}P_1 = P'_1$. And, let $[(P_i^3, q_i^3)_{i=1}^{N^3}] = [P_0^3, (1-\beta_{11}-\beta_{21})q'_1; P_1, \beta_{11}q'_1; P_2, \beta_{21}q'_1; P'_2, q'_2; \ldots; P'_{N'}, q'_{N'}]$ where $(1-\beta_{11}-\beta_{21})P_0^3 + \beta_{21}P_2 = (1-\beta_{11})P_0^2$. Similarly, construct $[(P_i^k, q_i^k)_{i=1}^{N^k}]$, for $k = 1, \ldots, N$. By construction, $\sum_{i=1}^{N-1}\beta_{i1}P_i + (1-\sum_{i=1}^{N-1}\beta_{i1})P_0^N = P'_1 = \sum_i \beta_{ij}P_i$. Hence, $P_0^N = P^N$, as desired. Construct $[(P_i^k, q_i^k)_{i=1}^{N^k}]$, for $k = jN+1, \ldots, (j+1)N$, $j = 2, \ldots, N'$ by "splitting" P'_j in a similar manner. □

The direction (iii)$\Rightarrow$(ii) follows from the definition of a linear bifurcation. ■

Proof of Proposition 3.1. The definition of a bifurcation can be extended to two-stage lotteries over outcomes in a natural way. By continuity, any pair of two-stage lotteries are ranked by linear bifurcation if and only if they can be approximated by a pair of two-stage lotteries which are discrete in the second stage and that are themselves ranked by linear bifurcation.

Lemma A.1 *Let* $X = [(F_1, q_1;\ \ldots;\ F_j, q_j;\ \ldots; F_N, q_N)]$ *and* $Y = [(F_1, q_1;\ \ldots;\ F_{j-1}, q_{j-1};\ F'_j, \beta q_j;\ F''_j, (1-\beta)q_j;\ F_{j+1}, q_{j+1};\ \ldots; F_N, q_N)]$ *in* $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$ *where* $\beta \in (0, 1]$. *Then* Y *is a linear bifurcation of* X *(that is,* $F_j = \beta F'_j + (1-\beta)F''_j$*) if and only if there exists a sequence of pairs of two-stage lotteries* X^n *and* Y^n *in* $\mathcal{L}_0(\mathcal{L}_0(\mathcal{X}))$ *such that* Y^n *is a linear bifurcation on* X^n *for all* n*, and such that* X^n *converges to* X *and* Y^n *converges to* Y.

Proof. Suppose $F_j = \beta F'_j + (1-\beta)F''_j$. For each $n = 1, 2, \ldots$. Fix the Prohorov metric μ on $\mathcal{L}(\mathcal{X})$. Since $\mathcal{L}_0(\mathcal{X})$ is dense in $\mathcal{L}(\mathcal{X})$, for each n, choose $(F_i)^n$ for $i \neq j$, and $(F'_j)^n$ and $(F''_j)^n \in \mathcal{L}_0(\mathcal{X})$ with the property: $\mu((F_i)^n, F_i) < \frac{1}{n}$, $\mu((F'_j)^n, F'_j) < \frac{1}{n}$, and $\mu((F''_j)^n, F''_j) < \frac{1}{n}$. Set $(F_j)^n = \beta(F'_j)^n + (1-\beta)(F''_j)^n$, and let $X^n = [((F_1)^n, q_1; \ldots; (F_N)^n, q_N)]$ and $Y^n = [((F_1)^n, q_1;\ \ldots;\ (F_{j-1})^n, q_{j-1};\ (F'_j)^n, \beta q_j;\ (F''_j)^n, (1-\beta)q_j;\ (F_{j+1})^n, q_{j+1};\ \ldots;\ (F_N)^n, q_N)]$. By construction, we have $X^n, Y^n \in \mathcal{L}_0(\mathcal{L}_0(\mathcal{X}))$ and Y^n is a linear bifurcation of X^n, and by the linearity of the topology, $\mu((F_i)^n, F_i) = \mu(\beta(F'_j)^n + (1-\beta)(F''_j)^n,\ \beta F'_j + (1-\beta)F''_j) < \frac{K}{n}$ for some constant K. Thus as $n \to \infty$, $X^n \to X$ and $Y^n \to Y$ in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$. The converse follows from the continuity of linear operation. □

We now prove the proposition. (i)⇒(ii) Suppose SAIL does not hold. Then, by Lemma A.1 and the continuity of the preference relation $\succeq_2$, there exists two lotteries that are discrete in the second stage, $X = [(F_1, q_1;\ \ldots;\ F_j, q_j;\ \ldots; F_N, q_N)]$ and $Y = [(F_1, q_1;\ \ldots;\ F_{j-1}, q_{j-1};\ F'_j, \beta q_j;\ F''_j, (1-\beta)q_j;\ F_{j+1}, q_{j+1};\ \ldots; F_N, q_N)]$ in $\mathcal{L}_0(\mathcal{L}_0(\mathcal{X}))$, where $q_j > 0$ and $F_j = \beta F'_j + (1-\beta)F''_j$, such that $Y \succ_2 X$. Choose c in $\mathcal{C}$ and a in $\mathcal{A}$ such that $c(\cdot, a)$ is a bijection from $[0, 1]$ to $\mathcal{X}$ and choose $(P_i)_{i=1}^N$, P'_j and P''_j, such that $(\tilde{c}(P_i, a))_{i=1}^N = (F_i)_{i=1}^N$, $\tilde{c}(P'_j, a) = F'_j$ and $\tilde{c}(P''_j, a) = F''_j$. By construction $\sum_i q_i P_i = \sum_{i \neq j} q_i P_i + \beta q_j P'_j + (1-\beta)q_j P''_j =: \pi$. Then by Lemma 2.1, there exists two signals (S, λ) and (S', λ') such that (S, λ) is more informative than (S', λ') with respect π, and such that the Blackwell commitment tree induced by (S', λ') given prior π, single action a and consequence function c, is strictly preferred to that induced by (S, λ). The proof for SAIN and for SAIA are similar. □

(ii) ⇒ (i): Commitment trees can be identified with the two-stage lotteries over consequences that they induce. By Lemma 2.1, the two-stage lottery induced by the more informative signal can be derived from that induced by the less informative signal by a sequence of linear bifurcations. The claim follows from transitivity of the preference relation $\succeq_2$. The proof for SAIN and for SAIA are similar. ■

Proof of Corollary 3.1.1. By construction, if X is obtained from Y by a simple linear bifurcation then $\rho(X) = \rho(Y)$, thus (i) implies (ii). Conversely, if $\rho(X) = \rho(Y)$, take Z to be the late-resolution lottery with $\rho(Z) = \rho(X) = \rho(Y)$. Then, both X and Y can be obtained from Z by a sequence of simple linear bifurcations. Thus SAIN implies $X \sim_2 Z \sim_2 Y$. It is immediate that (ii) implies (iii). For (iii) implies (ii), if SAIL applies but SAIN does not, then (by Lemma A.1 and the continuity of $\succeq_2$) there exists a pair of two-stage lotteries discrete in the second stage, X and Y in $\mathcal{L}_0(\mathcal{L}_0(\mathcal{X}))$ such that X is a linear bifurcation of Y, and $X \succ_2 Y$. Let $(x_i, r_i)_{i=1}^N := \rho(X) = \rho(Y)$. Repeated application of SAIL and the transitivity of preference imply that $[(\delta_{x_i}, r_i)_{i=1}^N] \succeq_2 X \succ_2 Y \succeq_2 [(x_i, r_i)_{i=1}^N, 1]$. But by TN, $[(x_i, r_i)_{i=1}^N, 1] \sim_2 [(\delta_{x_i}, r_i)_{i=1}^N]$. The proof for SAIA is similar. ■

Proof of Proposition 3.3. The first claim is straightforward given that both early- and late-resolution preferences respect first-order stochastic dominance. For the second claim, fix any $F = (x_i, r_i)_{i=1}^N$, $F' = (x'_i, r'_i)_{i=1}^{N'}$ and α. By the definition of $\mathrm{CCE}_{er}(F; F', \alpha)$, $[\delta_{\mathrm{CCE}_{er}(F;F',\alpha)}, \alpha; \delta_{x'_1}, (1-\alpha)r'_1; \ldots; \delta_{x'_{N'}}, (1-\alpha)r'_{N'}] \sim_{er} [\delta_{x_1}, \alpha r_1; \ldots; \delta_{x_N}, \alpha r_N; \delta_{x'_1}, (1-\alpha)r'_1; \ldots; \delta_{x'_{N'}}, (1-\alpha)r'_{N'}]$. By SAIL, $[\delta_{x_1}, \alpha r_1; \ldots; \delta_{x_N}, \alpha r_N; \delta_{x'_1}, (1-\alpha)r'_1; \ldots; \delta_{x'_{N'}}, (1-\alpha)r'_{N'}] \succeq_2 [F, \alpha; \delta_{x'_1}, (1-\alpha)r'_1; \ldots; \delta_{x'_{N'}}, (1-\alpha)r'_{N'}]$. And by compound independence and the definition of $\mathrm{CE}_{lr}(F)$, $[F, \alpha; \delta_{x'_1}, (1-\alpha)r'_1; \ldots; \delta_{x'_{N'}}, (1-\alpha)r'_{N'}] \sim_2 [\delta_{\mathrm{CE}_{lr}(F)}, \alpha; \delta_{x'_1}, (1-\alpha)r'_1; \ldots; \delta_{x'_{N'}}, (1-\alpha)r'_{N'}]$. Since the early-resolution preferences respect first-order stochastic dominance, the result follows. For the third claim, fix any F, F' in $\mathcal{L}(\mathcal{X})$, and α. By CI and SAIL it follows that $[\delta_{\mathrm{CE}_{lr}(F)}, \alpha; \delta_{\mathrm{CE}_{lr}(F')}, 1-\alpha] \succeq_2 [\delta_{\mathrm{CE}_{lr}(\alpha F + (1-\alpha)F'])}, 1]$. And by risk aversion of $\succeq_{er}$ it follows that $[\delta_{[\alpha\, \mathrm{CE}_{lr}(F) + (1-\alpha)\, \mathrm{CE}_{lr}(F')]}, 1] \succeq_{er} [\delta_{\mathrm{CE}_{lr}(F)}, \alpha; \delta_{\mathrm{CE}_{lr}(F')}, 1-\alpha]$. Since $\succeq_{er}$ respects first-order stochastic dominance we have $\alpha\, \mathrm{CE}_{lr}(F) + (1-\alpha)\, \mathrm{CE}_{lr}(F') \geq \mathrm{CE}_{lr}(\alpha F + (1-\alpha)F'])$, as required. The proof for SAIA is similar. ■

Proof of Proposition 3.4. Suppose there exists a lottery G_X in $\mathcal{L}_0(\mathcal{X})$ such that $u_{er}(\mathrm{CE}_{lr}(\cdot), G_X)$ is not convex. Then there are a pair of second-stage lotteries F' and F'' and an α in $(0, 1)$ such that $\alpha[u_{er}(\mathrm{CE}_{lr}(F'), G_X)] + (1-\alpha)u_{er}(\mathrm{CE}_{lr}(F''), G_X) - u_{er}(\mathrm{CE}_{lr}(\alpha F' + (1-\alpha)F''), G_X) < 0$. Let $X = [(F_i, q_i)_{i=1}^n]$ be the two-stage lottery such that $G_X = [(\delta_{\mathrm{CE}_{lr}(F_i)}, q_i)_{i=1}^n]$. Write $\overline{F} = \alpha F' + (1-\alpha)F''$, and set $x_i = CE_{lr}(F_i)$, $y_0 = CE_{lr}(\overline{F})$, $y_1 = CE_{lr}(F')$ and $y_2 = CE_{lr}(F'')$. Thus, $\alpha[u_{er}(y_1, G_X)] + (1-\alpha)u_{er}(y_2, G_X) - u_{er}(y_0, G_X) < 0$.

Let $G_0 = [(y_0, 1)]$ and $G_1 = [(y_1, \alpha); (y_2, 1-\alpha)]$. Let $Y_\epsilon := [\overline{F}, \epsilon; F_1, (1-\epsilon)q_1; \ldots; F_n, (1-\epsilon)q_n]$ and let $G_{Y_\epsilon} := [(y_0, \epsilon); (x_i, (1-\epsilon)q_i)_{i=1}^n] = \epsilon G_0 + (1-\epsilon)G_X$ denote the induced one stage lottery. Now, consider bifurcating Y_ϵ to create $Y'_\epsilon := [F', \alpha\epsilon; F'', (1-\alpha)\epsilon; F_1, (1-\epsilon)q_1; \ldots; F_n, (1-\epsilon)q_n]$ and let $G_{Y_\epsilon} = [(y_1, \alpha\epsilon); (y_2, (1-\alpha)\epsilon); (x_i, (1-\epsilon)q_i)_{i=1}^n] = \epsilon G_1 + (1-\epsilon)G_X$ denote the induced lottery. By the Gateaux differentiability of V_{er} at G, for any $\epsilon >$

0, $V_{er}(G_{Y_\epsilon}) - V_{er}(G_X) = V_{er}(\epsilon G_0 + (1-\epsilon)G) - V_{er}(G) = \int u_{er}(x;G)[\epsilon G_0(dx) - \epsilon G(dx)] + o_1(\epsilon)$ $= \epsilon u_{er}(y_0, G) - \epsilon\sum_{i=1}^{n} u_{er}(x_i, G) + o_1(\epsilon)$ where $\frac{o_1(\epsilon)}{\epsilon} \to 0$ as $\epsilon \to +0$. Similarly, $V_{er}(G_{Y'_\epsilon})$ $- V_{er}(G) = \alpha\epsilon u_{er}(y_1, G) + (1-\alpha)\epsilon u_{er}(y_2, G) - \epsilon\sum_{i=1}^{n} u_{er}(x_i, G) + o_2(\epsilon)$, where $\frac{o_2(\epsilon)}{\epsilon} \to 0$ as $\epsilon \to +0$. Combining these expressions, we have $V_{er}(G_{Y'_\epsilon}) - V_{er}(G_{Y_\epsilon}) = \alpha\epsilon u_{er}(y_1, G) + (1-\alpha)\epsilon u_{er}(y_2, G) + o_2(\epsilon) - \epsilon u_{er}(y_0, G) - o_1(\epsilon)$ for any $\epsilon > 0$. Hence, $\frac{1}{\epsilon}[V_{er}(G_{Y'_\epsilon}) - V_{er}(G_{Y_\epsilon})]$ < 0 if ϵ is small enough. But, by SAIL, $Y'_\epsilon \succeq_2 Y_\epsilon$ for any $\epsilon > 0$. Thus, by CI, we get $V_{er}(G_{Y'_\epsilon})$ $- V_{er}(G_{Y_\epsilon}) \geq 0$ for any $\epsilon > 0$: a contradiction.

Conversely, suppose $u_{er}(\mathrm{CE}_{lr}(\cdot), G)$ is convex for all G in $\mathcal{L}(\mathcal{X})$. Pick any pair of two-stage lotteries $Y_0 := [\bar{F}, \epsilon;\ F_1, (1-\epsilon)q_1;\ \ldots;\ F_n, (1-\epsilon)q_n]$ and $Y_1 := [F', \alpha\epsilon;\ F'', (1-\alpha)\epsilon;\ F_1, (1-\epsilon)q_1;\ \ldots;\ F_n, (1-\epsilon)q_n]$ such that $\epsilon > 0$ and $\bar{F} = \alpha F' + (1-\alpha)F''$. It suffices to show $Y_1 \succeq_2 Y_0$.

Write $x_i = CE_{lr}(F_i)$, $y_0 = CE_{lr}(\overline{F})$, $y' = CE_{lr}(F')$ and $y'' = CE_{lr}(F'')$. Let $G_X := [(x_i, q_i)_{i=1}^{n}]$, $G_0 = [(y_0, 1)]$ and $G_1 = [(y', \alpha);\ (y'', 1-\alpha)]$. Denote by G_{Y_0} and G_{Y_1} the induced one stage lotteries of Y_0 and Y_1; that is, $G_{Y_0} = [(y_0, \epsilon);\ (x_i, (1-\epsilon)q_i)_{i=1}^{n}] = \epsilon G_0 + (1-\epsilon)G_X$ and $G_{Y_1} = [(y', \alpha\epsilon);\ (y'', (1-\alpha)\epsilon);\ (x_i, (1-\epsilon)q_i)_{i=1}^{n}] = \epsilon G_1 + (1-\epsilon)G_X$. By CI, $Y_1 \succeq_2 Y_0$ holds if and only if $V_{er}(G_{Y_1}) - V_{er}(G_{Y_0}) \geq 0$, which we shall establish below.

For ξ in $[0,1]$, let $G_\xi = \xi G_{Y_1} + (1-\xi)G_{Y_0}$. Then since V_{er} is Gateaux differentiable, $\frac{d}{d\xi}V(G_\xi)$ is well defined and $V_{er}(G_{Y_1}) - V_{er}(G_{Y_0}) = V_{er}(G_\xi)|_{\xi=0}^{1} = \int_0^1 [\frac{d}{d\xi}V(G_\xi)|_{\xi=\beta}]d\beta$. From the definition of Gateaux differentiability, for any β in $[0,1]$, $\frac{d}{d\xi}V_{er}(G_\xi)|_{\xi=\beta} =$ $\int u_{er}(x; G_\beta)[G_{Y_1}(dx) - G_{Y_0}(dx)] = \epsilon\int u_{er}(x; G_\beta)[G_1(dx) - G_0(dx)] = \epsilon[\alpha u_{er}(y'; G_\beta) + (1-\alpha)u_{er}(y''; G_\beta) - u_{er}(y_0; G_\beta)] = \epsilon[\alpha u_{er}(CE_{lr}(F'); G_\beta) + (1-\alpha)u_{er}(CE_{lr}(F''); G_\beta) - u_{er}(CE_{lr}(F_0); G_\beta)]$. By the convexity assumption, the last expression is non-negative. Thus, $V_{er}(G_{Y_1}) - V_{er}(G_{Y_0}) = \int_0^1 [\frac{d}{d\xi}V(G_\xi)|_{\xi=\beta}]d\beta \geq 0$, as desired. ■

Proof of Proposition 3.6. Without loss of generality, let $u_{er}(\underline{x}) = u_{lr}(\underline{x}) = 0$ and $u_{er}(\overline{x})$ $= u_{lr}(\overline{x}) = 1$. From Chew & Nishimura (1992), we know that the local utility function of the RDEU functional $\int_0^1 u(x)g(G(dx))$ has the form $U(x; G, u, g) = -\int_0^x g'(G(z))u(dz)$. Moreover, U is convex in x if and only if u is convex and g is concave. By Proposition 3.4, SAIL is equivalent to the convexity of the function $U(CE_{lr}(F); G, u_{er}, g_{er})$ in F for all G.

(i): SAIL implies that $\succeq_{lr}$ must be quasi-convex in the probabilities. But Wakker's (1994) Theorem 24 shows that if a preference relation over one-stage lotteries satisfies RDEU then it is quasi-convex in probabilities if and only if its decumulative distortion function, g, is convex. Hence g_{lr} must be convex.

Next, fix any G and for any $r \in [0,1]$, consider a one-stage lottery $F_r = [\overline{x}, r;\ \underline{x}, 1-r]$. Since $V_{lr}(F_r) = g_{lr}(r)$, we obtain $U(CE_{lr}(F_r);\ G, u_{er}, g_{er}) = -\int_0^{u_{lr}^{-1}(g_{lr}(r))} g'_{er}(G(z))u_{er}(dz)$ $= -\int_0^r g'_{er}(G(u_{lr}^{-1}(g_{lr}(t)))u_{er}(u_{lr}^{-1}(g_{lr}(dt)))$, where the last equality is by a change of variable.

Notice that $G(u_{lr}^{-1}(g_{lr}(\cdot))$ by itself is the decumulative distribution function of some lottery, so the last expression is exactly $U(r;\ G(u_{lr}^{-1}(g_{lr}(\cdot)),\ u_{er}(u_{lr}^{-1}(g_{lr}(\cdot))), g_{er})$. SAIL implies that this function is convex in r, thus $u_{er}(u_{lr}^{-1}(g_{lr}(\cdot)))$ must be convex and g_{er} must be concave.

(ii): Fix any G, and set $\tilde{G} = G \circ u_{lr}^{-1}$. By a change of variable, $U(CE_{lr}(F);\ G, u_{er}, g_{er})$ $= -\int_0^{V_{lr}(F)} g'_{er}(G(u_{lr}^{-1}(t)))u_{er}(u_{lr}^{-1}(dt)) = U(V_{lr}(F);\ \tilde{G}, u_{er}(u_{lr}^{-1}(\cdot), g_{er})$. Since $u_{er}(u_{lr}^{-1}(\cdot))$ is convex and g_{er} is concave, $U(\cdot;\ \tilde{G}, u_{er}(u_{lr}^{-1}(\cdot), g_{er})$ is a convex function. But, by Wakker's (1994) Observation 2, since g_{lr} is convex, V_{lr} is convex in the probabilities; that is, $\alpha V_{lr}(F_1)$ $+\ (1-\alpha)V_{lr}(F_2) \geq V_{lr}(\alpha F_1 + (1-\alpha)F_2)$. Thus, $U(CE_{lr}(F);\ G, u_{er}, g_{er})$, as a composite of convex functions, is convex in F.

(iii) Fix any G and, this time, set $\tilde{G} = G(u_{lr}^{-1}(g_{lr}(\cdot)))$. By a change of variable, $U(CE_{lr}(F);$ $G,\ u_{er},\ g_{er}) = -\int_0^{u_{lr}^{-1}(g_{lr}(g_{lr}^{-1}(V_{lr}(F))))} g'_{er}(G(u_{lr}^{-1}(t)))u_{er}(u_{lr}^{-1}(dt)) =$ $-\int_0^{(g_{lr}^{-1}(V_{lr}(F))} g'_{er}(G(u_{lr}^{-1}(g_{lr}(t))))u_{er}(u_{lr}^{-1}(g_{lr}(dt))) = U(g_{lr}^{-1}(V_{lr}(F)); \tilde{G},\ u_{er}(u_{lr}^{-1}(g_{lr}(\cdot))), g_{er})$. Since $u_{er}(u_{lr}^{-1}(g_{lr}(\cdot)))$ is convex and g_{er} is concave, $U(\cdot;\ \tilde{G},\ u_{er}(u_{lr}^{-1}(g_{lr}(\cdot))),\ g_{er})$ is a convex function. So it suffices to show $g_{lr}^{-1}(V_{lr}(F))$ is convex in F. In fact, by Lemma A.1 and the continuity of $\succeq_2$, it is enough to show that SAIL holds for all two-stage lotteries that are discrete in the second stage. So it suffices to show that $g_{lr}^{-1}(V_{lr}(F))$ is convex in those F with finite support.

Assume that the supports of F_1 and F_2 are the same list of outcomes $\{x_1, \ldots, x_M\}$ where $x_1 = \overline{x}$, $x_M = \underline{x}$ and $x_j \leq x_{j-1}$ for all $j = 2, \ldots, M$. This is without loss of generality since we can always assign probability weights $p_j = 0$. So in particular, $g_{lr}^{-1} \circ V_{lr}(F_i)$ can be regarded as a function of $(p_1, ..., p_M)$ on the $M-1$ dimensional simplex. Algebraic manipulation yields that, for each G_i, $V_{lr}(G_i) = \sum_{j=1}^{M}(u_{lr}(x_j)[g_{lr}(\sum_{k=0}^{j} p_j) - g_{lr}(\sum_{k=0}^{j-1} p_j)])$ $= \sum_{j=1}^{M-1}(g_{lr}(\sum_{k=0}^{j} p_j)[u_{lr}(x_j) - u_{lr}(x_{j+1})]$ where we set $p_0 := 0$. Hence $g_{lr}^{-1} \circ V_{lr}(G_i) =$ $g_{lr}^{-1}(\sum_{j=1}^{M-1}(g_{lr}(\sum_{k=0}^{j} p_j)\ [u_{lr}(x_j) - u_{lr}(x_{j+1})])$ can be thought of as the generalized mean of $(p_1, ..., p_M)$ under the convex function $g_{lr} : [0,1] \rightarrow [0,1]$ where the "weight" on "outcome" p_j is given by $[u_{lr}(x_j) - u_{lr}(x_{j+1})]$. But, by Hardy, Littlewood and Polya's (1934) Theorem 106(i),[17] if g_{lr}'''' is continuous, g_{lr} is convex and g'_{lr}/g''_{lr} is concave then this generalized mean is convex. ∎

Proof of Proposition 3.7 Suppose that the agent satisfies SAIL. Then by the corollary to Proposition 3.4, $\succeq_{lr}$ must be convexifiable. But applying Kannai's (1976) general result, we have:

Lemma A.2 *Suppose that V represents preferences over $\mathcal{L}(X)$ that satisfy betweenness; that is, let V be a function from $\mathcal{L}(X)$ to the reals, such that for all lotteries F,*

[17] We thank Boaz Moselle for this reference.

$F' \in \mathcal{L}(X)$ and $\alpha \in (0,1)$, if $V(F) = V(F')$ then $V(\alpha F + (1-\alpha) F') = V(F)$. Then, if there exist a function h from $\mathbb{R}$ to $\mathbb{R}$ such that $h \circ V$ is convex, preferences satisfy independence.

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 CE_{lr}, $\succeq_{lr}$ must satisfy independence.

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 A.2. To be consistent with Kannai's notation, we shall show that if there is an h such that $h \circ V$ is *concave*, then independence must be satisfied.

Suppose preferences do not satisfy Independence, thus there exist F_1, F_2, $G \in \mathcal{L}(X)$ such that $V(F_1) = V(F_2)$ but $V(tF_1 + (1-t)G) > V(tF_2 + (1-t)G)$ for some $t \in (0,1)$. Let $\triangle$ be the convex hull of $\{F_1, F_2, G\}$, which is isomorphic to the set $\{(p_1,p_2) \in R^2_+ : p_1 + p_2 \leq 1\}$. The restriction of V on $\triangle$ satisfies Betweenness, so by construction, the indifference curves restricted on $\triangle$ are (isomorphic to) straight lines, but they are not parallel. Abusing notation, we shall identify $\triangle$ with the 2-dimensional set above, and we shall write $V(p)$ for $V(p_1F_1 + p_2F_2 + (1-p_1-p_2)G)$ for each $p = (p_1,p_2) \in \triangle$. The function $h \circ V$ is concave on $\triangle$. So we are done if we can show that a continuous function on R^2 whose level curves are linear but not parallel on a convex set with non-empty interior cannot be concavified.

The rest of argument is an application of Kannai's theorem 3.4.; we shall show that Kannai's necessary condition for concavifiability is violated. Without loss of generality, assume that at each $p \in \triangle$, there is $v(p) = (v_1(p), v_2(p)) \in \mathbb{R}^2$ with $|v(p)| = 1$ such that $V(p) = V(p')$ if and only if $v(p) \cdot (p - p') = 0$; that is, $v(p)$ is the normalized "gradient" of V at p. Since Independence is violated on $\triangle$, v is not a constant.

Pick $p, q \in \triangle$ with $V(p) > V(q)$ and let $\lambda(p,q) = \sup\{\frac{v(p)\cdot(p'-q')}{v(q)\cdot(p'-q')} : V(p) = V(p'),\ V(q) = V(q')\}$. 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, …, p_n with $V(p_0) > V(p_1) > \cdots > V(p_n)$. So, it is enough to show that we can set $\lambda(p,q)$ bounded away from 1 as p and q approach each other.

Pick p and q in $\triangle$ such that $v(p) \neq v(q)$, hence indifference curves (lines) intersects outside $\triangle$. Call the intersection $r(p,q)$. Notice that r is continuous, and not in $\triangle$, where ever it is well defined. Let $\theta(p,y)$ 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 $\lambda(p,q) = \sup\{\frac{|q'-r(p,q)|}{|p'-r(p,q)|} : V(p) = V(p'),\ V(q) = V(q')\}$. So λ is bounded away from 1 as

p and q get close to each other if r is bounded.

Thus, suppose that there exists $\bar{p}$ such that $r(p, \bar{p})$ is well defined for all p close enough to $\bar{p}$ with $V(p) > V(\bar{p})$ and $|r(p, \bar{p})|$ is unbounded as $p \to \bar{p}$. Without loss of generality, assume that the indifference curve corresponding to $\bar{p}$ is the horizontal axis of $\mathbb{R}^2$ and $r(p, \bar{p}) \in R^2_+$. Notice that $|r(p, \bar{p})| < |r(p', \bar{p})|$ and $V(p) > V(p')$ implies $|r(p, p')| < |r(p, \bar{p})|$. So the fact that $|r(p, \bar{p})|$ is unbounded around $\bar{p}$ implies that there are infinitely many points p_i, $i = 0, 1, \ldots$ such that $V(p_i) > V(p_{i+1})$ and $|r(p_i, p_{i+1})| < |r(p, \bar{p})|$, thus $\lambda(p_i, p_{i+1})$ is bounded away from 1, violating Kannai's condition.

So for any $\bar{p}$ if $r(p, \bar{p})$ is well defined for all p close enough to $\bar{p}$ with $V(p) > V(\bar{p})$, then $|r(p, \bar{p})|$ is bounded. So, if there is such a $\bar{p}$, then by continuity, $r(p, p')$ is bounded for p, p' close enough to $\bar{p}$, since $r(p, p')$ must be well defined in an open neighborhood of $\bar{p}$. Thus Kannai's condition is violated.

Thus for any $\bar{p}$, $r(p, \bar{p})$ is not well-defined (that is, $v(p) = v(\bar{p})$) for p arbitrary close to $\bar{p}$. But then, by continuity, $v(p) = v(q)$ for all p, q and this is a contradiction. □

It remains to show that given $\succeq_2$ satisfies CI, $\succeq_{lr}$ satisfies independence and $\succeq_{er}$ satisfies betweenness; the agent satisfies SAIL if and only if $v_{er}(u_{lr}^{-1}(\cdot), \overline{v})$ is convex for all $\overline{v}$ in the range of V_{er}. To see this, note that given $\succeq_2$ satisfies CI, $\succeq_{lr}$ satisfies independence and $\succeq_{er}$ satisfies betweenness; for any pair of two-stage lotteries of the form $X = [(F_1, q_1; \ldots; F_j, q_j; \ldots; F_N, q_N)]$ and $Y = [(F_1, q_1; \ldots; F_{j-1}, q_{j-1}; F'_j, \beta q_j; F''_j, (1-\beta)q_j; F_{j+1}, q_{j+1}; \ldots; F_N, q_N)]$ in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$, where $q_j > 0$ and $F_j = \beta F'_j + (1-\beta)F''_j$, $Y \succeq_2 X$ if and only if $\beta v_{er}(u_{lr}^{-1}(\bar{u}'), \overline{v}) + (1-\beta) v_{er}(u_{lr}^{-1}(\bar{u}''), \overline{v}) \geq v_{er}(u_{lr}^{-1}(\beta\bar{u}' + (1-\beta)\bar{u}'', \overline{v}))$ where $\bar{u}' = V_{lr}(F')$, $\bar{u}'' = V_{lr}(F'')$ and $\overline{v} = V_{er}(X)$. So the agent satisfies SAIL if and only if $\beta v_{er}(u_{lr}^{-1}(\bar{u}'), \overline{v}) + (1-\beta) v_{er}(u_{lr}^{-1}(\bar{u}''), \overline{v}) \geq v_{er}(u_{lr}^{-1}(\beta\bar{u}' + (1-\beta)\bar{u}'', \overline{v}))$ for all β in $(0,1)$, all $\bar{u}'$ and $\bar{u}''$ in the range of V_{lr}, and all $\overline{v}$ in the range of V_{er}. ■

Proof of Proposition 4.1. For both parts (i) and (ii), MAIL implies SAIL follows from the definition of MAIL. For the other direction, fix a prior belief π and two signals $((s_1, \ldots, s_N), \lambda)$ and $((s'_1, \ldots, s'_{N'}), \lambda')$ with the former signal more informative than the latter with respect to π. Let $[(P_i, q_i)_{i=1}^{N}]$ and $[(P'_j, q'_j)_{j=1}^{N'}]$ denote the distributions of posteriors on $[0,1]$ induced by the signals $((s_1, \ldots, s_N), \lambda)$ and $((s'_1, \ldots, s'_{N'}), \lambda')$ respectively. By Lemma 2.1(iii) there exists a sequence of distributions of posteriors $([(P_i^k, q_i^k)_{i=1}^{N^k}])_{k=1}^{K}$, with $[(P_i^1, q_i^1)_{i=1}^{N^1}] = [(P'_j, q'_j)_{j=1}^{N'}]$ and $[(P_i^K, q_i^K)_{i=1}^{N^K}] = [(P_i, q_i)_{i=1}^{N}]$, such that $[(P_i^{k+1}, q_i^{k+1})_{i=1}^{N^{k+1}}]$ is a linear bifurcation of $[(P_i^k, q_i^k)_{i=1}^{N^k}]$ for $k = 1, \ldots, K-1$. That is, in each step from k to $k+1$ of the sequence, a probability mass q_j^k on one of the posteriors P_j^k is 'split' into two parts, $\beta^k q_j^k$ and $(1-\beta^k)q_j^k$, and placed on two posteriors P_{j1}^k and P_{j2}^k, for which β^k is in $[0,1]$ and $\beta^k P_{j1}^k + (1-\beta^k)P_{j2}^k = P_j^k$. For each k, let $T^k = \left\langle (s_i^k, q_i^k, P_i^k, A)_{i=1}^{N^k}, c \right\rangle$ be the Blackwell tree associated with the distributions of posteriors $([(P_i^k,\, q_i^k)_{i=1}^{N^k}])_{k=1}^{K}$.

At the first step in the sequence, there are three cases to consider. For ease of notation but without loss of generality, assume that the probability mass on a posterior that is split in the first step is $[P_1', q_1']$ and write β for β^1, so that $\beta P_{11}' + (1-\beta)P_{12}' = P_1'$. If (a) $P_{11}' \neq P_j'$ and $P_{12}' \neq P_j'$ for all $j = 2, \ldots, N'$, then $N^2 = N' + 1$. If (b), without loss of generality, $P_{11}' = P_2'$ but $P_{12}' \neq P_j'$ for all $j = 2, \ldots, N'$, then $N^2 = N'$. If (c), without loss of generality, $P_{11}' = P_2'$ and $P_{12}' = P_3'$, then $N^2 = N' - 1$.

Proof of (i). Let $(m_i')_{i=1}^{N'}$, with each m_i in $\mathcal{L}_0(A)$, denote the agent's early-resolution behavioral strategy in the Blackwell tree $T' = \left\langle (s_i', q_i', P_i', A)_{i=1}^{N'}, c \right\rangle$ induced by the signal $((s_1', \ldots, s_{N'}'), \lambda')$. For any α in $[0,1]$, and any pair of randomized actions m and m' in $\mathcal{L}_0(A)$, the mixture $[\alpha m + (1-\alpha)m']$ is a well defined randomized action; that is, it can form part of an early-resolution behavioral strategy.

In case (a), $[(P_i^2,\, q_i^2)_{i=1}^{N^2}]$ can be written as $[P_{11}', \beta q_1';\, P_{12}', (1-\beta)q_1';\, P_2', q_2';\, \ldots;\, P_{N'}', q_{N'}']$. Since $N^2 = N'+1$, any signal that induces $[(P_i^2,\, q_i^2)_{i=1}^{N^2}]$ must have at least $N'+1$ resolutions. Thus, the early-resolution behavioral strategy with $N'+1$ components, $(m_1', m_1', m_2', \ldots, m_{N'}')$ is feasible in the tree $T^2 = \left\langle (s_i^2, q_i^2, P_i^2, A)_{i=1}^{N^2}, c \right\rangle$.

In case (b), $[(P_i^2,\, q_i^2)_{i=1}^{N^2}]$ can be written as $[P_2', q_2' + \beta q_1';\, P_{12}', (1-\beta)q_1';\, P_3', q_3';\, \ldots;\, P_{N'}', q_{N'}']$. Since $N^2 = N'$, any signal that induces $[(P_i^2,\, q_i^2)_{i=1}^{N^2}]$ must have at least N' resolutions. Thus, the early-resolution behavioral strategy with N' components, $([\frac{q_2'}{q_2' + \beta q_1'} m_2' + \frac{\beta q_1'}{q_2' + \beta q_1'} m_1'],\, m_1',\, m_3', \ldots, m_{N'}')$ is feasible in the tree $T^2 = \left\langle (s_i^2, q_i^2, P_i^2, A)_{i=1}^{N^2}, c \right\rangle$.

In case (c), $[(P_i^2,\, q_i^2)_{i=1}^{N^2}]$ can be written as $[P_2', q_2' + \beta q_1';\, P_3', q_3' + (1-\beta)q_1';\, P_4', q_4';\, \ldots;\, P_{N'}', q_{N'}']$. Since $N^2 = N' - 1$, any signal that induces $[(P_i^2,\, q_i^2)_{i=1}^{N^2}]$ must have at least $N' - 1$ resolutions. Thus, the early-resolution behavioral strategy with $N' - 1$ components, $([\frac{q_2'}{q_2' + \beta q_1'} m_2' + \frac{\beta q_1'}{q_2' + \beta q_1'} m_1'],\, [\frac{q_3'}{q_3' + (1-\beta) q_1'} m_3' + \frac{(1-\beta) q_1'}{q_3' + (1-\beta) q_1'} m_1'],\, m_4', \ldots, m_{N'}')$ is feasible in the tree $T^2 = \left\langle (s_i^2, q_i^2, P_i^2, A)_{i=1}^{N^2}, c \right\rangle$.

In each case, the prescribed strategy induces a two-stage lottery that is a simple linear bifurcation of the two-stage lottery resulting from the behavior $(m_i')_{i=1}^{N'}$ for the tree T'. Thus, SAIL and FI imply that $T^2 \succeq_T T'$. And, by a similar argument, $T^{k+1} \succeq_T T^k$ for all $k = 1, \ldots, K-1$. So by transitivity of $\succeq_T$ we have $T = T^K \succeq_T T'$, where T is the Blackwell tree induced by the more informative signal $((s_1, \ldots, s_N), \lambda)$. □

Proof of (ii). Let $(a_i')_{i=1}^{N'}$ in $A^{N'}$ denote the agent's pure strategy in the Blackwell tree T' and suppose that this strategy induces the two-stage lottery $[(\tilde{c}(P_i', a_i'), q_i')_{i=1}^{N'}]$.

In case (a), $[(P_i^2,\, q_i^2)_{i=1}^{N^2}]$ can be written as $[P_{11}', \beta q_1';\, P_{12}', (1-\beta)q_1';\, P_2', q_2';\, \ldots;\, P_{N'}', q_{N'}']$. Thus, the $N'+1$ component pure strategy $(a_1', a_1', a_2', \ldots, a_{N'}')$ is feasible in the tree $T^2 = \left\langle (s_i^2, q_i^2, P_i^2, A)_{i=1}^{N^2}, c \right\rangle$. This strategy induces a two-stage lottery that is a simple linear bifurcation of the two-stage lottery resulting from the strategy $(a'_i)_{i=1}^{N'}$ for the tree T'. Thus, SAIL and FI imply that $T^2 \succeq_T T'$.

In case (b), $[(P_i^2, q_i^2)_{i=1}^{N^2}]$ can be written as $[P'_2, q'_2+\beta q'_1; P'_{12}, (1-\beta)q'_1; P'_3, q'_3; \ldots; P'_{N'}, q'_{N'}]$. Thus, both the N' component pure strategy $(a'_1, a'_1, a'_3, \ldots, a'_{N'})$ and the N' component pure strategy $(a'_2, a'_1, a'_3, \ldots, a'_{N'})$ are feasible in the tree $T^2 = \left\langle (s_i^2, q_i^2, P_i^2, A)_{i=1}^{N^2}, c \right\rangle$. Let X be the two-stage lottery $[\tilde{c}(P'_2, a'_1), \beta q'_1 + q_2; \tilde{c}(P'_{12}, a'_1), (1-\beta)q'_1; \tilde{c}(P'_3, a'_3), q'_3; \ldots; \tilde{c}(P'_{N'}, a'_{N'}), q'_{N'}]$ induced by the former strategy and let X' be the two-stage lottery $[\tilde{c}(P'_2, a'_2), \beta q'_1 + q_2;$ $\tilde{c}(P'_{12}, a'_1), (1-\beta)q'_1; \tilde{c}(P'_3, a'_3), q'_3; \ldots; \tilde{c}(P'_{N'}, a'_{N'}), q'_{N'}]$ induced by the latter strategy. Let Y be the two-stage lottery $[\tilde{c}(P'_2, a'_1), \beta q'_1; \tilde{c}(P'_{12}, a'_1), (1-\beta)q'_1; \tilde{c}(P'_2, a'_2), q'_2; \ldots; \tilde{c}(P'_{N'}, a'_{N'}), q'_{N'}]$. Since Y is an early-resolution mixture of X and X', by CQV, at least one of X or X' is (weakly) preferred to Y. Moreover, Y is a simple linear bifurcation of the two-stage lottery resulting from the strategy $(a'_i)_{i=1}^{N'}$ for the tree T'. Thus, transitivity, SAIL and FI imply that $T^2 \succeq_T T'$.

In case (c), $[(P_i^2, q_i^2)_{i=1}^{N^2}]$ can be written as $[P'_2, q'_2 + \beta q'_1; P'_3, q'_3 + (1-\beta)q'_1; P'_4, q'_4; \ldots;$ $P'_{N'}, q'_{N'}]$. Thus, all four $N'-1$ component pure strategies $\mathbf{a}^{13} = (a'_1, a'_3, \ldots, a'_{N'})$, $\mathbf{a}^{23} = (a'_2, a'_3, \ldots, a'_{N'})$, $\mathbf{a}^{11} = (a'_1, a'_1, \ldots, a'_{N'})$ and $\mathbf{a}^{21} = (a'_2, a'_1, \ldots, a'_{N'})$ are feasible in the tree $T^2 = \left\langle (s_i^2, q_i^2, P_i^2, A)_{i=1}^{N^2}, c \right\rangle$. Let X^{jk} be the two-stage lottery $[\tilde{c}(P_2, a'_j), \beta q'_1 + q_2;$ $\tilde{c}(P'_3, a'_k), (1-\beta)q'_1 + q'_3; \ldots; \tilde{c}(P'_{N'}, a'_{N'}), q'_{N'}]$ induced by the strategy $\mathbf{a}^{jk}$, for $j = 1, 2$ and $k = 1, 3$. Let Y^{12k} be the two-stage lottery $[\tilde{c}(P'_2, a'_1), \beta q'_1; \tilde{c}(P'_2, a'_2), q'_2; \tilde{c}(P'_3, a'_k), (1-\beta)q'_1 + q'_3;$ $\ldots; \tilde{c}(P'_{N'}, a'_{N'}), q'_{N'}]$, for $k = 1, 3$. But Y^{12k} is an early-resolution mixture of X^{1k} and X^{2k}. Thus by CQV, at least one of X^{1k} or X^{2k} is (weakly) preferred to Y^{12k}. Let Z be the two-stage lottery $[\tilde{c}(P'_2, a'_1), \beta q'_1; \tilde{c}(P'_2, a'_2), q'_2; \tilde{c}(P'_3, a'_1), (1-\beta)q'_1; \tilde{c}(P'_3, a'_3), q'_3; \ldots;$ $\tilde{c}(P'_{N'}, a'_{N'}), q'_{N'}]$. But Z is an early-resolution mixture of Y^{121} and Y^{123}. Thus, by CQV, at least one of Y^{121} or Y^{123} is (weakly) preferred to Z. Hence by transitivity at least one of the two-stage lotteries X^{13}, X^{23}, X^{11}or X^{21} is weakly preferred to Z. Moreover, Z is a simple linear bifurcation of the two-stage lottery resulting from the strategy $(a'_i)_{i=1}^{N'}$ for the tree T'. Thus, transitivity, SAIL and FI imply that $T^2 \succeq_T T'$.

By analogous reasoning, $T^{k+1} \succeq_T T^k$ for all $k = 1, \ldots, K-1$, and so by FI and transitivity, we have $T \succeq_T T'$. ■

Proof of Proposition 4.2. Follows by the argument used in the proof of Proposition 4.1, part (ii), case (a).

References

Blackwell, David (1953): "Equivalent Comparison of Experiments", *Annals of Mathe-*

matics and Statistics, 24, 265-272.

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 G. 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, Larry G. Epstein and Uzi Segal (1991): "Mixture Symmetry and Quadratic Utility", *Econometrica*, 59, 139-163.

Chew, Soo Hong and T. Ho (1994): "Hope: An Empirical Study of Attitude Toward the Timing of Uncertainty Resolution", *Journal of Risk and Uncertainty*, 8 267-288.

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.

Chew, Soo Hong and Naoko Nishimura (1992): "Differentiability, Comparative Statics and Non-Expected Utility Preferences", *Journal of Economic Theory*, 56, 294-312.

Dekel, Eddie (1986): "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom", *Journal of Economic Theory*, 40, 304-318.

Epstein, Larry G. and Stanley E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework", *Econometrica*, 57, 937-969.

Epstein, Larry G. and Stanley E. 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 (1992a): "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 (1992b): "Many Good Risks: An Interpretation of Multivariate Risk and Risk Aversion without the Independence Axiom", *Journal of Economic Theory*, 56, 338-351.

Hardy, Godfrey H., John E. Littlewood and George Polya (1934): *Inequalities*, Cambridge University Press.

Johnsen, Thore H. and John B. Donaldson (1984): "The Structure of Intertemporal Preferences under Uncertainty and Time Consistent Plans", *Econometrica*, 6, 1451-1458.

Kannai, Yakar, (1976): "Concavifiability and Constructions of Concave Utility Functions", *Journal of Mathematical Economics*, 4, 1-56.

Karni, Edi and David Schmeidler (1991): "Atemporal Dynamic Consistency and Expected Utility Theory", *Journal of Economic Theory*, 54, 401-408.

Kelsey, David (1989): "Comment on 'Nonexpected utility as aversion of information"', *Journal of Behavioral Decision Making*, 2: 199-200.

Kreps, David M. and Evan L. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory", *Econometrica*, 46, 185-200.

Kreps, David M. and Evan L. Porteus (1979): "Temporal von Neumann-Morgenstern and Induced Preferences", *Journal of Economic Theory*, 20, 81-109.

LaValle, Irving H. and Kenneth R. Wapman (1985): "Rolling Back Decision Trees Requires the Independence Axiom!", *Management Science*, 32, 382-385.

Machina, Mark (1982): "Expected Utility Analysis without the Independence Axiom", *Econometrica*, 50, 277-323.

Machina, Mark (1984): "Temporal Risk and the Nature of Induced Preferences", *Journal of Economic Theory*, 33, 199-231.

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.

McClennen, Edward F. (1990): *Rationality and Dynamic Choice: Foundational Explorations*, Cambridge University Press.

Quiggin, John (1982): "A Theory of Anticipated Utility", *Journal of Economic Behavior and Organization*, 3, 323-43.

Sarin, Rakesh and Peter Wakker (1993): "Consistency in Dynamic Choice Situations with Nonexpected Utility Models", mimeo.

Sarin, Rakesh and Peter Wakker (1994): "Folding Back in Decision Tree Analysis", *Management Science*, 40, 625-628.

Schlee, Edward (1991): "The Value of Perfect Information in Nonlinear Utility Theory", *Theory and Decision*, 30, 127-131.

Schlee, Edward (1990): "The Value of Information in Anticipated Utility Theory", *Journal of Risk and Uncertainty*, 3, 83-92.

Segal, Uzi (1990): "Two-Stage Lotteries without the Reduction Axiom", *Econometrica*, 58, 349-377.

Siebert, Charles (1995): "Living with Toxic Knowledge: The DNA We've Been Dealt", *New York Times Magazine*, Sept 17.

Wakker, Peter (1988): "Nonexpected Utility as Aversion of Information", *Journal of Behavioral Decision Making*, 1, 169-175.

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.