

Intrinsic Preference for Information*

Simon Grant

*Faculty of Economics and Commerce, Australian National University,
Canberra ACT, 0200 Australia
Simon.Grant@anu.edu.au*

Atsushi Kajii

*Institute of Policy and Planning Sciences, University of Tsukuba,
Tsukuba, Japan*

and

Ben Polak

Economics Department, Yale University, New Haven, Connecticut

Received August 22, 1996; revised April 13, 1998

Suppose agents value information not only to make contingent plans but also intrinsically. How are such attitudes toward information related to attitudes toward risk? We generalize the Kreps–Porteus recursive expected utility model, dropping both recursivity and expected utility. There is a geometric analogy between risk and information. We characterize intrinsic information loving, in general, by a substitution property analogous to multivariate risk loving; and, for smooth preferences, by the convexity of Gateaux derivatives. Even with recursivity, preference for information does not imply expected utility: we provide an example. We examine connections between information loving and risk aversion for early- and late-resolving risks. *Journal of Economic Literature* Classification Numbers: D80, D81. © 1998

Academic Press

Key Words: information; risk-aversion; anxiety; non-expected utility; Gateaux.

* An earlier version of this paper included a detailed analysis of intrinsic preference for information in the recursive rank-dependent and betweenness models, and also included instrumental aspects of preference for information. For these, see Grant, Kajii, and Polak [20, 21, 22]. The geometric interpretation of information used here was suggested to us in some classes of Jerry Green. A sub-editor and an anonymous referee provided helpful comments. We also thank Eddie Dekel, Al Klevorick, Ngo Van Long, Stephen Morris, Boaz Moselle, David Pearce, Matthew Ryan, David Schwartz, Tim van Zandt, Peter Wakker, and seminar participants in four continents. Parts of this research were undertaken when Grant was visiting Yale, Kajii was visiting CORE, and Polak was visiting ANU. We thank these institutions for their hospitality. Kajii also gratefully acknowledges support from the U. Penn Institute for Economics Research Fund. We are responsible for any errors.

1. INTRODUCTION

In the standard model of individual choice, all agents at least weakly prefer more information to less. This idea was formalized by Blackwell [4, 5]. In the standard account, however, the agents' preference for information is only instrumental. That is, they like information only because it lets them design better strategies. If they do not or cannot condition their actions on what they learn, then information is of no value to them. Introspection suggests, however, that we sometimes intrinsically (that is, for its own sake) prefer more information to less, even in the absence of any instrumental purpose. Moreover, even in the absence of strategic concerns, we sometimes choose not to be informed. Consider, for example, the decision of whether to be tested for an incurable genetic disorder. A director of a genetic counseling program 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 [41, p. 52]

Most of us can find examples from our own lives either where information had some intrinsic value or where we would have preferred not to know something. Anxious job market candidates might pay good money to be told whether they are destined for the dole queue, not because they plan to act on the information but simply because they prefer to know. Expectant parents, on the other hand, might pay good money not to be told the sex of their future child, simply preferring to remain ignorant. Further afield, there is a debate among US Tort lawyers regarding whether plaintiffs should be able to claim damages for the current anxiety of not knowing whether they will suffer future physical effects from, say, exposure to the HIV virus.² Recognizing that people may be willing to pay for information to avert such anxiety may justify assessing damages for such "psychic costs." Including the intrinsic value of information may affect cost-benefit

¹ We thank David Pearce for this reference. David Kelsey [24] discusses the example of incurable diseases in a comment on a paper by Peter Wakker [43] on aversion to information.

² See, for example, Gale and Goyer [17] and Franklin and Rabin [16, pp. 226–234 and 311–322].

studies of scientific research. Including the intrinsic value of ignorance may affect cost-benefit studies of holiday gift giving.³

Kreps and Porteus's [25] recursive expected utility model, unlike the standard atemporal expected utility model, allows for the idea that agents might have intrinsic preference for information. The model is still restrictive, however, in that it maintains the independence axiom, the central assumption of expected utility theory. There are at least two reasons to want to consider attitudes toward information beyond the realm of expected utility. First, the examples above suggest that intrinsic information loving may be related to risk aversion. The expected utility model, however, is quite narrow in the way it models attitudes towards risk. Thus, to explore the connections between attitudes toward risk and toward information, it is restrictive and perhaps misleading to confine attention to expected utility.

Second, Blackwell showed that if we make standard assumptions about preferences, we can conclude that more information is always preferred to less. The other direction, however, is an open question. That is, suppose we assume that an individual always (at least weakly) prefers more information to less. What implications does this have for the overall shape of their preferences? For example, Wakker [43] suggests that such preference for information might imply expected utility. We cannot explore this question unless we start from a broader class of preferences.

This paper, then, considers preference for information without assuming expected utility. We focus on attitudes towards information in the absence of contingent choices, that is, where the agent has only one action available. In this setting, instrumental purposes for information are excluded so any preference for (or aversion to) information is intrinsic. We consider both the connections between intrinsic attitudes toward information and risk, and the implications of assuming that more information is always (weakly) preferred. A companion paper (Grant, Kajii, and Polak [20]) extends the analysis to the many-action, intrinsic cum instrumental, case.

There are two quite different connections between information and risk. First, there is a direct geometric analogy. This analogy has long been known. It is implicit, for example, in Blackwell and Girshick [6]. Adding risk to a lottery causes the distribution of possible outcomes to be more "spread out" in the outcome space. We can think of signals as lotteries over induced posterior beliefs. Given the action taken by the agent, each posterior induced by a signal is itself associated with a lottery over outcomes. Thus, signals induce lotteries over lotteries over outcomes: two-stage lotteries.

³ See Waldfogel [44]. Chew and Ho [9] discuss how both hope and anxiety affect well-being. They suggest that the shape of agents' preferences for information may affect the design of news programs, of punishments, of contracts, and of financial instruments.

Adding information to a signal causes the distribution of posteriors to be more “spread out” in the simplex. Thus, adding information can be thought of as adding risk to the first-stage distribution over the second-stage lotteries.

This analogy is useful. It allows us to translate known results about intrinsic attitudes toward risk aversion into new results about intrinsic attitudes toward information. For example, assuming an agent is risk averse or risk loving places restrictions on the shape of her preferences over one-stage lotteries. Assuming an agent is information averse or information loving places analogous restrictions on the shape of her preferences over two-stage lotteries. Below, using this trick, we provide necessary and sufficient conditions for intrinsic preference for information for a wide range of non-expected utility models. With this characterization result in hand, we construct an example to show that always preferring more information to less does not imply (even recursive) expected utility.

The second connection between risk and information concerns preference itself. Part of an agent’s aversion to risk may be because she dislikes living with risk over time.⁴ If so, she has reason to want information not (or not only) to fine-tune her decisions but to resolve uncertainty earlier. For instance, suppose that an agent is deciding today between a fixed payout at the time of her retirement (adjusted for inflation so there is no risk), or receiving at that time a payout based on her, as yet unknown, final year’s salary. One may view the difference between the expected payout of the second option and the fixed payout as a risk premium. Allowing for intrinsic attitudes toward information, the risk premium she would be willing to pay depends on the time between now and her retirement, the time she would have to live with the risk. Specifically, if she is an information lover, then the further into the future is her retirement date, the greater is the risk premium she would be willing to accept to eliminate the retirement payout risk. For the case of recursive expected utility, this implication extends to partial removals of future risks, and the implication also works the other way: if she would pay more for the removal of risk when her retirement is distant then she is an information lover. Below, we show to what degree these results do and do not generalize.

In this paper, all results are stated primarily for information loving but this is just for expositional ease. The corresponding results for information aversion are implicit. More generally, whether an agent prefers more or less information might depend on the outcomes at stake. At some notational

⁴ Pope [33, 34] stresses the importance of distinguishing between what she calls the “pre-outcome period” (when a decision maker is still uncertain) and the “outcome period” (when all uncertainty has been resolved). She contends that, by conflating these two periods, the standard models omit welfare considerations that stem directly from an individual’s not knowing the outcome she will receive.

cost, however, most of the results could be adapted to “local” statements; for example, we could characterize what it means to be an information lover for a particular set of outcomes. At still greater notational cost, we could even localize the results to subsets of simplexes. If, however, an agent is, say, an information lover for “losses” but an information averter for “gains,” we would need our notation to keep track of some “status quo” outcome or expected ex post utility level. This would take us significantly beyond our current framework.

For most of the analysis below, we have in mind that the resolution of uncertainty takes place in real time. Indeed, each of the examples with which we began involved some passage of time. 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. In this paper, we do not explicitly model the passage of time.⁵ This allows our formal analysis to apply equally to the pure sequence case, the case of short time delays between stages, or the case of long time delays. Of course, the plausibility of each assumed property of preference may change according to the interpretation. Some might prefer to use the term “preference for early resolution” for the real time case, reserving “preference for information” for the pure sequence case. We do not feel strongly about this, and since both cases are reduced to the same analysis in our framework, we use both terminologies.⁶

Related work includes Chew and Epstein [7], who extend Kreps and Porteus’s [25] analysis of preference for early or late resolution of uncertainty to a recursive betweenness model. Epstein and Zin [14, 15] extend both models and apply them to data. Cook [12], Chew and Ho [9] and Ahlbrecht and Weber [1, 2] present experimental evidence on temporal uncertainty resolution. Wakker [43], Machina [31], Schlee [36, 37], and Safra and Sulganik [35] each consider preference for information in non-expected utility models. They confine attention, however, to instrumental preference. Schlee [38] and Skiadas [42] extend this instrumental analysis to subjective uncertainty.

Some of the formal ideas about preferences over two-stage lotteries used below were first introduced in Segal [39]. Among many other things, Segal considers preferences over two-stage lotteries that respect different notions

⁵ Indeed, we confine attention to two notional stages, “early” and “late,” in which uncertainty could be resolved. The extension to any finite number of stages, however, is straightforward; see Grant, Kajii, and Polak [21].

⁶ Our notation could be extended in the standard way to keep track of discrete time. In that case, we could also keep track of consumption in earlier periods. We abstract from this for simplicity. Notice, however, that our framework is already sufficiently rich to consider payments for information in terms of the final period consumption good.

of first-order stochastic dominance. Information ranking is a notion of second-order stochastic dominance. Thus, our work can be seen as an extension of Segal's study.

2. GROUNDWORK

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

DEFINITION 1. Suppose \mathcal{Z} is itself a convex set in a linear space. An *elementary linear bifurcation* of a measure $[(\zeta_j, p_j)_{j=1}^M]$ in $\mathcal{L}_0(\mathcal{Z})$ is a measure of the form $[\zeta_1, p_1; \dots; \zeta_{j-1}, p_{j-1}; \zeta'_j, \beta p_j; \zeta''_j, (1 - \beta)p_j; \zeta_{j+1}, p_{j+1}; \dots; \zeta_M, p_M]$ in $\mathcal{L}_0(\mathcal{Z})$, where β is in $[0, 1]$, and $\zeta_j = \beta\zeta'_j + (1 - \beta)\zeta''_j$.

Before we talk about preference for information, we should say what we mean by information. We use the standard definition, due to Blackwell [4, 5]. Let Ω be a set of states of nature. In the following, unless otherwise stated, Ω is taken to the interval $[0, 1]$. Thus, $\mathcal{L}(\Omega)$ is the set of probability measures on the unit interval. A signal consists of a (finite) set of possible realizations, denoted S , and a likelihood function, $\lambda: S \times \Omega \rightarrow [0, 1]$, where for any state ω in Ω , $\sum_{s \in S} \lambda(s | \omega) = 1$. We then define "increasing information" as follows.

DEFINITION 2. The signal (S, λ) is *more informative than* the signal (S', λ') with respect to the prior belief π in $\mathcal{L}(\Omega)$ if there exists a function $\alpha: S' \times S \rightarrow [0, 1]$ such that $\sum_{s'} \alpha(s', s) = 1$ for all realizations s in S , and $\lambda'(s' | \omega) = \sum_s \alpha(s', s) \lambda(s | \omega)$ for π -almost all states ω in Ω and all realizations s' in S' .

The definition says we can construct the likelihood function, λ' , of the “less informative” signal directly from the likelihood function, λ , of the “more informative” signal and some function α , which does not itself depend on the state ω . We can think of $\alpha(s', s)$ as the conditional probability of observing realization s' from the less informative signal given that we would have observed realization s from the more informative signal. That is, loosely speaking, we can think of the “less informative” signal as a “garbled” version of the “more informative” signal.⁷ If (S, λ) is more informative than the signal (S', λ') with respect to all prior beliefs π in $\mathcal{L}(\Omega)$ then (S, λ) is a “sufficient statistic” for (S', λ') .

Blackwell considered decision problems in which an agent has a prior over the states of nature and has available to her a (closed) set of actions. The agent gets to observe the realization of a signal before choosing an action from this set. The final outcome or consequence depends jointly on the action taken by the agent and on the state of nature. The issue Blackwell addressed was which signals an agent would prefer. Formally, 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, unless otherwise stated, we take \mathcal{X} be the interval $[0, 1]$.⁸ Let \mathcal{C} be the set of measurable functions, $c: \Omega \times \mathcal{A} \rightarrow \mathcal{X}$, that assign a consequence to each state-action pair.

Loosely speaking, one direction of Blackwell’s theorem⁹ says that, regardless of their prior beliefs π in $\mathcal{L}(\Omega)$, their actions sets $A \subset \mathcal{A}$, and the consequence function c in \mathcal{C} , all atemporal expected utility maximizers¹⁰ (weakly) prefer more informative to less informative signals. Moreover, if the agent’s action set is a singleton, then she is indifferent between signals. In this paper, we drop the assumption of atemporal expected utility theory to escape this second conclusion; that is, to allow for intrinsic preference for information. We do not discuss the case where the agent has many actions from which to choose. An extension of the analysis here to this many-action case, however, is available in Grant, Kajii, and Polak [22].

If agents have only one action available to them, they cannot condition their actions on what they observe. In this case, given the action, a prior and a consequence function, each signal may be identified with a two-stage lottery. The first stage is a probability distribution over the possible realizations

⁷ Strictly speaking, the definition above is more general than a garbling: see Marschak and Radner [32, pp. 64–67].

⁸ The analysis can readily be extended to outcome sets that are general compact metric spaces if we assume that all welfare-relevant risk can be characterized as risk over the ranks of outcomes; see Grant, Kajii, and Polak [18].

⁹ Bohnenblust, Shapley, and Sherman [2] derived a closely related result.

¹⁰ By atemporal expected utility maximizer, we mean that the agent’s preferences satisfy the von Neuman–Morgenstern assumptions plus, in the context of multi-stage lotteries, the reduction axiom (defined below).

of the signal, and each second-stage lottery is the distribution over outcomes associated with the posterior induced by a possible realization.

To formalize this, first notice that, with the weak topology, $\mathcal{L}(\mathcal{X})$ is a compact metric space. So, $\mathcal{L}(\mathcal{L}(\mathcal{X}))$ is a compact metric space with the weak topology. With slight abuse of terminology, we will often refer to elements of $\mathcal{L}(\mathcal{X})$ as distributions or one-stage lotteries, and to elements of $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$ as simple two-stage lotteries. Now, fix a prior belief π in $\mathcal{L}(\Omega)$, a single action a in \mathcal{A} , and a consequence function c in \mathcal{C} . From the prior belief π and the signal (S, λ) , we can compute the unconditional probability of observing each realization. We can ignore signal realizations that occur with zero probability, so let $(s_1, \dots, s_N) \subseteq S$ denote the list of signal realizations that occur with positive probability. For each such s_i , let q_i in $(0, 1]$ denote this probability, and let P_i in $\mathcal{L}(\Omega)$ denote the posterior belief over states induced by that realization. Given the posterior belief P_i , the action a induces a probability measure over consequences via the consequence function c . Let $H(P_i, c, a)$ be this probability measure where, for any Borel subset B of \mathcal{X} , $H(P_i, c, a)(B) = P_i(\{\omega \in \Omega: c(\omega, a) \in B\})$. Thus $[(H(P_i, c, a), q_i)_{i=1}^N]$ is a two-stage lottery in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$.¹¹

Let \succsim_2 denote an agent's (complete and transitive) preference relation over the set of simple two-stage lotteries, $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$. We assume that an agent's attitude towards information can be characterized by these preferences. That is, the agent uses \succsim_2 to construct a preference ordering over signals according to the two-stage lotteries that the signals induce. This assumption excludes framing effects. For example, if two signals induce the same two-stage lottery, the agent is indifferent between them. We assume throughout that \succsim_2 is continuous as a relation on $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$.

In standard models of individual choice (including Blackwell's), it is implicitly assumed that preferences over two-stage lotteries satisfy the reduction of compound lottery axiom. For any two-stage lottery, $X = [(F_i, q_i)_{i=1}^N]$ in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$, let $\rho(X)$ in $\mathcal{L}(\mathcal{X})$ be the reduced one-stage lottery such that, for each Borel subset B of \mathcal{X} , $\rho(X)(B) = \sum_{i=1}^N q_i F_i(B)$.

DEFINITION 3. An agent's preference relation over two-stage lotteries, \succsim_2 , satisfies *reduction* if for all pairs of two-stage lotteries X and X' in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$: $\rho(X) = \rho(X')$ implies $X \sim_2 X'$.

In this paper we do not assume reduction. The reason is that reduction imposes intrinsic indifference for information. Suppose, for example, that a student has just taken an exam which she believes she has passed with probability three-quarters. She has the choice between two delivery systems (signals): a fast one that delivers the result tonight, or a slow one that

¹¹ This is the element of $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$ that for each F in $\mathcal{L}(\mathcal{X})$ assigns probability $\sum_i: H(P_i, c, a) = F$ q_i .

delivers it tomorrow. That is, tonight, under the fast system she will be informed but under the slow system she will not. She has no decisions to make tonight that depend in any way on her information. Both of the signals induce the same probability over the final outcomes, pass or fail; that is, $(\frac{3}{4}, \frac{1}{4})$. Therefore, if we impose that student's preferences satisfy reduction, then she must be indifferent between getting the information tonight or not. We want to allow, however for the possibility that, tonight, the student would strictly prefer to be informed.¹²

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. In the above example, the two-stage lottery induced by the fast delivery system is an early-resolution lottery, while that induced by the slow system is a late-resolution lottery. Let \succsim_{er} denote the restriction of two-stage lottery preferences, \succsim_2 , to early-resolution lotteries, and let \succsim_{lr} denote the restriction of \succsim_2 to late-resolution lotteries. Both \succsim_{er} and \succsim_{lr} inherit continuity from \succsim_2 .

Both early- and late-resolution lotteries are subsets of the set of two-stage lotteries. However, since early-resolution lotteries are degenerate in the second stage, and late-resolution lotteries are degenerate in the first stage, both sets are isomorphic to the set of one-stage lotteries.¹³ That is, each one-stage lottery $F = [(x_i, p_i)_{i=1}^N]$ in $\mathcal{L}_0(\mathcal{X})$ is naturally associated with the early-resolution lottery $[(\delta_{x_i}, p_i)_{i=1}^N]$ and the late-resolution lottery $[(x_i, p_i)_{i=1}^N, 1]$, both of which have F as their reduction. As we are not assuming reduction, the agent need not be indifferent between $[(\delta_{x_i}, p_i)_{i=1}^N]$ and $[(x_i, p_i)_{i=1}^N, 1]$.

Given these isomorphisms the preference relations \succsim_{er} and \succsim_{lr} may (or may not) be endowed with familiar properties of preferences over one-stage lotteries such as independence. More precisely, let \succsim_1 be a generic preference relation over sets that are isomorphic to $\mathcal{L}(\mathcal{X})$.

DEFINITION 4. The preference relation \succsim_1 satisfies *independence* if, for all F_1, F_2 and F_3 in $\mathcal{L}(\mathcal{X})$ and all α in $(0, 1]$, $F_1 \succsim_1 F_2$ if and only if $\alpha F_1 + (1 - \alpha) F_3 \succsim_1 \alpha F_2 + (1 - \alpha) F_3$.

Unless explicitly stated, we do not assume that the preferences over either early- or late-resolution lotteries satisfy independence. We assume throughout, however, that both \succsim_{er} and \succsim_{lr} respect first-order stochastic dominance.

¹² Some writers do not use the term reduction where time is involved. Notice, however, that in this example, all preferences and information refer to the same moment in time, "tonight."

¹³ Strictly speaking, the set of early-resolution lotteries are isomorphic to $\mathcal{L}_0(\mathcal{X})$ and the set of late-resolution lotteries are isomorphic to $\mathcal{L}(\mathcal{X})$.

In this paper, we reserve the term independence to refer to a substitution property of p preferences over sets isomorphic to the set of one-stage lotteries.¹⁴ Thus, independence involves “within-stage” probability mixtures of either early- or late-resolution lotteries. By contrast, the following substitution property involves replacing one second-stage lottery with another within a two-stage lottery.

DEFINITION 5. An agent’s preference relation over two-stage lotteries, \succsim_2 , satisfies *recursivity* if for all pairs of two-stage lotteries of the form $X = [(F_i, q_i)_{i=1}^N]$ and $Y = [(F_1, q_1; \dots; F_{j-1}, q_{j-1}; F'_j, q_j; F_{j+1}, q_{j+1}; \dots; F_N, q_N)]$ in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$, with $q_j > 0$: $X \succsim_2 Y$ if and only if $[F_j, 1] \succsim_{tr} [F'_j, 1]$.^{15, 16}

Recursivity superficially resembles the independence axiom. Indeed, if the agent satisfies reduction, then recursivity implies that she is an atemporal expected utility maximizer. Without reduction, however, recursivity does not imply that the agent’s preferences over either early- or late-resolution lotteries satisfy independence. If an agent’s preferences over two-stage lotteries are recursive and both her preferences over early- and over late-resolution lotteries satisfy independence, we say that the agent satisfies recursive expected utility. Notice, however, that recursive expected utility, first studied by Kreps and Porteus [25], does not imply atemporal expected utility; that is, reduction need not hold.¹⁷

Recursivity is a technically convenient assumption, especially in the absence of reduction. Either property allows preferences over two- (or indeed multi-) stage lotteries to be analyzed in terms of preference relations over one-stage lotteries. Given reduction, we can first reduce each two-stage lottery, then evaluate the reduced lottery using preferences over one-stage lotteries. Given recursivity, even without reduction, we can break the evaluation of a two-stage lottery into two parts, where the first part uses only preferences over late-resolution lotteries and the second part uses only preferences over early-resolution lotteries. First, we replace each second-stage lottery by its late-resolution certainty equivalent. For any F in $\mathcal{L}(\mathcal{X})$, the certainty equivalent of F with respect to the preference relation \succsim_1 , $CE_1(F)$ is the outcome in \mathcal{X} such that $\delta_{CE_1(F)} \sim_1 F$. That is, by definition,

¹⁴ Segal [39] refers to this as “mixture independence.”

¹⁵ The $(F_i)_{i=1}^N$, and F'_j are not necessarily distinct. Thus, for example, given recursivity, $[F_1, 1] \succsim_2 [F_1, \frac{1}{2}; F_2, \frac{1}{2}]$ if and only if $[F_1, 1] \succsim_{tr} [F_2, 1]$.

¹⁶ Segal [39] refers to this property as “compound independence.” It is closely related to Kreps and Porteus’s [25] “temporal consistency,” LaValle and Wapman’s [27] “recursive analysis” or “rolling-back” property, Chew and Epstein’s [7] “consistency,” and Epstein’s [13] “weak recursivity.”

¹⁷ For a detailed discussion of the relation between (mixture) independence, recursivity (compound independence) and reduction, see Segal [39].

for all F in $\mathcal{L}(\mathcal{X})$, $[\delta_{\text{CE}_h(F)}, 1] \sim_{lr} [F, 1]$. So, given recursivity, any two-stage lottery $X = [(F_i, q_i)_{i=1}^N]$ is indifferent to the early-resolution lottery $[(\delta_{\text{CE}_h(F_i)}, q_i)_{i=1}^N]$. We then evaluate this lottery using \succcurlyeq_{er} . This trick is well-known and is sometimes called the fold-back or recursive method.¹⁸

While some (see, for example, Segal [40]), find recursivity intuitively appealing, others (see, for example, Machina [31]) argue that the property is too consequentialist. Therefore, in the following, recursivity is not assumed unless explicitly stated. Sometimes we will make do with the following weaker property.

DEFINITION 6. An agent's preference relation over two-stage lotteries, \succcurlyeq_2 , satisfies *conditional quasi-convexity (CQV)* if for all pairs of two-stage lotteries of the form $X = [(F_i, q_i)_{i=1}^N]$ and $Y = [(F_1, q_1; \dots; F_{j-1}, q_{j-1}; F'_j, q_j; F_{j+1}, q_{j+1}; \dots; F_N, q_N)]$ in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$, with $q_j > 0$ if $X \succcurlyeq_2 Y$ then, for all α in $(0, 1)$, $X \succcurlyeq_2 [(F_1, q_1; \dots; F_{j-1}, q_{j-1}; F_j, \alpha q_j; F'_j, (1-\alpha)q_j; F_{j+1}, q_{j+1}; \dots; F_N, q_N)]$.

Although conditional quasi-convexity is strictly weaker than recursivity (and also strictly weaker than quasi-convexity of the preference relation \succcurlyeq_2 ¹⁹), it does not meet all of Machina's [31] objections to recursivity. In particular, an agent who satisfies CQV never strictly prefers to randomize. Again, CQV is not assumed unless explicitly stated.

We will also consider the case where preferences are "smooth," by which we mean Gateaux differentiable.²⁰ More formally,

DEFINITION 7. Let W be a function on the set of probability measures, $\mathcal{L}(\mathcal{Z})$, over \mathcal{Z} , a non-empty closed subset of a metric space. For each μ in $\mathcal{L}(\mathcal{Z})$, we say that W is *Gateaux-differentiable at μ* in $\mathcal{L}(\mathcal{Z})$ if there is a measurable function $v(\cdot; \mu)$ on \mathcal{Z} such that for any ν in $\mathcal{L}(\mathcal{Z})$ and any α in $(0, 1)$,

$$W(\alpha\nu + (1-\alpha)\mu) - W(\mu) = \alpha \int v(\zeta; \mu)(\nu(d\zeta) - \mu(d\zeta)) + o(\alpha),$$

where $o(\alpha)$ is a function with the property $o(\alpha)/\alpha \rightarrow 0$ as $\alpha \rightarrow 0$. The function $v(\cdot, \mu)$ is said to be the *Gateaux derivative of W at μ* . We say that W is

¹⁸ The fact that this method can be used recursively has more bite if there are more than two stages. Notice that this method can be applied for any forms of preference we choose for \succcurlyeq_{er} and \succcurlyeq_{lr} (given continuity and respect for first order stochastic dominance), confirming that recursivity does not imply expected utility.

¹⁹ With slight abuse of terminology, by quasi-convexity of \succcurlyeq_2 we mean that $X \succcurlyeq_2 Y$ implies that $X \succcurlyeq_2 [\alpha X + (1-\alpha)Y]$ for any measures X and Y in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$ and any α in $[0, 1]$.

²⁰ See, for example, Chew and Nishimura [11], or Wang [45].

Gateaux differentiable if W is Gateaux differentiable at all μ in $\mathcal{L}(\mathcal{X})$. We refer to $v: \mathcal{X} \times \mathcal{L}(\mathcal{X}) \rightarrow \mathbb{R}$ as the Gateaux derivative of W .

For the special case where $\mathcal{L}(\mathcal{X})$ is isomorphic to the set of one-stage lotteries $\mathcal{L}(\mathcal{X})$, we use $u_1(\cdot; \cdot): \mathcal{X} \times \mathcal{L}(\mathcal{X}) \rightarrow \mathbb{R}$ to denote the Gateaux derivative associated with a Gateaux differentiable representation of the preference relation \succsim_1 . Following Machina [28], for each G in $\mathcal{L}(\mathcal{X})$, we refer to $u_1(\cdot; G)$ as the local utility function at G .

3. CHARACTERIZING INTRINSIC PREFERENCE FOR INFORMATION

We now introduce a new substitution property for preferences over two-stage lotteries.

DEFINITION 8. An agent’s preference relation over two-stage lotteries, \succsim_2 , satisfies *single-action information loving (SAIL)* if, for any pair of two stage lotteries X and Y in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$ such that Y is an elementary linear bifurcation of X , $Y \succsim_2 X$. *Single-action information aversion* and *single-action information neutrality* can be defined similarly, mutatis mutandis.

The name single action information loving is justified by the proposition below. Loosely speaking, recursivity involves the replacement of one “branch” of a two-stage lottery by another single branch. On the other hand, SAIL involves ‘splitting’ one branch into two. The latter substitutions have a natural geometric interpretation, illustrated in Fig. 1.

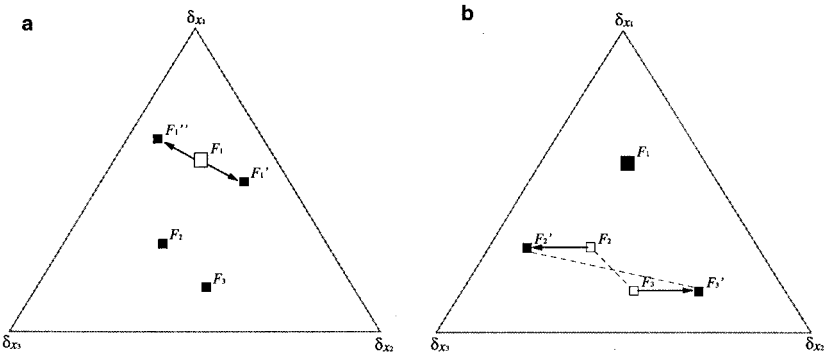


FIG. 1. An example of an elementary bifurcation and an example that is not.

The simplex in Figs. 1a and 1b is the set of lotteries over the outcomes, x_1 , x_2 and x_3 . We can think of each point, F , in this simplex as representing a second-stage lottery over these three outcomes. Each square represents a first-stage probability mass, q , on such second-stage lottery, F . Thus, a distribution of such squares represents a two-stage lottery. For example, the corners of the simplex represent degenerate second-stage lotteries, so a distribution with squares only at the corners would be an early-resolution two-stage lottery. Conversely, a distribution consisting of only one large square at some F in the simplex represents the late-resolution two-stage lottery $[F, 1]$. Consider the two-stage lottery $X = (F_1, \frac{1}{2}; F_2, \frac{1}{4}; F_3, \frac{1}{4})$. This is shown in Fig. 1a by the two small (closed) squares at F_2 and F_3 and the larger (open) square at F_1 . Next consider the two-stage lottery $Y = (F'_1, \frac{1}{4}; F''_1, \frac{1}{4}; F_2, \frac{1}{4}; F_3, \frac{1}{4})$. This is shown, also in Fig. 1a, by the same two small (closed) squares at F_2 and F_3 , and the two new small (closed) squares at F'_1 and F''_1 . Notice that Y may be obtained from X by “splitting” the first-stage probability mass of $\frac{1}{2}$ at F_1 into two masses of $\frac{1}{4}$ each at F'_1 and at F''_1 . That is, $F_1 = \frac{1}{2}F'_1 + \frac{1}{2}F''_1$. Thus, the two-stage lottery Y is an elementary linear bifurcation of X . If an agent’s preferences over two-stage lotteries satisfy SAIL then she prefers Y to X .

Figure 1b also shows a pair of two-stage lotteries. The two-stage lottery $(F_1, \frac{1}{2}; F_2, \frac{1}{4}; F_3, \frac{1}{4})$ is represented by the large (closed) square at F_1 and the two smaller (closed) squares at F_2 and F_3 . The two-stage lottery $(F_1, \frac{1}{2}; F'_2, \frac{1}{4}; F'_3, \frac{1}{4})$ is represented by the same large (closed) square at F_1 and the two new smaller (closed) squares at F'_2 and F'_3 . This second two-stage lottery looks more spread out than the first, but it can not be reached from the first by a sequence of elementary linear bifurcations. In particular, the second-stage lotteries F_2 , F_3 , F'_2 and F'_3 are not colinear. In this case, SAIL is not enough to imply preference for the second lottery over the first. That is, there exists a (continuous) preference relation over two-stage lotteries that satisfies SAIL but for which the seemingly more spread out lottery is preferred.²¹ Thus, SAIL is a discontinuous property, but it is exactly what we need to describe intrinsic preference for more informative signals.

The following proposition collects together implications of intrinsic information loving: for general preferences, for conditionally quasi-convex preferences, for smooth preferences; for smooth recursive preferences, and for recursive expected utility. The proof is relegated to the appendix but some informal intuition for the results is discussed below.

²¹ To construct an example, loosely speaking, first take a convex piece-wise linear function defined on the simplex of second-stage lotteries $\mathcal{L}(\{x_1, x_2, x_3\})$, with the function’s only “kink” along the line containing F'_2 and F'_3 . Then, construct the preferences over $\mathcal{L}_0(\mathcal{L}(\{x_1, x_2, x_3\}))$ represented by first evaluating this function at every second-stage lottery and then taking the expectation with respect to the first-stage probabilities.

PROPOSITION 1. (i) *The following two conditions are equivalent:*

(A) *For all priors π in $\mathcal{L}(\Omega)$, all c in \mathcal{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 two-stage lottery induced by (S, λ) is weakly preferred to that induced by (S', λ') .*

(B) *The agent's preference relation over two-stage lotteries, \succcurlyeq_2 , satisfies SAIL.*

(ii) *Suppose the agent's preference relation \succcurlyeq_2 satisfies conditional quasi-convexity. Then, \succcurlyeq_2 satisfying SAIL implies that the agent's late resolution preference relation \succcurlyeq_{lr} is quasi-convex in the probabilities; that is, for all F_1 and F_2 in $\mathcal{L}(\mathcal{X})$ and all α in $[0, 1]$, if $[F_1, 1] \succcurlyeq_{lr} [F_2, 1]$ then $[F_1, 1] \succcurlyeq_{lr} [\alpha F_1 + (1 - \alpha) F_2, 1]$.*

(iii) *Suppose the agent's preference relation \succcurlyeq_2 can be represented by a Gateaux differentiable function, $W: \mathcal{L}_0(\mathcal{L}(\mathcal{X})) \rightarrow \mathbb{R}$, with Gateaux derivative $v: \mathcal{L}(\mathcal{X}) \times \mathcal{L}_0(\mathcal{L}(\mathcal{X})) \rightarrow \mathbb{R}$. Then, \succcurlyeq_2 satisfies SAIL if and only if $v(\cdot; X)$ is convex at any X in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$.*

(iv) *Suppose the agent's preference relation \succcurlyeq_2 satisfies recursivity and her preference relation over early-resolution lotteries, \succcurlyeq_{er} can be represented by a Gateaux differentiable utility function with Gateaux derivative $u_{er}(\cdot; \cdot): \mathcal{X} \times \mathcal{L}_0(\mathcal{X}) \rightarrow \mathbb{R}$. Then, \succcurlyeq_2 satisfies SAIL if and only if the compound function $u_{er}(\text{CE}_{lr}(\cdot); G)$ from the set of late-resolution lotteries to the reals is convex for all lotteries G in $\mathcal{L}_0(\mathcal{X})$.*

(v) (Kreps and Porteus, [25]) *Suppose the agent's preference relation \succcurlyeq_2 satisfies recursivity and her preference relations over early- and late-resolution lotteries, \succcurlyeq_{er} and \succcurlyeq_{lr} , satisfy independence with von Neumann–Morgenstern utility indices u_{er} and u_{lr} , respectively (that is, the agent satisfies recursive expected utility). Then, \succcurlyeq_2 satisfies SAIL if and only if the compound function $u_{er} \circ u_{lr}^{-1}$ is convex*

Similar results hold for single action information aversion and neutrality, mutatis mutandis.

To get some intuition for these results, consider again the pair of two-stage lotteries in Fig. 1a. We can think of the two-stage lottery $Y = (F'_1, \frac{1}{4}; F''_1, \frac{1}{4}; F_2, \frac{1}{4}; F_3, \frac{1}{4})$ as being induced, given a prior and a consequence function, by a signal with four possible realizations, s'_1, s''_1, s_2 and s_3 . We can think of the two-stage lottery $X = (F_1, \frac{1}{2}; F_2, \frac{1}{4}; F_3, \frac{1}{4})$ being induced by a signal that is similar except that now the agent can not distinguish between realizations s'_1 and s''_1 , so that the second-stage lottery F_1 induced by realization s_1 (the “convolution” of s'_1 and s''_1) lies between F'_1 and F''_1 . Clearly, this second signal is less informative than the first. Thus, an agent who prefers more to less informative signals must prefer Y to X . The proof that SAIL is necessary for preference for information

simply formalizes this argument. To prove that SAIL is also sufficient, it is enough to show that, for any prior and consequence function, the two-stage lottery induced by a less informative signal can be obtained from that induced by a more informative signal by a finite sequence of elementary linear bifurcations. This is accomplished by means of a lemma in the appendix.

Figure 1a is also suggestive of the analogy between risk and information, mentioned in the introduction. The pair of two-stage lotteries, X and Y , looks a lot like a pair of multivariate distributions, where the multi-dimensional simplex becomes some subset of a multi-dimensional commodity space. Thus, each second-stage lottery, F_i in some two-stage lottery $[(F_i, q_i)_{i=1}^n]$, becomes a multi-dimensional commodity bundle, and the associated first-stage probability, q_i , becomes the probability of getting that bundle in a one-stage lottery over the commodity space. With this interpretation, the multi-commodity lottery Y is riskier than the multi-commodity lottery X according to a notion of multi-commodity (or many-good) risk based on elementary linear bifurcations. Single-action information loving is analogous to the substitution property we called many-good bifurcation risk loving. Moreover, there is a similar analogy between recursivity and the substitution property we called degenerate independence.²²

Proposition 1 exploits this analogy. 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 recursivity imply that the preferences over late-resolution lotteries are quasi-convex. (Quasi-concavity becomes quasi-convexity because risk aversion becomes information loving.) This is the intuition for part (ii) of the proposition, though we do not need the full strength of recursivity for this result: conditional quasi-convexity is enough. This result concerns intrinsic preference for information and does not depend on the particular form of the preferences over early or late resolution lotteries. Kreps and Porteus [26] and Machina [29], however, show that if an agent has more than one action available and is a recursive expected utility maximizer then the *induced* preferences over lotteries are, again, quasi-convex. Thus, intrinsic and instrumental preference for information both imply quasi-convexity.²³

Parts (iii) and (iv) also use the analogy. Machina [28] exploits the fact that if preferences over one-stage lotteries are sufficiently smooth, even if

²² See Grant, Kajii, and Polak [19] and [18], respectively.

²³ Green [23] and Safra and Sulganik [35] show that quasi-convexity is also necessary for an agent always to prefer more information where she has many actions available, and satisfies reduction but not dynamic consistency.

independence is violated, the preferences can be approximately represented, at least locally, by expected utility functionals. Machina shows that many global comparative-static properties of preferences could be characterized in terms of properties of the local representations. Although Machina assumed Fréchet differentiability, Chew, Karni and Safra [10] and Chew and Nishimura [11] extend the idea to Gateaux differentiability.

We can think of Proposition 1 part (iv) as combining Machina's [28] method with the recursive method discussed in Section 2. Just as Machina's local expected utility functions map from commodities to the reals, so the compound functions $u_{er}(CE_{lr}(\cdot), G)$ map from second-stage lotteries to the reals. Loosely speaking, the inner part of this function, $CE_{lr}(\cdot)$, "replaces" each second-stage lottery with its late-resolution certainty equivalent. The agent then evaluates the resultant early-resolution lottery using her preference relation \succsim_{er} , which (by assumption) is smooth, and so can be represented using the local expected utility functions, $u_{er}(\cdot, G)$ at each G in $\mathcal{L}_0(\mathcal{X})$. 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, an agent always prefers more informative signals if and only if the compound function $u_{er}(CE_{lr}(\cdot), G)$ is convex for all G in $\mathcal{L}_0(\mathcal{X})$; that is, the agent is everywhere "locally information loving." Part (iii) of the proposition shows that this idea can be generalized and does not depend at all on recursivity.

Part (v) of the proposition specializes further to the recursive expected utility model, originally introduced by Kreps and Porteus [25]. The result can be obtained as a corollary of their Theorem 3 or as a corollary of part (iv) here. In the appendix, however, we provide a direct proof. For an intuition, notice that 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 \succsim_{lr} is greater than that for \succsim_{er} . For the case of recursive expected utility, to say that an agent is more risk-averse for late- than for early-resolution lotteries is the same as saying that she dislikes risk more the longer she has to wait for its resolution. Hence, she is willing to pay for information that resolves risk early. That is, she is an intrinsic information lover.

This intuition seems so natural that it suggests the result might extend beyond expected utility. That is, one might conjecture that, given recursivity, regardless of their particular form, preferences over early-resolution lotteries are less risk-averse than those over late-resolution lotteries if and only if the agent satisfies SAIL. This conjecture, however, turns out to be false.

Recall that a signal that is fully revealing induces an early-resolution lottery (so that all risk is resolved in the first stage), and a signal that reveals nothing induces a late-resolution lottery (so that no risk is resolved in the first stage). Therefore, an agent always prefers full information to no

information if and only if $CE_{er}(F) \geq CE_{lr}(F)$ for all simple one-stage lotteries F in $\mathcal{L}_0(\mathcal{X})$. Now, suppose an agent's preferences over late-resolution lotteries are more risk-averse than those over early-resolution lotteries. Then, in particular, the agent would pay more for the complete elimination of a late-resolving risk than for the complete elimination of the corresponding early-resolving risk. That is, for each risk, her late-resolution certainty equivalent must be less than her early-resolution certainty equivalent. Thus, if an agent's 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, in general, this is not sufficient to conclude that the agent prefers any partial increase in information.²⁴ That is, this does not imply SAIL.

Similarly, suppose an agent satisfies SAIL. Then, she prefers more to less informative signals so, in particular, she prefers full to no information. Thus, for each risk, her late-resolution certainty equivalent must be less than her early-resolution certainty equivalent. That is, if an agent satisfies SAIL then, she would pay more for the complete elimination of a risk that resolves late than for that of the corresponding risk that resolves early. But, this is not sufficient to rank the preferences over early- and late-resolution lotteries by their attitudes to all partial decreases in risk. For the special case of expected utility, to adduce whether one preference relation is more risk averse than another with regard to all reductions in risk, partial or complete, it is enough to compare their willingness to pay for the complete elimination of risks, that is to compare certainty equivalents. For general preferences, however, we need also to compare conditional willingness to pay for partial reduction of risk, that is, to compare conditional certainty equivalents.

DEFINITION 9. For any pair of one stage lotteries F and F' in $\mathcal{L}(\mathcal{X})$, and any α in $(0, 1]$ the *conditional certainty equivalent* of F given F' and α with respect to the preference relation \succsim_1 , $CCE_1(F; F', \alpha)$, is the outcome in \mathcal{X} such that $\alpha F + (1 - \alpha) F' \sim_1 \alpha \delta_{CCE_1(F)} + (1 - \alpha) F'$.

The following proposition formalizes (and slightly strengthens) the conclusions of the above discussion. The proof, including counter-examples for the negative statements, is in the appendix.

PROPOSITION 2. *Given recursivity:*

(i) *If an agent's preference relation over two-stage lotteries, \succsim_2 , satisfies SAIL then for all pairs of discrete one-stage lotteries F and F' in*

²⁴ Ahlbrecht and Weber's [2] experimental evidence suggests that a preference for partial increases in information does not always coincide with a preference for complete information.

$\mathcal{L}_0(\mathcal{X})$, and all α in $(0, 1]$, $CE_{lr}(F) \leq CCE_{er}(F; F', \alpha)$. But, this agent's preference relation over late-resolution lotteries, \succsim_{lr} , need not be more risk averse than her preference relation over early-resolution lotteries, \succsim_{er} .

(ii) If an agent's preference relation over late-resolution lotteries, \succsim_{lr} , is more risk averse than her preference relation over early-resolution lotteries, \succsim_{er} , then for each one-stage lottery $(x_i, p_i)_{i=1}^N$ in $\mathcal{L}_0(\mathcal{X})$ the early-resolution lottery $(\delta_{x_i}, p_i)_{i=1}^N$ is weakly preferred to the late-resolution lottery $[(x_i, p_i)_{i=1}^N, 1]$. But this agent's preference relation over two-stage lotteries, \succsim_2 , need not satisfy SAIL.

Earlier studies, such as Wakker [43], have argued that if an agent always prefers more information to less then her preferences must satisfy independence. Safra and Sulganik [35] show that for almost all pairs of information-ranked signals there exists a non-expected utility maximizer who, given some set of actions, strictly prefers the less informative signal. Both theirs and Wakker's assumptions, however, implicitly exclude agents who have strict intrinsic preferences about information. That is, they assume the reduction axiom. Safra and Sulganik leave as an open question whether, if we drop reduction, there exist recursive non-expected utility preferences that always prefer more information to less.

In a companion paper (Grant, Kajii, and Polak [21]), we have shown that both the two most-studied families of non-expected utility preferences (recursive betweenness and recursive rank-dependence) almost collapse back to recursive expected utility if we impose preference for information. The following example, however, shows that, even given recursivity, preference for information does not imply that either the preferences over early- or over late-resolution lotteries need satisfy independence. Indeed, in the example, preferences are strictly quasi-convex in both stages. Moreover, both preferences are risk averse, and both come from the same family of static preferences.²⁵

EXAMPLE 1. Assume recursivity. Let preferences over early-resolution lotteries be represented by $W_{er}(F) = \frac{1}{2}[\int_0^1 x^{1/2}F(dx) + (\int_0^1 x^{1/4}F(dx))^2]$ and let preferences over late-resolution lotteries be represented by $W_{lr}(F) = \frac{1}{2}[\int_0^1 x^{1/4}F(dx) + (\int_0^1 x^{1/8}F(dx))^2]$.

Given Proposition 1 part (iv), to show that these preferences exhibit SAIL, it is enough to show that the local utility function $u_{er}(CE_{lr}(\cdot); G)$ is

²⁵ In itself, this example only deals with the case where the agent has just one action available and so can not condition her actions on what she learns. That is, it deals only with intrinsic, not instrumental, preference for information. Since we have recursivity, however, it is straightforward to extend to the many-action case, see Grant, Kajii, and Polak [22].

convex for all G in $\mathcal{L}(\mathcal{X})$. The local utility function of W_{er} is given by $u_{er}(x; G) = \frac{1}{2}x^{1/2} + [\int_0^1 y^{1/4}G(dy)] \times x^{1/4}$. The certainty equivalent function derived from the preference for late-resolution lotteries, W_{lr} , is given by $CE_{lr}(F) = h_{lr} \circ W_{lr}(F)$, where $h_{lr}(z) = z^4$. Since $W_{lr}(\cdot)$ is convex for probability mixtures, it follows that $CE_{lr}(F)$ is also convex. Hence the compound function $u_{er}(CE_{lr}(\cdot); G)$ from $\mathcal{L}(\mathcal{X})$ to the reals is convex.

The functionals W_{er} and W_{lr} are examples of Chew, Epstein, and Segal's [8] quadratic utility; hence Example 1 is a recursive quadratic utility model. Recall that increasing information is analogous to increasing risk in the space of second-stage lotteries. 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 or quasi-concavity in the probabilities; in particular, preferences can be both risk averse and quasi-concave. It is this flexibility that allows the recursive quadratic utility model to incorporate both intrinsic preference for information and violations of the independence axiom at both stages. We know from Proposition 2, 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.

APPENDIX

An increase in the informational content of a signal can be represented by a sequence of elementary linear bifurcations in the distribution of induced posteriors. This is formalized by the third equivalence statement in the lemma below. The equivalence of the first two statements comes from Blackwell and Girshick [6] Theorem 12.2.2.

LEMMA A.1. *Suppose that $[(P_i, q_i)_{i=1}^N]$ and $[(P'_j, q'_j)_{j=1}^{N'}]$ in $\mathcal{L}_0(\mathcal{L}(\Omega))$ are the distributions of posteriors on Ω induced by the signals (S, λ) and (S', λ') respectively. Suppose that both distributions of posteriors have the same prior, that is, $\sum_i q_i P_i = \sum_j q'_j P'_j = \pi \in \mathcal{L}(\Omega)$. 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 .*

(ii) *There exist weights $\beta_{ij} \geq 0$, $i = 1, \dots, N$, $j = 1, \dots, N'$, 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 states ω in Ω , and all j .*

(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* $k = 1, \dots, K - 1$.

Proof of Lemma A.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 states ω in 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, $\{s \in S \mid \int_{\omega} \lambda(s | \omega) \pi(d\omega) = 0\}$. Similarly define S'_0 . We can then identify $S \setminus S_0$ with $(s_i)_{i=1}^N$, and $S' \setminus S'_0$ with $(s'_j)_{j=1}^{N'}$. Since for π -almost all states ω in 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 $\sum_{s' \in S' \setminus S'_0} \alpha(s', s_0) = 1$, 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 $[(P_i^2, q_i^2)_{i=1}^{N^2}] = [P_0^2, (1 - \beta_{11}) q'_1; P_1, \beta_{11} q'_1; P_2, q'_2; \dots; P_{N'}, q'_{N'}]$ where $(1 - \beta_{11}) P_0^2 + \beta_{11} P_1 = P'_1$. And, let $[(P_i^3, q_i^3)_{i=1}^{N^3}] = [P_0^3, (1 - \beta_{11} - \beta_{21}) q'_1; P_1, \beta_{11} q'_1; P_2, \beta_{21} q'_1; P_2, q'_2; \dots; P_{N'}, q'_{N'}]$ 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 $k = 1, \dots, N$. 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 $k = jN + 1, \dots, (j+1)N$, $j = 2, \dots, N'$ by “splitting” P'_j in a similar manner.

(iii) \Rightarrow (ii) This follows from the definition of an elementary linear bifurcation. ■

The next lemma shows that, by continuity, any pair of two-stage lotteries are ranked by an elementary 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 an elementary linear bifurcation.

LEMMA A.2. *Let* $X = [(F_i, q_i)_{i=1}^N]$ *and* $Y = [(F_1, q_1; \dots; F_{j-1}, q_{j-1}; F'_j, \beta q_j; F''_j, (1 - \beta) q_j; F_{j+1}, q_{j+1}; \dots; F_N, q_N)]$ *in* $\mathcal{L}_0(\mathcal{L}(X))$ *where* $\beta \in [0, 1]$.

Then Y is an elementary 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 an elementary 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 $n = 1, 2, \dots$. 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: $\mu((F_i)^n, F_i) < 1/n$, $\mu((F'_j)^n, F'_j) < 1/n$, and $\mu((F''_j)^n, F''_j) < 1/n$. Set $(F_j)^n = \beta(F'_j)^n + (1 - \beta)(F''_j)^n$, and let $X^n = [((F_i)^n, q_i)_{i=1}^N]$ and $Y^n = [((F_1)^n, q_1; \dots; (F_{j-1})^n, q_{j-1}; (F'_j)^n, \beta q_j; (F''_j)^n, (1 - \beta) q_j; (F_{j+1})^n, q_{j+1}; \dots; (F_N)^n, q_N)]$. By construction, we have $X^n, Y^n \in \mathcal{L}_0(\mathcal{L}_0(\mathcal{X}))$ and Y^n is an elementary linear bifurcation of X^n , and by the linearity of the topology, $\mu((F_i)^n, F_i) = \mu(\beta(F'_j)^n + (1 - \beta)(F''_j)^n, \beta F'_j + (1 - \beta) F''_j) < K/n$ for some constant K . Thus as $n \rightarrow \infty$, $X^n \rightarrow X$ and $Y^n \rightarrow Y$ in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$. The converse follows from the continuity of linear operation. ■

Proof of Proposition 1. (i). (A) \Rightarrow (B) Suppose SAIL does not hold. Then, by Lemma A.2 and the continuity of the preference relation \succcurlyeq_2 , there exists a pair of two-stage lotteries both discrete in the second stage, $X = [(F_i, q_i)_{i=1}^N]$ and $Y = [(F_1, q_1; \dots; F_{j-1}, q_{j-1}; F'_j, \beta q_j; F''_j, (1 - \beta) q_j; F_{j+1}, q_{j+1}; \dots; F_N, q_N)]$ 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 $X \succ_2 Y$. Choose c in \mathcal{C} and a in \mathcal{A} such that $c(\cdot, a)$ is a bijection from Ω to \mathcal{X} and choose $(P_i)_{i=1}^N$, P'_j and P''_j , such that $(H(P_i, c, a))_{i=1}^N = (F_i)_{i=1}^N$, $H(P'_j, c, a) = F'_j$ and $H(P''_j, c, a) = F''_j$. 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 A.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 two-stage lottery induced by (S', λ') given prior π , single action a and consequence function c , is strictly preferred to that induced by (S, λ) .

(B) \Rightarrow (A) By Lemma A.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 \succcurlyeq_2 .

(ii) For all pairs of discrete one-stage lotteries F, F' in $\mathcal{L}(\mathcal{X})$ such that $[F, 1] \succcurlyeq_{lr} [F', 1]$, CQV implies that $[F, 1] \succcurlyeq_2 [F, (1 - \alpha); F', \alpha]$ for all α in $(0, 1)$. SAIL implies $[F, (1 - \alpha); F', \alpha] \succcurlyeq_2 [(1 - \alpha) F + \alpha F', 1]$. Hence, $[F, 1] \succcurlyeq_{lr} [(1 - \alpha) F + \alpha F', 1]$.

(iii) The idea of the proof is similar to Machina [28]. Suppose that \succcurlyeq_2 satisfies SAIL but that $v(\cdot; X)$ is not convex for some X in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$. Then, we can find F, F' , and F'' in $\mathcal{L}(\mathcal{X})$ and α in $(0, 1)$ such that

$F = \alpha F' + (1 - \alpha) F''$ but $v(F; X) > \alpha v(F'; X) + (1 - \alpha) v(F''; X)$. Let $Y := \alpha \delta_{F'} + (1 - \alpha) \delta_{F''} \in \mathcal{L}_0(\mathcal{L}(\mathcal{X}))$; that is, Y is the two-stage lottery that gives F' with probability α and F'' with probability $(1 - \alpha)$. For each ε in $(0, 1)$, let $Z_0(\varepsilon)$ and $Z_1(\varepsilon)$ be the elements of $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$ given by $Z_0(\varepsilon) := \varepsilon \delta_F + (1 - \varepsilon) X$ and $Z_1(\varepsilon) := \varepsilon Y + (1 - \varepsilon) X$ respectively. By construction, $Z_1(\varepsilon)$ is an elementary linear bifurcation of $Z_0(\varepsilon)$, so SAIL implies $W(Z_0(\varepsilon)) - W(Z_1(\varepsilon)) \leq 0$ for each ε in $(0, 1)$.

Using the Gateaux differentiability of W at X , however, $W(Z_0(\varepsilon)) - W(X) = W(\varepsilon \delta_F + (1 - \varepsilon) X) - W(X) = \int_{\mathcal{L}(\mathcal{X})} v(\mu; X) [\varepsilon \delta_F(d\mu) - \varepsilon X(d\mu)] + o_1(\varepsilon)$, and similarly, $W(Z_1(\varepsilon)) - W(X) = \int_{\mathcal{L}(\mathcal{X})} v(\mu; X) [\varepsilon Y(d\mu) - \varepsilon X(d\mu)] + o_2(\varepsilon)$ where $o_1(\varepsilon)$ and $o_2(\varepsilon)$ satisfy $o_1(\varepsilon)/\varepsilon \rightarrow 0$ and $o_2(\varepsilon)/\varepsilon \rightarrow 0$ as $\varepsilon \rightarrow +0$. Combining these two expressions, we get $W(Z_0(\varepsilon)) - W(Z_1(\varepsilon)) = \int_{\mathcal{L}(\mathcal{X})} v(\mu; X) [\varepsilon \delta_F(d\mu) - \varepsilon Y(d\mu)] + o_1(\varepsilon) - o_2(\varepsilon) = \varepsilon [v(F, X) - (\alpha v(F', X) + (1 - \alpha) v(F'', X))] + o_1(\varepsilon) - o_2(\varepsilon)$. Hence, $1/\varepsilon [W(Z_0(\varepsilon)) - W(Z_1(\varepsilon))] > 0$ for small enough ε : a contradiction.

Conversely, assume $v(\cdot; X)$ is convex at any X in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$. Fix a two-stage lottery X in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$, weights p and q both in $(0, 1)$, and two one-stage lotteries F' and F'' in $\mathcal{L}(\mathcal{X})$ and let $F := pF' + (1 - p) F''$. Let Z_0 and Z_1 be the elements of $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$ given by $Z_0 := q\delta_F + (1 - q) X$ and $Z_1 := q(p\delta_{F'} + (1 - p) \delta_{F''}) + (1 - q) X$. By construction, Z_1 is an elementary linear bifurcation of Z_0 . Since our choice of X , F' , F'' , p and q was arbitrary, it is enough to show that $W(Z_1) - W(Z_0) \geq 0$.

Let $Z(\alpha) := \alpha Z_1 + (1 - \alpha) Z_0$. We first show that $\alpha \mapsto W(Z(\alpha))$ is a real valued differentiable function on $(0, 1)$; that is, $(d/d\alpha) W(Z(\alpha))|_{\alpha=\beta}$ exists for any β in $(0, 1)$. By the definition of Gateaux differentiability, for any β in $(0, 1)$ and for any γ in $(0, 1)$, $W(\gamma Z_1 + (1 - \gamma) Z(\beta)) - W(Z(\beta)) = \gamma \int_{\mathcal{L}(\mathcal{X})} v(\mu; Z(\beta)) (Z_1(d\mu) - Z(\beta)(d\mu)) + o(\gamma)$. Set $\alpha = \gamma + (1 - \gamma) \beta$, so that α approaches β as γ approaches zero and $Z(\alpha) = \gamma Z_1 + (1 - \gamma) Z(\beta)$. By Gateaux differentiability, $(d/d\alpha) W(Z(\alpha))|_{\alpha=\beta} = \lim_{\alpha \rightarrow \beta} 1/(\alpha - \beta) [W(Z(\alpha)) - W(Z(\beta))] = \lim_{\gamma \rightarrow 0} 1/(\gamma(1 - \beta)) [W(\gamma Z_1 + (1 - \gamma) Z(\beta)) - W(Z(\beta))] = 1/(1 - \beta) \int_{\mathcal{L}(\mathcal{X})} v(\mu; Z(\beta)) (Z_1(d\mu) - Z(\beta)(d\mu))$. Substituting for $Z(\beta)$, this derivative becomes $\int_{\mathcal{L}(\mathcal{X})} v(\mu; Z(\beta)) (Z_1(d\mu) - Z_0(d\mu))$.

Substituting for Z_1 and Z_0 , then subtracting, yields $q \int_{\mathcal{L}(\mathcal{X})} v(\mu, Z(\beta)) [(p\delta_{F'} + (1 - p) \delta_{F''}) - \delta_F](d\mu) = q[pv(F', Z(\beta)) + (1 - p) v(F'', Z(\beta)) - v(F, Z(\beta))] \geq 0$, where the last inequality is by the convexity of $v(\cdot, Z(\beta))$. Thus we have $(d/d\alpha) W(Z(\alpha))|_{\alpha=\beta} \geq 0$ for any β in $(0, 1)$. Integrating with respect to β yields $\int_0^1 [(d/d\alpha) W(Z(\alpha))|_{\alpha=\beta}] d\beta = W(Z(1)) - W(Z(0)) \geq 0$ by the fundamental theorem of calculus. We are done since $W(Z(1)) = W(Z_1)$ and $W(Z(0)) = W(Z_0)$.

(iv) Let $V_{er}: \mathcal{L}_0(\mathcal{X}) \rightarrow \mathbb{R}$ be the Gateaux differentiable representation of the agent's preferences over early-resolution lotteries with Gateaux derivative $u_{er}(\cdot; \cdot): \mathcal{X} \times \mathcal{L}_0(\mathcal{X}) \rightarrow \mathbb{R}$. For any two-stage lottery in $X = [(F_i, q_i)_{i=1}^N]$

in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$, we denote by G_X the image measure $X \circ \text{CE}_{lr}^{-1}$ of lottery X via the function $\text{CE}_{lr}(\cdot)$. That is, G_X is the element of $\mathcal{L}_0(\mathcal{X})$ given by $[(\text{CE}_{lr}(F_i), q_i)_{i=1}^N]$, which is the one-stage lottery such that, for each Borel subset B of \mathcal{X} , $G_X(B) = X(\{F \in \mathcal{L}(\mathcal{X}) : \text{CE}_{lr}(F_i) \in B\})$. So from the standard theory of integration, we can use the change of variables formula: $\int_{\mathcal{X}} f(x) G_X(dx) = \int_{\mathcal{L}(\mathcal{X})} f(\text{CE}_{lr}(\mu)) X(d\mu)$ for any integrable function f on \mathcal{X} .

Since the agent satisfies recursivity, there exists a representation of the agent's preferences over two-stage lotteries $W: \mathcal{L}_0(\mathcal{L}(\mathcal{X})) \rightarrow \mathbb{R}$ such that, for all X in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$, $W(X) = V_{er}(G_X)$. Given part (iii) of the proposition, it is enough to show that this representation W is Gateaux differentiable with Gateaux derivative $v: \mathcal{L}(\mathcal{X}) \times \mathcal{L}_0(\mathcal{L}(\mathcal{X})) \rightarrow \mathbb{R}$ given by $v(F, X) := u_{er}(\text{CE}_{lr}(F); G_X)$ for each F in $\mathcal{L}(\mathcal{X})$ and each X in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$. To see this, fix any α in $(0, 1)$ and any pair of two-stage lotteries, X and Y in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$, and let G_X and G_Y be the corresponding one-stage lotteries as defined above. By recursivity, we have $W(\alpha Y + (1 - \alpha) X) - W(X) = V_{er}(\alpha G_Y + (1 - \alpha) G_X) - V_{er}(G_X)$. By the Gateaux differentiability of V_{er} , this is equal to $\int_{\mathcal{X}} u_{er}(x, G_X)(\alpha G_Y(dx) - \alpha G_X(dx)) + o(\alpha) = \int_{\mathcal{L}(\mathcal{X})} u_{er}(\text{CE}_{lr}(u), G_X)(\alpha Y(d\mu) - \alpha X(d\mu)) + o(\alpha)$, where the last equality is by a change of variable. But the last expression is indeed the definition of Gateaux differentiability.

(v) Let the functional that represents \succsim_{lr} be given by let $U_{lr}(F) = \int_0^1 u_{lr}(x) F(dx)$, for all F in $\mathcal{L}(\mathcal{X})$. Notice that, for all F in $\mathcal{L}(\mathcal{X})$, $\text{CE}_{lr}(F) = u_{lr}^{-1}(U_{lr}(F))$. Let $X = [(F_i, q_i)_{i=1}^N]$ and $Y = [(F_1, q_1; \dots; F_{j-1}, q_{j-1}; F'_j, \beta q_j; F''_j, (1 - \beta) q_j; F_{j+1}, q_{j+1}; \dots; F_N, q_N)]$ in $\mathcal{L}_0(\mathcal{L}(\mathcal{X}))$, where $q_j > 0$ and $F_j = \beta F'_j + (1 - \beta) F''_j$. Since \succsim_{lr} satisfies independence, $\beta U_{lr}(F'_j) + (1 - \beta) U_{lr}(F''_j) = U_{lr}(F_j)$. Thus, given recursivity and the fact that \succsim_{er} satisfies independence, $X \succsim_2 Y$ 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. ■

Proof of Proposition 2. (i) Fix any $F = [(x_i, p_i)_{i=1}^N]$, $F' = [(x'_i, p'_i)_{i=1}^{N'}]$ in $\mathcal{L}_0(\mathcal{X})$, and any α in $(0, 1]$. By the definition of $\text{CCE}_{er}(F; F', \alpha)$, $[\delta_{\text{CCE}_{er}(F; F', \alpha)}, \alpha; \delta_{x'_1}, (1 - \alpha) p'_1; \dots; \delta_{x'_{N'}}, (1 - \alpha) p'_{N'}] \sim_{er} [\delta_{x_1}, \alpha p_1; \dots; \delta_{x_N}, \alpha p_N; \delta_{x'_1}, (1 - \alpha) p'_1; \dots; \delta_{x'_{N'}}, (1 - \alpha) p'_{N'}]$. By SAIL, $[\delta_{x_1}, \alpha p_1; \dots; \delta_{x_N}, \alpha p_N; \delta_{x'_1}, (1 - \alpha) p'_1; \dots; \delta_{x'_{N'}}, (1 - \alpha) p'_{N'}] \succsim_2 [F, \alpha; \delta_{x'_1}, (1 - \alpha) p'_1; \dots; \delta_{x'_{N'}}, (1 - \alpha) p'_{N'}]$. And by recursivity and the definition of $\text{CE}_{lr}(F)$, $[F, \alpha; \delta_{x'_1}, (1 - \alpha) p'_1; \dots; \delta_{x'_{N'}}, (1 - \alpha) p'_{N'}] \sim_2 [\delta_{\text{CE}_{lr}(F)}, \alpha; \delta_{x'_1}, (1 - \alpha) p'_1; \dots; \delta_{x'_{N'}}, (1 - \alpha) p'_{N'}]$. Since the early-resolution preferences respect first-order stochastic dominance, the claim follows.

It remains to show that SAIL does not imply that preferences over late-resolution lotteries are more risk averse than those over early-resolution

lotteries. Given recursivity, we can characterize a preference relation over two-stage lotteries, \succcurlyeq_2 , by the representations of its restriction to early- and late resolution lotteries, denoted V_{er} and V_{lr} respectively.²⁶ With slight abuse of notation, let F represent both the cumulative distribution function of a lottery in $\mathcal{L}(\mathcal{X})$ and the one-stage lottery itself.

COUNTEREXAMPLE 1. Consider the recursive preferences over two-stage lotteries where for any F in $\mathcal{L}(\mathcal{X})$, $V_{er}(F) = \int_0^1 x^{1/2} dF(x)$ and $V_{lr}(F) = \int_0^1 x d(f \circ F(x))$ where f is the increasing function given by $f(p) = 1 - (1 - p)^2$ for all p in $[0, 1]$.

These preferences over early-resolution lotteries satisfy the expected utility theory while the preferences over late-resolution lotteries satisfy Yaari's [46] dual theory.

We first show that these preferences over two-stage lotteries satisfy SAIL. Given Proposition 1(iv), it is enough to show that, for all lotteries G in $\mathcal{L}(\mathcal{X})$, the compound function $u_{er}(\text{CE}_{lr}(\cdot); G): \mathcal{L}(\mathcal{X}) \rightarrow \mathbb{R}$ is convex. Since the preferences over early resolution lotteries are just expected utility, however, for all G in $\mathcal{L}(\mathcal{X})$, $u_{er}(\text{CE}_{lr}(\cdot); G) = (\text{CE}_{lr}(\cdot))^{1/2} = (\int_0^1 x d(f \circ F(x)))^{1/2} = (\int_0^1 [1 - F(x)]^2 dx)^{1/2}$ where the second equality is derived using integration by parts. But this last expression is convex in F by Minkowski's inequality.

We need to show that these preferences over late-resolution lotteries are not more risk averse than the preferences over early-resolution lotteries; that is, there exists a pair of lotteries F, G in $\mathcal{L}(\mathcal{X})$ and an α in $(0, 1]$ such that $\text{CCE}_{lr}(F; G, \alpha) > \text{CCE}_{er}(F; G, \alpha)$. Let $F = G = [0, \frac{1}{2}; 1, \frac{1}{2}]$. Then a simple calculation shows that, for all $\alpha < 1$, $\text{CCE}_{lr}(F; F, \alpha) = [(1 + \alpha)^2 + (1 - \alpha)^2]^{-1} > [2^2]^{-1} = \text{CCE}_{er}(F; F, \alpha)$.

(ii) Fix a lottery $F = [(x_i, p_i)_{i=1}^N]$ in $\mathcal{L}_0(\mathcal{X})$. By the definition of certainty equivalent, for the late resolution lottery we have $[F, 1] \sim_{lr} [\delta_{\text{CE}_{lr}(F)}, 1]$, and for the early resolution-lottery $[(\delta_{x_i}, p_i)_{i=1}^N] \sim_{er} [\delta_{\text{CE}_{er}(F)}, 1]$. Since the agent's preferences over late-resolution lotteries are more risk that those over early resolution lotteries, $\text{CE}_{lr}(F) \leq \text{CE}_{er}(F)$. The claim follows from first-order stochastic dominance.

It remains to show that preferences over late-resolution lotteries being more risk averse than those over early-resolution lotteries does not imply SAIL. As before, given recursivity, we can characterize a preference relation over two-stage lotteries, \succcurlyeq_2 , by the representations of its restriction to

²⁶ Strictly speaking, preferences over early-resolution lotteries, \succcurlyeq_{er} , are only defined on $\mathcal{L}_0(\mathcal{X})$. For ease of notation, however, let V_{er} denote the extension of a representation to $\mathcal{L}(\mathcal{X})$. That is, we say that $V_{er}: \mathcal{L}(\mathcal{X}) \rightarrow \mathbb{R}$ represents \succcurlyeq_{er} if for all F, G in $\mathcal{L}_0(\mathcal{X})$, $V_{er}(F) \geq V_{er}(G)$ if and only if $F \succcurlyeq_{er} G$.

early- and late resolution lotteries, denoted V_{er} and V_{lr} respectively. With slight abuse of notation, let F represent both the cumulative distribution function of a lottery in $\mathcal{L}(\mathcal{X})$ and the one-stage lottery itself.

COUNTEREXAMPLE 2. Consider the recursive preferences over two-stage lotteries where for any F in $\mathcal{L}(\mathcal{X})$, $V_{er}(F) = V_{lr}(F) = \int_0^1 x d(f \circ F(x))$ where $f(p) = p^2$.

Since the preferences over early- and late-resolution lotteries are the same, trivially, the latter is weakly more risk averse than the former.²⁷ Thus, we only need to show that these preferences do not satisfy SAIL. Since the preferences are recursive (and hence satisfy conditional quasi-convexity), given Proposition 1 part (ii), it is enough to show that the preferences

over late-resolution lotteries violate quasi-convexity. Consider the three one-stage lotteries, $F = [0, \frac{1}{2}; 1, \frac{1}{2}]$, $G = [\frac{3}{4}, 1]$ and $\frac{1}{2}F + \frac{1}{2}G$. A simple calculation shows that $V_{lr}(F) = V_{lr}(G) = \frac{3}{4} < \frac{13}{16} = V_{lr}(\frac{1}{2}F + \frac{1}{2}G)$. ■

REFERENCES

1. M. Ahlbrecht and M. Weber, The resolution of uncertainty: An experimental study, *J. Inst. Theor. Econ.* **152** (1996), 593–607.
2. M. Ahlbrecht and M. Weber, Preference for the gradual resolution of uncertainty, *Theory and Decision* **43** (1997), 167–185.
3. H. Bohnenblust, L. Shapley, and S. Sherman, Reconnaissance in game theory, working paper rm-208, The RAND Corporation, 1949.
4. D. Blackwell, The composition of experiments, in “Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability,” Univ. of California Press, Berkeley, 1951.
5. D. Blackwell, Equivalent comparison of experiments, *Ann. Math. Statist.* **24** (1953), 265–272.
6. D. Blackwell and M. Girshick, “Theory of Games and Statistical Decisions,” Wiley, New York, 1954.
7. S. H. Chew and L. Epstein, The structure of preferences and attitudes towards the timing of the resolution of uncertainty, *Int. Econ. Rev.* **30** (1989), 103–117.
8. S. H. Chew, L. Epstein, and U. Segal, Mixture symmetry and quadratic utility, *Econometrica* **59** (1991), 139–163.
9. S. H. Chew and J. Ho, Hope: An empirical study of attitude toward the timing of uncertainty resolution, *J. Risk Uncertainty* **8** (1994), 267–288.
10. S. H. Chew, E. Karni, and Z. Safra, Risk aversion in the theory of expected utility with rank-dependent preferences, *J. Econ. Theory* **42** (1987), 370–381.

²⁷ In this example, both preferences are risk loving. Nothing, however, depends on this or on the preferences being the same. Grant, Kajii, and Polak [20] give necessary and sufficient conditions for information loving in the general recursive rank dependent model (of which the dual model is a special case), allowing the construction of a more general example.

11. S. H. Chew and N. Nishimura, Differentiability, comparative statics and non-expected utility preferences, *J. Econ. Theory* **56** (1992), 294–312.
12. V. Cook, "The Effects of Temporal Uncertainty Resolution on the Overall Utility and Suspense of Risky Monetary and Survival Gambles," unpublished Ph.D. thesis, McGill University, 1989.
13. L. Epstein, Behavior under risk: Recent developments in theory and applications, in "Advances in Economic Theory, Sixth World Congress. Vol. 2" (Jean-Jacques Laffont, Ed.), pp. 1–63, Cambridge Univ. Press, Cambridge, UK, 1992.
14. L. Epstein and S. Zin, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework, *Econometrica* **57** (1989), 937–969.
15. L. Epstein and S. Zin, Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis, *J. Polit. Econ.* **99** (1991), 263–286.
16. M. Franklin and R. Rabin, "Tort Law and Alternatives," 6th ed., Foundation Press, New York, 1996.
17. F. Gale and J. Goyer, Recovery for cancerphobia and increased risk of cancer, *Defense Law J.* **35** (1986), 443–464.
18. S. Grant, A. Kajii, and B. Polak, Many good choice axioms: When can many be treated as one? *J. Econ. Theory* **56** (1992), 313–337.
19. S. Grant, A. Kajii, and B. Polak, Many good risks: An interpretation of multivariate risk and risk aversion without the independence axiom, *J. Econ. Theory* **56** (1992), 338–351.
20. S. Grant, A. Kajii, and B. Polak, Preference for information, Australian National University, Department of Economics working paper no. 298, 1996.
21. S. Grant, A. Kajii, and B. Polak, Temporal resolution of uncertainty and recursive non-expected utility models, Australian National University, Department of Economics working paper no. 324, 1997.
22. S. Grant, A. Kajii, and B. Polak, Dynamic consistency and preference for information, Australian National University, Department of Economics working paper no. 326, 1997.
23. J. Green, Making book against oneself: The independence axiom and non-linear utility theory, *Quart. J. Econ.* **102** (1987), 785–796.
24. D. Kelsey, David, comment on "nonexpected utility as aversion of information," *J. Behav. Decision Making* **2** (1989), 199–200.
25. D. Kreps and E. Porteus, Temporal resolution of uncertainty and dynamic choice theory, *Econometrica* **46** (1978), 185–200.
26. D. Kreps and E. Porteus, Temporal von Neumann–Morgenstern and induced preferences, *J. Econ. Theory* **20** (1979), 81–109.
27. I. LaValle and K. Wapman, Rolling back decision trees requires the independence axiom! *Manage. Sci.* **32** (1985), 382–385.
28. M. Machina, Expected utility analysis without the independence axiom, *Econometrica* **50** (1982), 277–323.
29. M. Machina, Temporal risk and the nature of induced preferences, *J. Econ. Theory* **33** (1984), 199–231.
30. M. Machina, Choice under uncertainty: Problems solved and unsolved, *J. Econ. Perspectives* **1** (1987), 121–154.
31. M. Machina, Dynamic consistency and non-expected utility models of choice under uncertainty, *J. Econ. Lit.* **27** (1989), 1622–1668.
32. J. Marschak and R. Radner, "Economic Theory of Teams," Yale Univ. Press, New Haven, 1972.
33. R. Pope, The pre-outcome period and the utility of gambling, in "Foundations of Utility and Risk Theory with Applications" (B. Stigum and F. Wenstøp, Eds.), pp. 137–177, Reidel, Dordrecht, 1983.

34. R. Pope, Timing contradictions in von Neumann and Morgenstern's axioms and in Savage's "surething" proof, *Theory and Decision* **18** (1985), 229–261.
35. Z. Safra and E. Sulganik, On the nonexistence of Blackwell's theorem-type results with general preference relations, *J. Risk Uncertainty* **10** (1995), 187–201.
36. E. Schlee, The value of information in anticipated utility theory, *J. Risk Uncertainty* **5** (1990), 83–92.
37. E. Schlee, The value of perfect information in nonlinear utility theory, *Theory and Decision* **30** (1991), 127–131.
38. E. Schlee, The sure thing principle and the value of information, mimeo, Arizona State University, 1995.
39. U. Segal, Two-stage lotteries without the reduction axiom, *Econometrica* **58** (1990), 349–377.
40. U. Segal, The independence axiom versus the reduction axiom: Must we have both? in "Utility Theories: Measurements and Applications" (W. Ewards, Ed.), Kluwer Academic, Dordrecht, 1992.
41. C. Siebert, Living with toxic knowledge: The DNA we've been dealt, *New York Times Magazine* Sept. 17, 1995.
42. C. Skiadas, Recursive utility and preferences for information, working paper no. 196, Department of Finance, Kellogg Graduate School of Management, Northwestern University, 1997.
43. P. Wakker, Nonexpected utility as aversion of information, *J. Behav. Decision Making* **1** (1988), 169–175.
44. J. Waldfogel, The deadweight loss of christmas, *Amer. Econ. Rev.* **83** (1993), 1328–1336.
45. T. Wang, L_p -Fréchet differentiable preference and "local utility" analysis, *J. Econ. Theory* **61** (1993), 139–159.
46. M. Yaari, The dual theory of choice under risk, *Econometrica* **55** (1987), 95–115.