Inter-temporal Preference for Flexibility and Risky Choice*

Alan Kraus†

Jacob S. Sagi‡

First Version: November 1999

This Version November 2002

Abstract

We derive an inter-temporal theory, in the spirit of Kreps and Porteus (1978), of changing tastes and unforeseen contingencies from normative primitives by weakening the completeness axiom of von Neumann and Morgenstern. Our agent can only partially order *future* decisions, yet desires inter-temporal consistency. Our formulation contrasts with existing literature in several ways: (i) the theory is inherently inter-temporal; (ii) a time consistency condition and therefore the *normative* connection between ex-ante and ex-post choice forces a 'utility for flexibility'; (iii) we deduce conditions under which a subjective state space for unforeseen contingencies is *topologically unique* and derive its existence from preference primitives as opposed to the representation; Finally, our theory reduces to standard recursive utility when the agent can completely order future decisions.

Keywords: Inter-temporal Choice, Flexibility, Non-Expected Utility, Decision Theory, Incomplete Preference Relations, Utility Representation, Partial Orders, Time Consistency, Endogenous State Space

Journal of Economic Literature **Classification number:** D11, D81, D91.

*Acknowledgments: later

†Faculty of Commerce, University of British Columbia, 2053 Main Mall, Vancouver, B.C., Canada V6T 1Z2. email: kraus@commerce.ubc.ca, Phone: (604)-822-8342

‡Haas School of Business, University of California at Berkeley, 545 Student Services Building, Berkeley, California, 94720-1900. email: sagi@haas.berkeley.edu, phone: (510)-642-3442

1 Introduction

In a seminal paper, Kreps (1979) showed that a preference for flexibility implies that an agent acts as if she possesses an endogenous state space. For example, if a menu from which the agent will later consume, $\{a, b\}$, is strictly preferred to both the menu $\{a\}$ *and* the menu $\{b\}$, Kreps demonstrates, under parsimonious assumptions on preference over choice sets, that the agent has a utility representation that suggests an endogenous state space of possible future tastes. Specifically, Kreps derives a representation for preference over subsets of a finite space of prospects, X, with the following structure:

$$x \succ y \Leftrightarrow \sum_{s \in S} \max_{d \in x} U(s, d) \; > \sum_{s \in S} \max_{d \in y} U(s, d)$$

x and y are menus (subsets of X), S is an index set derived endogenously, and for each $s \in S$, $U(s, d)$ is a real-valued function over X. The existence of the index set, S (i.e., the set of utility functions), is interpreted as an endogenous state space of tastes. This space, unfortunately, is not unique and the formulation is not normative. In particular, the theory does not rule out the possibility that the agent simply prefers choice sets with higher cardinality regardless of their content, thus no connection is made to eventual choice from menus or to future preference over the constituents of menus. Moreover, there are many equivalent utility representations that specify different index sets. These, in general, take the form:

$$U(x) = u\big(\max_{d \in x} U(1, d),\ \max_{d \in x} U(2, d),\ \ldots\ \max_{d \in x} U(S, d)\big)$$

where u is increasing in all its arguments (it can be viewed as an aggregator of future utilities). Dekel, Lipman and Rustichini (2001) show that the state space can be essentially pinned down by considering menus containing lotteries over X and insisting that an agent is indifferent between a menu and its convex hull. The latter assumption leads to aggregated $U(s, d)$'s (or ex-post utilities) that are expected utility functionals. Although alternative representations exist, a state space of future expected utility tastes has a weakly smaller cardinality than any other representation whenever $U(x)$ is monotonically increasing in the expected utility functionals.

Both Kreps (1979, 1992) and Dekel, Lipman and Rustichini (2001) describe a setting where choice is essentially static. The agent chooses among menus, uncertainty over her subjective states is assumed to resolve and then the agent selects from the menu. However, there is no explicit modeling of ex-post choice and no role for consistency between realized tastes and tastes inferred from ex-ante preferences - the theories are essentially static.[1] One reason that an inter-temporal theory of unforeseen contingencies is viewed as 'problematic' is due to the interpretation of the rankings that appear in the representation (i.e., the $U(s,d)$'s) as endogenous states: in an inter-temporal setting one would like to establish that the rankings are ex-post justified, or at the very least to impose a connection between the states and ex-post choice. Ideally, one would like to establish that the agent's ex-post ranking is actually one of the $U(s,d)$'s. A 'simple' way to achieve this is to explicitly impose an axiom stating that the agent's ex-post ranking is one of the utility functions that appears in the ex-ante representation. This, however, is completely unsatisfactory: axioms must be imposed directly on agents' choice behavior and not on the mathematical representation of their behavior. Aside from the tautological approach, it is not easy to conceive of a way to directly tie the $U(s,d)$'s to realized ex-post rankings. The goal of this paper is to study and address this problem.

To understand how we confront these issues, consider the top decision tree in Figure 1. Squares (circles) denote a decision (nature) node, the $z_{i,j}$'s correspond to date i payoffs in state j, while the $x_{i,j}$'s to date i menus in state j. Assume that at date 0 the agent agrees to commit to the lottery awarding $z_{2,9}$ or $z_{2,10}$ in the (future) choice node $x_{1,2}$, but refuses to commit to any one of the top two lotteries coming out of $x_{1,1}$. One can represent this situation with the bottom tree of Figure 1: the agent 'eliminates' future dominated choice branches, and turns choice nodes in which flexiblity is valued into *endogenous state* nodes (bold circle around $x_{1,1}$). Branches coming out of the bold circle represent unforeseen contingencies which may impact the agent's tastes. Alternatively, the same branches can be interpreted as the *Pareto frontier* of the set $x_{1,1}$ generated by the possible tastes at date 1. Time consistency requires that the agent, at date 1, select something from the Pareto

[1]By contrast, a relationship between ex-ante and ex-post is important in Gul and Pessendorfer (2001). Their goal, however, is to model self-control problems while we only focus on unforeseen contingencies and preference for flexibility. Consequently their representation includes the negative weighting of some future tastes while ours is almost everywhere strictly positive.

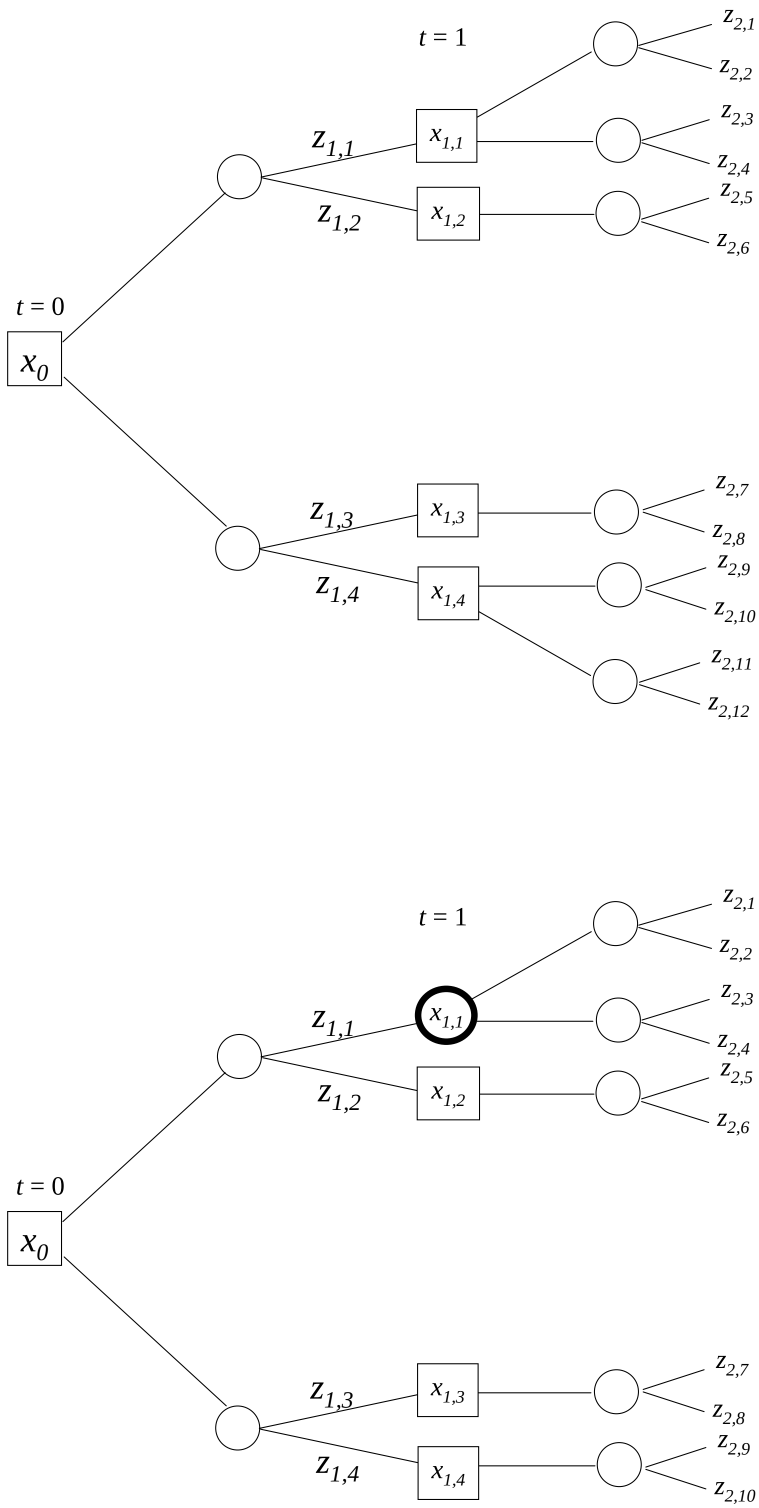

Figure 1: Turning a choice node into an endogenous-states node.

frontier inferred by the date 0 preferences (i.e., she should not ex-post select something that she was willing to give up ex-ante). Note two important observations:

1. The presence of endogenous states and their number depends on date 0 preferences and not on the date 0 representation of preferences.

2. To achieve consistency with the inferred states, it is not necessary to explcitly specify ex-post rankings. Instead, it is sufficient to demand that whatever the realized ex-post ranking, it must be consistent with the Pareto frontier (i.e., branches coming out of the converted node) inferred earlier.

Part of the novelty of our analysis consists of a generalization of the above intuition to a *representation-free* definition of an endogenous state space. Our definition guarantees uniqueness (though not always existence) and is homeomorphic to the endogenous state space of Dekel, Lipman and Rustichini (2001) under fairly general conditions. Moreover, by abstracting away from a representation, we avoid having to associate endogenous states with ex-post rankings. Instead, states are directly associated with elements of ex-ante Pareto frontiers - elements of a menu, up to indifference, to which the agent assigns ex-ante flexibility value. Time consistency and a preference for flexibility is achieved by demanding that ex-post choice must be from the ex-ante Pareto frontier. Thus we do not require realized ex-post rankings to coincide with any one ranking appearing in the representation of ex-ante choice.

Our main arguments and contributions are as follows:

i) The assumptions we make *weaken* those of the standard recursive utility theory of Kreps and Porteus (1978). We start by postulating that the agent's realized date-t ranking of temporal trees (such as the tree in Figure 1) is a completion of some partial ordering, $\succeq_t^P$; the sequence of partial orderings, indexed by the date, limits the set of possible revealed rankings. Each of the partial orderings obeys a transitivity, continuity and independence axiom, thus our point of departure from Kreps and Porteus (1978) consists of, first and foremost, relaxing their completeness axiom. The independence axiom implies that each $\succeq_t^P$ is the intersection of a set of von Neumann-Morgenstern (expected utility) preferences. We denote the associated set of date-t von Neumann-

Morgenstern utility functions as Ψ_t. Revealed preference must be some function of expected utility functions from Ψ_t, and the ex-ante Pareto frontier of a date t menu is generated by the closed convex hull of Ψ_t.

ii) By imposing a time consistency condition that ties ex-post choice to the ex-ante Pareto frontier, we show that each of the von Neumann-Morgenstern functions in Ψ_t exhibits a preference for flexibility analogous to the ordinal EU form derived in Dekel, Lipman and Rustichini (2001).[2] Specifically, any element of Ψ_t positively aggregates the optimal expected utilities from the closed convex hull of the set Ψ_{t+1}. This property is passed down to revealed preferences.

iii) We demonstrate that the sequence of partial orderings, $\{\succeq_s^P\}_{s>t}$, can be inferred from revealed preferences at date t. This is a consequence of time consistency and ultimately justifies the imposition of behavioral axioms on $\{\succeq_s^P\}_{s>t}$.

iv) We formalize the notion of 'transforming choice nodes into endogenous state nodes' illustrated earlier and in Figure 1. Roughly, we define a topological endogenous state space as the set of indifference classes that cannot be deleted from a menu with 'maximal' flexibility without causing welfare loss. In contrast with the literature on preference for flexibility, such a state space derives from preference primitives and not the utility representation used. More importantly, representation independence allows us to sensibly refer to 'topological uniqueness'. We provide conditions under which it is meaningful to view the closed convex hull of Ψ_t as a *topologically unique* endogenous state space. This refines the uniqueness results of Dekel, Lipman and Rustichini (2001).

The benefit or utility that a decision-maker may derive from having flexibility is the subject of research for the social choice and welfare literature as well as that on unforeseen contingencies. Notable references not already mentioned include Bossert, Pattanaik and Xu (1994), Pattanaik and Xu (1998), Puppe (1995, 1996), Nehring and Puppe (1996, 1999), Bossert (1997), Nehring (1999), Ozdenoren (2002) and Al-Najjar, Ozdenoren and

[2]As in Dekel, Lipman and Rustichini (2001), we also require that the agent is free to use mixing strategies when choosing from a choice set.

Casadesus-Masanell (2001).[3] Puppe (1995), in particular, notes that preference for flexibility is intimately related to discontinuous and/or partial orderings over singleton choice sets. If one wishes to derive a preference ordering over choice sets that exhibits a preference for flexibility from a more primitive ordering on individual prospects, then the inducing ordering must be either discontinuous or incomplete. If one wishes to retain continuity, the implication is that a normative theory of changing tastes must arise from primitives that partially order the set of future prospects. This last point forms the basis of our theory.

Preference for flexibility, although implicitly appealing to inter-temporal considerations, has been divorced from the literature on inter-temporal choice. In the latter, preference for the timing of resolution of risk and uncertainty has been a much discussed subject in recent economic and decision theoretic literature.[4] To frame a context for our contribution, we note that our axiomatization of preferences corresponds to recursive utility where the agent exhibits preferences for the timing of resolution of both risk and choice. In particular, when an agent possesses a preference for flexibility she may wish to 'defer' choice nodes. Although an aversion to flexibility (i.e., strict preference for commitment) makes intuitive sense in some contexts, such behavior is precluded from our theory because of our insistence on dynamic consistency.[5]

Finally, we note that partial orderings play an important role in the social choice and decision theory literature. For instance, Aumann (1962) and Bewley (1986, 1987) are seminal references in considering partial orderings in decision theory and economics, while more recent work relevant to our discussion includes Mitra and Ok (2000), Dubra and Ok (2000), Ok (2000) and especially Baucells and Shapley (1998), Dubra, Maccheroni and Ok (2001) and Sagi (2000). Other related references include Levi (1980) and Seidenfeld, Schervish and Kadane (1995).

Although the agents in this paper generally have a 'utility for flexibility', the repre-

[3] Of these references, Nehring (1999) and Ozdenoren (2001) are the closest in spirit to our work. These also axiomatize static representations for a preference for flexibility. Al-Najjar, Ozdenoren and Casadesus-Masanell (2001) relates the presence of a preference for flexibility to the existence of 'complexity' in decision making.

[4] Such as in Kreps and Porteus (1978, 1979), Chew and Epstein (1989, 1991), Epstein and Zin (1989) Machina (1989), Grant, Kajii and Polak (1998, 2000ab), and Skiadas (1997, 1998).

[5] For an axiomatic approach to aversion to flexibility, see Gul and Pesendorfer (2000, 2002ab). Their work is set in the context of self-control problems as opposed to unforeseen contingencies.

sentation of inter-temporal preference need not involve a positive linear weighting of future tastes as in Kreps (1979) and Nehring (1999). In particular, the representation does not require a probabilistically sophisticated approach to the uncertainty surrounding future tastes. This is a desirable feature, since it allows one to model Knightian uncertainty over future tastes (i.e., referring to subjective states without reference to subjective probability). We provide necessary and sufficient conditions for an additive representation as well as multi-prior (or maxmin) representation.

The rest of the paper is organized as follows. Section 2 introduces the basic axioms and concepts, and derives the main results on preference for flexibility. Section 3 presents a definition of a subjective state space motivated by local properties of preferences, as opposed to the representation of those preferences; it also provides a uniqueness result for the subjective state space and its associated topology. Before concluding, we present axioms for additive and maxmin representations.

2 Theory

2.1 Formulation of the Choice Problem and Agents' Preferences

We use the temporal-lottery framework of Kreps and Porteus (1978). Intuitively, an inter-temporal decision problem is a temporal lottery (such as the one in Figure 1) where, following a nature node, the agent receives an allotment for consumption ($z_{t,j}$'s in Figure 1) plus a choice set ($x_{t+1,j}$ - the boxes in Figure 1) of probability distributions (d_{t+1} - the circles in Figure 1). To review the formalism of Kreps and Porteus (1978), consider an arbitrary finite sequence of dates, $t \in 1, \ldots, T$, where at each date an agent must choose a distribution or lottery, d_t, from a current choice set, x_t. The lottery, d_t, is a probability measure over outcomes. Each outcome takes the form of a pair, (z_t, x_{t+1}), where $z_t \in Z_t$ is a bundle of goods in the compact metric space, Z_t, representing the goods available for consumption at date t. x_{t+1} is a future choice set containing distributions. Specifically, d_t is an element of D_t, the set of all probability measures over the Borel sets of $Z_t \times X_{t+1}$. In turn, X_{t+1}, representing all possible $t+1$ choice sets, is the set of all closed subsets in D_{t+1} endowed with the Hausdorff metric.

Since Z_T is metrizable and compact, and assuming $X_{T+1} \equiv \{\emptyset\}$, D_T is metrizable and compact in the *weak** topology (the topology of weak convergence in probability measures). Kuratowski (1950 - cf. §42) proves that X_T, the set of all closed subsets in D_T, is also a compact metric space. Thus $Z_{T-1} \times X_T$ is compact, meaning that D_{T-1} is metrizable and compact in the *weak** topology. Clearly, this can be continued recursively to $t = 0$, when the agent must choose a distribution, d_0 from a closed subset, x_0 of D_0.

An agent faced with a dynamic choice problem must select a distribution, d_t, from $x_t \subseteq D_t$ consistent with some ordering over D_t. The choice behavior of the agent at date t can thus be summarized by a preference relation, $\succeq_t$, over D_t. In the presence of unforeseen contingencies it does not make sense to assume that the agent's preference relation at each date is known since her choice behavior can change when an unforeseen event occurs. In contrast with Kreps and Porteus (1978), who assume that $\succeq_t$ is complete, negatively transitive, continuous and invariant under mixture (the von-Neumann and Morgenstern axioms), we impose a weaker structure on the agent's revealed preferences. Namely, we require that revealed preferences, denoted as $\succeq^*$,[6] are consistent with some partial ordering and place structure over that partial ordering:

Axiom 1. *(Revealed Preferences)*
The agent's revealed preference at date t, denoted as $\succeq_t^*$, is complete, transitive, continuous[7]*, and $\forall\ d, d \in D_t$*

$$d \succeq_t^P d' \Rightarrow d \succeq_t^* d'$$

where $\succeq_t^P$ is a reflexive and transitive partial ordering over D_t.

If the partial ordering, $\succeq_t^P$, is known it places limitations on revealed ranking of lottery trees. In the classic approach, the agent's date-t ordering (i.e., $\succeq_t^*$) is only actually observable at date t, but a time-consistency condition allows an observer to correctly *infer* future orderings from current choice behavior. The same will be partially true in our approach: a time consistency condition will allow an observer of revealed choice (i.e., $\succeq_t^*$) to correctly infer $\succeq_{t'}^P$

[6]The $\succeq_t^*$'s are *actual* ex-post preferences, in the language of Dekel, Lipman and Rustichini (2001).

[7]Continuity means that upper and lower contour sets of $\succeq_t^*$ are weak* closed.

for $t' > t$. Note that $\succeq_t^P$, in principle, can depend on a history. For notational simplicity, however, we suppress such dependence.[8]

We also impose the following:

Axiom 2. *(Continuity)*
For any $d_t \in D_t$, $\{d_t' \in D_t \mid d_t' \succeq_t^P d_t\}$ is weak closed. Further, if $d_n \to d$ and $d_n' \to d'$ are weak* convergent sequences in D_t, with $d_n \succeq_t^P d_n'$ for every n, then $d \succeq_t^P d'$.*

Axiom 3. *(Independence)*
For any $d_t, d_t', c_t \in D_t$ and $\alpha \in (0, 1]$,

$$d_t \succeq_t^P d_t' \Leftrightarrow \alpha d_t + (1-\alpha)c_t \succeq_t^P \alpha d_t' + (1-\alpha)c_t$$

Axiom 2 is a technical condition on the 'at-least-as-good-as' and 'no-better-than' sets of $\succeq_t^P$ while Axiom 3 is the Independence axiom applied to the partial ordering defined by $\succeq_t^P$. The interpretation of the Independence Axiom is standard: if d_t is preferred to d_t' for any conceivable realization of revealed preference, then this remains true for identical probabilistic mixtures of d_t and d_t' with some 'noise' variable (i.e., c_t). We do not take a strong stance on the normative value of this condition. However, as we demonstrate later, it is generally essential if one wishes to interpret the resulting endogenous state space as one containing 'expected utility states'.

Theorem 1. *Axioms 1-3 are necessary and sufficient for the following representation: for every $p, q \in D_t$,*

$$q \succeq_t^P p \iff \underset{U_t \in \Psi_t}{inf} \left(E_q[U_t] - E_p[U_t]\right) \geq 0 \tag{1}$$

where Ψ_t is a non-empty subset of $\mathcal{C}(Z_t \times X_{t+1})$, the set of real-valued, continuous and bounded functions on $Z_t \times X_{t+1}$, and $E_q[\cdot]$ denotes an expectation taken over the distribution q. The elements of Ψ_t are defined up to an affine (positive linear) transformation. Moreover, the closed convex cone spanned by Ψ_t is unique.

[8]Only trivial but notationally burdensome, modifications must be made to the Axioms to account for path-dependence (this was done in earlier versions of this paper).

All proofs are relegated to Appendix 2. The Theorem is derived in Dubra, Maccheroni and Ok (2001).[9] The affine equivalence and boundedness of elements of Ψ_t allows us to suppose without loss of generality that it is closed and that $0 \leq U_t \leq 1$ for any $U_t \in \Psi_t$. Note that with the normalization, the closed convex hull of Ψ_t is unique and compact (in the *sup* topology).

Ψ_t is a utility function set. In words, $\succeq_t^P$ is a Paretian ordering generated by the utility functions in Ψ_t: $\succeq_t^P$ ranks q and p if and only if all utility functions in Ψ_t agree on the relative ranking of q and p. In turn, the agent's revealed preference is consistent with the Paretian ordering induced by the many ordinary von Neumann-Morgenstern utility functions in Ψ_t. One can therefore interpret $\succeq_t^*$ as a 'social planner's ranking'.

It is both useful and intuitive to define several other relations, derived from $\succeq_t^P$:

Definition 1.

i) $\succ_t^{NP}$ *is the complement of the inverse of* $\succeq_t^P$ *(i.e.,* $d_t \succ_t^{NP} d_t'$ *if and only if it is not the case that* $d_t' \succeq_t^P d_t$*).*

ii) $\succeq_t^{NP}$ *is the weak* closure of* $\succ_t^{NP}$*.*

iii) $\succ_t^P$ *is the complement of the inverse of* $\succeq_t^{NP}$ *(i.e.,* $d_t \succ_t^P d_t'$ *if and only if it is not the case that* $d_t' \succeq_t^{NP} d_t$*).*

Applying Theorem 1 to the definitions and using the normalization of Ψ_t, it immediately follows that

$$q \succ_t^{NP} p \iff \max_{U_t \in \Psi_t} \left(E_q[U_t] - E_p[U_t]\right) > 0 \tag{2a}$$

$$q \succeq_t^{NP} p \iff \max_{U_t \in \Psi_t} \left(E_q[U_t] - E_p[U_t]\right) \geq 0 \tag{2b}$$

$$q \succ_t^{P} p \iff \min_{U_t \in \Psi_t} \left(E_q[U_t] - E_p[U_t]\right) > 0 \tag{2c}$$

$\succ_t^{NP}$ (*resp.* $\succeq_t^{NP}$) is a non-Paretian type of strict (*resp.* weak) preference where some, but not necessarily all, of the utility functions agree on a strict (*resp.* weak) ranking.[10] The

[9] The theorem was also independently derived in Sagi (2000) and (in a finite dimensional setting) in Baucells and Shapley (2001).

[10] It is both intuitive and true that $\succeq_t^{NP}$ contains $\succeq_t^P$.

relation $\succ_t^P$ corresponds to strict Pareto dominance.[11]

To ensure that $\succeq_t^P$ actually places non-trivial constraints on revealed preferences, we require that the strict partial order, $\succ_t^P$, is non-trivial (in particular, this implies that the constant utility function is not in any of the utility function sets):

Axiom 4. *(Non-triviality)*
$\succ_t^P$ *is non-empty for every* $t \in \{0, ..., T\}$*.*

Recall that the set $\succeq_t^P$ limits the potential choice behavior exhibited by revealed preferences. It is therefore reasonable to expect that whatever $\succeq_t^*$ turns out to be, its representation must be a function elements of Ψ_t.

Proposition 1. *Assume Axioms 1-4 and let* $V_t^* : D_t \mapsto \mathbb{R}$ *be a continuous representation for* $\succeq_t^*$*. Then for any* $d \in D_t$*,*

$$V_t^*(d) \equiv v_t^*\Big(\big(E_d[U]\big)_{U \in \Psi_t}\Big) \tag{3}$$

where $v_t^* : [0,1]^{\Psi_t} \mapsto \mathbb{R}$ *is continuous and non-decreasing in its arguments, and if* $E_d[U] > E_{d'}[U]$ *for every* $U \in \Psi_t$ *then* $V_t^*(d) > V^*(d')$*.*

According to Proposition 1 Ψ_t is a singleton if and only if $\succeq_t^P = \succeq_t^*$; in this case, the Axioms reduce to those of Kreps and Porteus (1978) and $\succeq_t^*$ has an expected utility representation. If Ψ_t contains more than one element, then $\succeq_t^*$ is not fully specified, but Theorem 1 implies that the set of $\succeq_t^*$'s consistent with $\succeq_t^P$ is non-empty (any expected utility functional with von Neumann-Morgenstern index from Ψ_t does the job). It is important to note that revealed (or 'ex-post') preference need not be linear in the U_t's and thus the space of revealed rankings consistent with $\succeq_t^P$ is far larger than the set of linear functionals generated by Ψ_t. we remark here that although one may be tempted to interpret any function of the form given in Eqn. (3) as an 'endogenous state' for tastes, this interpretation will generally not lead to a unique endogenous state space or the notion of endogenous states we described in the Introduction. To model the *realization* of choice behavior, one must make

[11]Note that $\succ_t^P$ does *not* generally coincide with the asymmetric part of $\succeq_t^P$. $\succ_t^P$, however, is contained in the asymmetric part of $\succeq_t^p$.

some assumptions about the nature of $\succeq_t^*$.[12] Other than the assumptions made in Axioms 1-4 we refrain from imposing any more structure on $\succeq_t^*$ or how it is realized.[13]

Finally, we emphasize that, without imposing additional structure, the fact that Ψ_t is non-singleton does not imply unforeseen contingencies in and of itself. An alternative possibility is that $\succeq_t^*$ is known by the agent at all dates previous to t to be an expected utility function, and $\succeq_t^P$ has no useful information. Another possibility is that $\succeq_t^*$ is predetermined, but is not a von Neumann-Morgenstern utility function over D_t. In this latter case, Ψ_t might be defined as the set of Gateaux derivatives of the functional representing $\succeq_t^*$ - assuming it is sufficiently well behaved.[14] Unforeseen contingencies can only be inferred from the agent's revealed choice behavior, and in particular, her concern with making sure that she does not commit to an ex-post inferior contingent plan. Such concern is specified by a time consistency requirement that we impose in the next subsection.

2.2 Time Consistency

At date t, the agent chooses a distribution, $d_t \in D_t$, from a menu of distributions, $x_t \in X_t$. If $d_t \in x_t$ is degenerate, then it will be henceforth identified with its outcome: a consumption-menu pair, (z_t, x_{t+1}). Thus if at date t the degenerate distribution that awards $(z, f) \in Z_t \times X_{t+1}$ weakly '$\succeq_t^P$-dominates' the degenerate distribution that awards $(z', g) \in Z_t \times X_{t+1}$, then we write,

$$(z, f) \succeq_t^P (z', g)$$

If a choice set, x_t, consists of only a single choice, say d_t, then it is denoted $x_t = \{d_t\}$. For any choice set, $x_t \in X_t$, define the *Pareto undominated* subset:

$$x_t^P \equiv \{d \in x_t \mid d \succeq_t^{NP} d', \forall d' \in x_t\} \tag{4}$$

[12]See Rigotti and Shannon (2001) and Bewley (1986, 1987) for examples in which one can deduce much without committing to a complete order.

[13]As we shall soon demonstrate, the fact that Ψ_t is not a singleton will lead to a well-defined endogenous state space filtration of 'tastes'. We are no more interested in making *objective* statements about the probabilistic evolution of these states than we would be in a Savage setting (where the states are exogenously specified).

[14]For instance, if revealed preference at any date, t, can always be represented by $V_t(d) = \min_{U \in \Psi_t} E_d[U]$ for any $d \in D_t$, then it is easy to check the validity of Axioms 1-3.

If the utility functions in Ψ_t are viewed as a set of agents, then x_t^P can be interpreted as the Pareto frontier of x_t. A more material interpretation, however, is that x^P contains all distributions that cannot be *ex-ante* ruled out by the strict partial ordering, $\succ_t^P$. Since, by Axiom 1, revealed choice, $\succeq_t^*$, is consistent with (i.e., is a completion of) $\succeq_t^P$, x^P can also be seen to contain all conceivable ex-post revealed choices.

Henceforth, we shall also assume, as do Dekel, Lipman and Rustichini (2001), that the agent can add to any menu of distributions, x, probabilistic mixtures of elements of x. In other words, by 'tossing coins' in selecting lotteries from x, the agent effectively convexifies x and, therefore, does not distinguish between x and Hull($\overline{x}$) (Hull($\overline{x}$) denotes the closed convex hull of x).

Axiom 5. *(Convexification)*
For any t, $z_t \in Z_t$ and $x_t \in X_{t+1}$,

$$(z_t, x_t) \succeq_t^P (z_t, Hull(\overline{x}_t)) \quad \textit{and} \quad (z_t, Hull(\overline{x}_t)) \succeq_t^P (z_t, x_t)$$

Aside from analytical convenience, this relatively weak assumption, along with Axiom 1, implies that any one revealed choice action at date $t+1$ will always be consistent with the date $t+1$ maximization of *some* element of Hull($\overline{\Psi}_{t+1}$). In other words, if one and only one choice is revealed at date $t+1$, one cannot empirically refute the position that the agent chose by maximizing the expected utility of some utility function from Hull($\overline{\Psi}_{t+1}$).[15]

Should the agent's behavior prior to date $t+1$ reflect beliefs that at date $t+1$ she will choose a distribution from the Pareto undominated set? Such a normative connection between ex-post and ex-ante choice is established through the following:

Axiom 6. *(Time Consistency)*
For any t, $z \in Z_t$, and sets $f, g \in X_{t+1}$, if for every $d \in g$ there is some $c(d) \in f$ such that $c(d) \succeq_{t+1}^P d$, and the relation is strict for some element of g^P, then $(z, f) \succ_t^P (z, g)$.

Intuitively, the Axiom states that f is ex-ante preferred to g whenever under every

[15]This is, essentially, a type of 'Second Welfare Theorem': any Pareto optimal allocation can be achieved in a convex set of alternatives by maximizing a 'social welfare' function that is affine in the agents' utilities. Note that if many choices are simultaneously elicited (e.g., by introspection or in an appropriate experimental setting), one can refute the hypothesis that revealed choice is expected utility in the same way that one can refute the linearity of a social planner's objective function.

(ex-post) contingency each element in g is no better than some corresponding element in f, *and* at least one element of the Pareto frontier of g (i.e., g^P) is strictly worse than some element in f. From a normative point of view, this seems hardly questionable. However, there is a sense in which Axiom 6 is too strong. We discuss this further in Section 2.4.

Continuity guarantees that Axiom 6 has a weak-preference counterpart:

Lemma 1. *Assume Axioms 1-6. For any t, $z \in Z_t$, and sets $f, g \in X_{t+1}$, if for every $d \in g$ there is some $c(d) \in f$ such that $c(d) \succeq^P_{t+1} d$, then $(z, f) \succeq^P_t (z, g)$.*

The following Theorem is the first major consequence of the time consistency condition:

Theorem 2. *Given Axioms 1-5, Axiom 6 is equivalent to the following:*
Fix any t, $z \in Z_t$, and $f, g \in X_{t+1}$ where $\hat{d} \succ^P_{t+1} q$ for some $\hat{d} \in f$ and $q \in D_{t+1}$. If $\max_{d \in f} E_d[\psi] \geq \max_{d \in g} E_d[\psi]$ for every $\psi \in Hull(\overline{\Psi}_{t+1})$ and strict inequality holds for at least one ψ, then $(z, f) \succ^P_t (z, g)$.

Remark 1. *If there exists some 'worst' distribution that is strictly $\succ^P_{t+1}$-dominated by every other distribution, then the condition* "where $\hat{d} \succ^P_{t+1} q$ for some $\hat{d} \in f$ and $q \in D_{t+1}$," *automatically holds. Thus the latter condition is of consequence only when there is no way to clearly determine what is a 'worst possible outcome'.*

Remark 2. *Taking stock of the development thus far, Axioms 1-5 imply that any* **single** *choice that the agent makes at date $t + 1$ is consistent with maximizing the expected utility of some element of the set $Hull(\overline{\Psi}_{t+1})$; Theorem 2, by way of Axiom 6, establishes that the agent fully anticipates this in her date t (or ex-ante) preferences.*

Remark 3. *Using the assumed continuity of preferences from Axiom 2, it is easy to prove that Theorem 2 has an analogue in the case of weak dominance. I.e., if $\max_{d \in f} E_d[\psi] \geq \max_{d \in g} E_d[\psi]$ for every $\psi \in Hull(\overline{\Psi}_{t+1})$ then $(z, f) \succeq^P_t (z, g)$.*

Remark 4. *Since $Hull(\overline{\Psi}_{t+1})$ is closed and convex, if $\max_{d \in f} E_d[U_{t+1}] > \max_{d \in g} E_d[U_{t+1}]$ for some $U_{t+1} \in Hull(\overline{\Psi}_{t+1})$, then the same strict inequality holds for a measurable subset of $Hull(\overline{\Psi}_{t+1})$ (i.e., all utility functions in a neighborhood of U_{t+1}).*

Remark 5. *Finally, note that if the utility function set at date $t+1$ contains only a single element, then Axiom 6 is equivalent to the Temporal Consistency Axiom in Kreps and Porteus (1978), and the theory reduces to their axiomatic formulation.*

The next consequence of Axiom 6 is that every date t utility function very nearly has the ordinal EU form discussed by Dekel, Lipman and Rustichini (2001); in other words, holding the date t consumption bundle fixed, each of the utility functions in Ψ_t is (almost everywhere) strictly increasing in each of the maximal utilities from $\text{Hull}(\overline{\Psi}_{t+1})$ attainable at date $t+1$. Formally, define the continuous mapping $w_{t+1} : \text{Hull}(\overline{\Psi}_{t+1}) \times X_{t+1} \mapsto \mathbb{R}$ as

$$w_{t+1}(\psi, x) \equiv \max_{d \in x} E_d[\psi] \tag{5}$$

Roughly speaking, $w_{t+1}(\cdot, x)$ denotes an infinite dimensional vector whose elements are indexed by elements of $\text{Hull}(\overline{\Psi}_{t+1})$. Fixing $x \in X_{t+1}$, it is clear that $w_{t+1}(\cdot, x)$ is an element of $\mathcal{C}(\text{Hull}(\overline{\Psi}_{t+1}))$, the space of bounded continuous real-valued functions defined over $\text{Hull}(\overline{\Psi}_{t+1})$. The space of such functions generated by all possible menus in X_{t+1} is defined via

$$W_{t+1} \equiv \Big\{ w_{t+1}(\cdot, x) \mid \forall x \in X_{t+1} \Big\}$$

Since $\text{Hull}(\overline{\Psi}_{t+1})$ is compact in its *sup* topology,[16] W_{t+1} is a compact subspace of $\mathcal{C}(\text{Hull}(\overline{\Psi}_{t+1}))$. Denote $w_{t+1}(\cdot, x) \geq w_{t+1}(\cdot, x')$ whenever $w_{t+1}(\psi, x) \geq w_{t+1}(\psi, x')$ for every $\psi \in \text{Hull}(\overline{\Psi}_{t+1})$; and denote $w_{t+1}(\cdot, x) > w_{t+1}(\cdot, x')$ whenever $w_{t+1}(\cdot, x) \geq w_{t+1}(\cdot, x')$ *and* there is some $\psi \in \text{Hull}(\overline{\Psi}_{t+1})$ for which the inequality is strict. The next theorem characterizes the functional form of each of the elements of Ψ_t (i.e., the basis set of utility functions that, via Proposition 1, characterizes the set of possible revealed preferences).

Theorem 3. *Assume Axioms 1-6 and fix $U_t \in \Psi_t$. Then for any $z \in Z_t$ and $x \in X_{t+1}$,*

$$U_t(z, x) \equiv u_t(z, w_{t+1}(\cdot, x)) \tag{6}$$

where $u_t : Z_t \times W_{t+1} \mapsto \mathbb{R}$ is a continuous function and $w_{t+1}(\cdot, x) \geq w_{t+1}(\cdot, x') \Rightarrow u_t(z, w_{t+1}(\cdot, x)) \geq u_t(z, w_{t+1}(\cdot, x'))$. Moreover, if there exists $d \in x$ such that $d \succ^P_{t+1} q$ for some $q \in D_{t+1}$ then $w_{t+1}(\cdot, x) > w_{t+1}(\cdot, x') \Rightarrow u_t(z, w_{t+1}(\cdot, x)) > u_t(z, w_{t+1}(\cdot, x'))$.

[16] Recall that elements of Ψ_{t+1} are normalized so that their image set is $[0, 1]$.

Remark 6. *Note that Axiom 4 implies that u is strictly increasing in $w_{t+1}(\cdot, x)$ almost everywhere. Also, as with Remark 1, if there exists some 'worst' distribution, q, that is strictly P-dominated by every other distribution, then the representation is strictly increasing everywhere.*

Remark 7. *Proposition 1 implies that $\succeq_t^*$ shares similar properties with respect to the $w_{t+1}(\cdot, x)$. In particular, Theorem 3 applies to the representation for $\succeq_t^*$ when it is restricted to degenerate distributions (i.e., date-t lotteries that pay (z, x) for sure).*

Remark 8. *If Ψ_{t+1} is a singleton then the representation resembles standard recursive utility - see Kreps and Porteus (1978). To truly reduce to their theory, Ψ_s must be a singleton for every $s > t$.*

The properties of U_t outlined in Theorem 3 are essentially those possessed by the utility functions defined in Kreps (1979) when agents have a preference for flexibility. Indeed, Theorem 3 implies that $U_t(z, x \cup x') \geq U_t(z, x)$ for any $x, x' \in X_{t+1}$; moreover, $U_t(z, x \cup x') = U_t(z, x) \ \Rightarrow U_t(z, x'' \cup x \cup x') = U_t(z, x'' \cup x)$. The result is even more closely related to the Ordinal EU representation of Dekel, Lipman and Rustichini (2001) in which 'ex-ante' utility is an aggregate of maximal expected 'ex-post' utilities.[17] The main differences are as follows: first, in rationalizing their representation, Dekel, Lipman and Rustichini (2001) informally appeal to the idea that ex-post ranking is expected utility. As Proposition 1 indicates, revealed 'ex-post' preference (i.e., $\succeq_{t+1}^*$) need not be expected utility. Theorem 3, on the other hand, says that aggregating ex-post expected utility functionals is sensible even when ex-post choice is not expected utility. Our approach therefore provides the behavioral rationale (in terms of axioms 3 and 6) for why it is that aggregating ex-post expected utility is a normatively sensible representation for utility for flexibility *whether or not* actual ex-post choice is expected utility. Another difference is that the normative nature of our axioms explicitly prevents an aversion to flexibility: in the language of Dekel, Lipman and Rustichini (2001) or Gul and Pesendorfer (2001), we only allow for positive 'states', meaning that the

[17] The Ordinal EU representation theorem in Dekel, Lipman and Rustichini (2001) contains a slight error. Our representation also satisfies their ordinal EU axiom, but the statement of their representation theorem should be qualified, as is ours, to deal with the case in which there is no 'worst' distribution.

representation is increasing in the vector of optimal ex-post expected utility functionals. Finally, the time consistency condition also imposes additional structure on the set of utility functions that acts as a subjective state space in Dekel, Limpan and Rustichini (2001): namely, the set must be closed and convex (i.e., Hull($\overline{\Psi}_{t+1}$)).

2.3 Observational Equivalence of the partial order and revealed preferences

Earlier we promised to demonstrate that the partial ordering, $\succeq_t^P$ can be observed by the agent's choice behavior. As mentioned before, in the classic approach, the agent's date-t ordering (i.e., $\succeq_t^*$) is only actually observable at date t, but time-consistency allows an observer to correctly *infer* future orderings from current choice behavior. The analogue in our theory is that the time consistency condition allows an observer of $\succeq_t^*$ to correctly infer $\succeq_{t'}^P$ for $t' > t$. This is formally contained in the next result:

Theorem 4. *Assume Axioms 1-6, fix $z \in Z_t$ and $x, x' \in X_{t+1}$, and assume that there exists $c \in x \cup x'$ such that $c \succ_{t+1}^P q$ for some $q \in D_{t+1}$. Then*

$$(z, x \cup x') \succ_t^* (z, x') \;\Leftrightarrow\; \exists d \in x \textit{ such that } d \succ_{t+1}^{NP} p' \;\; \forall p' \in x'$$

I.e., the agent strictly prefers the union of two sets to a subset if and only if the Pareto frontier of the union (i.e., the joint set of NP-undominated alternatives) is larger than that of the subset. The Theorem answers the question: when will the agent be willing to commit (i.e., agree to limit or reduce her choice set)? The answer, quite obviously, is only whenever restricting her choice set does not reduce the choice set's 'Pareto frontier'. Note that by letting $x \equiv \{d\}$ and $x' \equiv \{d'\}$, one can deduce $\succ_{t+1}^{NP}$ from $\succ_t^*$ almost everywhere.[18] This is sufficient to completely characterize $\succeq_{t+1}^P$. An immediate corollary is:

Corollary to Theorem 4:

Assume Axioms 1-6, and fix any $z \in Z_t$ and $d, d' \in D_t$ such that $d \succ_{t+1}^P q$ for some $q \in D_{t+1}$. Then

$$(z, \{d, d'\}) \sim_t^* (z, \{d\}) \textit{ and } (z, \{d, d'\}) \succ_t^* (z, \{d'\}) \;\Leftrightarrow\; d \succeq_{t+1}^P d' \textit{ and } d \succ_{t+1}^{NP} d'$$

[18] Axiom 4 guarantess that the condition, "...there exists $c \in x \cup x'$ such that $c \succ_{t+1}^P q$ for some $q \in D_{t+1}$" is satisfied almost everywhere.

Thus knowledge of $\succeq_t^*$ imparts full knowledge of $\succeq_{t+1}^P$ (and thus Ψ_{t+1}). Note that, by Theorem 3, this implies that Ψ_{t+2} is also known, etc.; once $\succeq_t^*$ is fully characterized, so are all the future *partial orderings*. This result places the partial ordering on the same observational footing as revealed preferences, and consequently justifies our claim that the axioms over $\succeq_t^P$ are indeed 'behavioral'.

Note that if Ψ_{t+1} contains only one function, Eq. (6) in Theorem 3 reduces to the time consistent recursive inter-temporal utility introduced by Kreps and Porteus (1978). In this case, Theorem 4 is a simple consequence of their Temporal Consistency Axiom (since $\succ_{t+1}^{NP}$ is complete and transitive); thus there is no preference for flexibility or deferment of choice. These observations motivate characterizing any agent facing a dynamic choice problem satisfying Axioms 1-6, as having an *Inter-temporal Flexibility Preference* (IFP). A formal definition is given by:

Definition 2. *A sequence of revealed preference relations over D_t, $\{\succeq_t^*\}$, is an* **Inter-termporal Flexibility Preference** *(IFP) if and only if it induces a temporal sequence of partial ordering, $\{\succeq_s^P\}_{s>t}$ that together with $\{\succeq_s^*\}_{s\geq t}$ obey Axioms 1-6.*

2.4 Extreme Tastes

At first blush, Axiom 6 seems entirely unobjectionable. On closer inspection, however, it does rule out behavior that might be deemed reasonable. For instance, assume that at date t the agent's preferences are such that a date $t+1$ menu containing a Shakespeare drama and an opera buffa is strictly preferred to a singleton menu containing either. Is it irrational for an agent to claim that although she might find herself in the mood for one or the other form of entertainment, the moods are mutually exclusive (i.e., under no circumstance she will be indifferent (or close to indifferent) between the two choices at date $t+1$)? In other words, is it reasonable for the agent to assume that she doesn't know her future tastes, but does know that they will be *extreme*?

Axiom 6 rules out extreme tastes because the agent must give weight to all 'convex combinations' of a set of basis tastes.[19] This means that she should consider the possibility

[19]This is implied by the fact that every element of Hull($\overline{\Psi}_{t+1}$) has weight in the agent's date t preferences

of being indifferent (and close to indifferent) between the opera and the play. Recall that strict preference between (z, f) and (z, g) is implied in Axiom 6 whenever any $d' \in g$ is ($\succeq^P_{t+1}$) dominated by something in f **and** an element on the 'Pareto frontier' of g (i.e., $d' \in g^P$) is *strictly* ($\succ^P_{t+1}$) dominated. The reason that all convex combinations of the basis rankings (say, the ranking that strictly prefers opera with that which strictly prefers the play) are weighted ex-ante is that the 'Pareto frontier' (as defined in Eqn. (4)) is rich. To see this and how it leads to a contradiction with extreme tastes, return to the example and suppose that g contains the play, the opera and all probabilistic mixtures of the two; suppose, further, that f is composed of three prospects: the play, the opera, and a gamble that awards the play, the opera, or \$10,000 with respective probabilities $(.5-\frac{\epsilon}{2}, .5-\frac{\epsilon}{2}, \epsilon)$, for some arbitrarily small $\epsilon > 0$. Finally, assume that $(.5-\frac{\epsilon}{2}, .5-\frac{\epsilon}{2}, \epsilon)$ strictly (i.e. $\succ^P_{t+1}$) dominates $(.5, .5, 0)$.[20]. Note the following line of reasoning:

i) Theorem 1 implies that the distribution $(.5, .5, 0)$ is not dominated by anything else in g, therefore, by definition, $(.5, .5, 0)$ is in g^P.

ii) f is equivalent to its convex hull, by Axiom 5.

iii) Since g is contained in the convex hull of f, every element of g is weakly dominated by an element of the convex hull of f.

iv) Since $(.5-\frac{\epsilon}{2}, .5-\frac{\epsilon}{2}, \epsilon) \succ^P_{t+1} (.5, .5, 0)$, Axiom 6 implies that the convex hull of f strictly dominates g (an implication of (iii) and the fact that an element of the convex hull of f strictly dominates an element of g^P).

v) f strictly dominates g since f is equivalent to its convex hull.

However, if the agent has extreme tastes, she must be indifferent between f and g. To see this consider that in her 'ex-post' decision, she will never select $(.5, .5, 0)$ (or any other mixture) from g, since she will either be in a mood for a comic opera or in a mood for a tragic play (and never close to indifferent). Likewise, she will never select $(.5-\frac{\epsilon}{2}, .5-\frac{\epsilon}{2}, \epsilon)$ from f. Thus it is not fair to claim that $(.5, .5, 0)$ is part of her date $t+1$ 'Pareto frontier', and

[20]This would be true, by Theorem 1, if \$10,000 dominated the opera or play regardless of contingency - a rather mild assumption, unless one lived in New York.

even less fair to claim that f strictly dominates g because $(.5 - \frac{\epsilon}{2}, .5 - \frac{\epsilon}{2}, \epsilon) \in f$ dominates $(.5, .5, 0) \in g$. The problem disappears if $(.5, .5, 0)$ and other mixtures are removed from the 'Pareto frontier' of g^P. That is precisely how Axiom 6 can be weakened to allow extreme tastes.

When D_t contains a single 'best' element, consider the following alternative definition for x^P: for any choice set, $x_t \in X_t$,

$$x_t^P \equiv \{d \in x_t \mid d \succeq_t^{NP} d^*, \text{ for some } d^* \in D_t \text{ s.t. } d^* \succeq_t^P d' \ \ \forall d' \in x_t\} \tag{4'}$$

d^* in the definition is anything that is at least as good as the entire set, x_t. The 'Pareto frontier' is composed of all elements of x_t that are not *strictly* dominated by *every* such d^*. For instance, setting d^* to be the date $t+1$ degenerate lottery that awards *both* opera and play, it's easy to see that the opera and the play are (separately) in g^P (since at date $t+1$ one of the opera and play awarded by d^* will be deemed worthless[21]). Note, however, that for the same reason, this is not true for the mixture $(.5, .5, 0) \in g$: intuitively, anything that $\succeq_{t+1}^P$-dominates both the opera and play will $\succ_{t+1}^P$-dominate their mixture. Hence g^P contains *only* the opera and play. Under this definition of g^P Axiom 6 implies that the agent is indifferent between f and g.

It is straight forward to show that under the alternative definition for x_t^P all previous results follow by replacing $\text{Hull}(\overline{\Psi}_{t+1})$ with $\mathcal{E}(\overline{\Psi}_{t+1})$, where $\mathcal{E}(\overline{x})$ is the closure of the extremal set of x. We leave the proof of this claim to the reader. Our point is that the time consistency axiom itself has universal appeal, but the definition of x^P on which it relies may be changed according to the application at hand. Note that any reasonable alteration of the original definition of x^P corresponds to reducing its size. Thus the original definition characterizes the broadest range of potential tastes in a 'changing tastes' interpretation.

2.5 Violations of the Independence Axiom

Although Axiom 3 has a normative flavor, it very much constrains the representation. Clearly, the aggregate expected utility representation arises due to this axiom. We illustrate the limitations of our axiomatization by an example.

[21] We assume free the disposal of tickets but that they are not fungible.

Suppose that payoffs are denominated in terms of wealth (i.e., the Z_t's correspond to a closed interval in $\mathbb{R}$) and the agent knows that at date T she will select among distributions based on some mean-variance criteria. At earlier dates, she does not know precisely what that criteria will be, but by observing her revealed preferences at date $T-1$ it is possible to back out (through the relationship in Theorem 4, for instance) that $\succeq_T^P$ is generated by the family: $\left\{U(d_T) = \mu(d_T) - \frac{A}{2}\sigma^2(d_T) \mid A \in [1,2]\right\}$, where $\mu(d_T)$ and $\sigma^2(d_T)$ denote, respectively, the mean and variance of a distribution, d_T. Consider three date T distributions: q, r, and s, with $\mu(q) = \bar{q}, \sigma^2(q) > 0$, $\mu(r) = \bar{q} - \sigma^2(q)$, $\mu(s) < \mu(r)$ and $\sigma^2(r) = \sigma^2(s) = 0$. According to these assumptions, the agent's revealed preferences indicate that q weakly dominates r, which strictly dominates s. In other words, $q \succeq_T^P r \succ_T^P s$. It is easy to show, however, that a probabilistic mixture of q with s does *not* generally $\succeq_T^P$-dominate the same mixture of r with s (in particular, for $A = 2$). Moreover, $T-1$ revealed preference would reflect that, in general, it is not the case that $\lambda q + (1-\lambda)s \succeq_T^P \lambda r + (1-\lambda)s$ for $\lambda < 1$. This is a direct violation of Axiom 3 and the agent's preferences cannot be described within our framework.

Since the mean-variance preferences used above are locally smooth, one can approximate them with convex combinations of two expected utility functionals:

$$\left\{U^{\xi}(d_T) = E_{d_T}[-\frac{A}{2}(x-\xi)^2 \mid \xi \in \{\mu_l + \frac{1}{2}, \mu_h + 1\}\right\}$$

where μ_l and μ_h are, respectively, the lowest and highest payoffs available. It is therefore possible to construct a representation of the agent in the example that aggregates maximal expected utility functions. This representation, however, will not be strictly increasing in all of the optimized expected utilities (see Theorem 3) or even in any of the generating functionals (even though when choosing wealth lotteries there generally is a 'worst' element).[22]

The reason that the independence axiom is violated is *not* because the agent's revealed preferences are sure to be non-expected utility - recall Proposition 1. The problem with the example is that the Pareto frontier implied by revealed preferences is not equivalent to the Pareto frontier generated by a single set of ex-post expected utility functionals. In the case described it would be far more sensible to derive a representation that aggregates

[22]In other words, one could represent the agent's behavior as in Eqn. (6), but the utility function set will vary with the choice set, x_t.

non-Expected Utility functionals.

3 Preference for Flexibility and a Subjective States Filtration

In an atemporal setting, Dekel, Lipman and Rustichini (2001) consider a representation for preferences over menus of distributions of the form

$$u\Big(\max_{d\in x} E_d[\psi^1], \max_{d\in x} E_d[\psi^2], \ldots, \max_{d\in x} E_d[\psi^\alpha], \ldots\Big)$$

where α is an index in a set, S, the ψ^α's are bounded functions over a *finite* set of payoffs and u is 'strictly increasing' in the ψ^α's. Such a representation is said to be an 'Ordinal EU Representation' and, by their Theorem 3, is equivalent to requiring that the preference ordering satisfy several axioms: weak order, continuity, non-triviality, monotonicity and weak independence. Of these, only the last two merit explanation. Monotonicity says that $x' \subseteq x$ implies that x is at least as good as x'. Weak independence requires that if x is strictly preferred to x' *and* $x' \subset x$, then $\lambda x + (1-\lambda)\bar{x}$ is also strictly preferred to $\lambda x' + (1-\lambda)\bar{x}$ for any $\lambda \in (0, 1]$. Dekel, Lipman and Rustichini (2001) establish several additional results: (i) if the index set, S, is finite, then any other representation that involves only expected utility functionals must use the same set of utility functions (i.e., the set of expected utility functions indexed by S is unique); (ii) any other representation that makes use of non-expected utility functionals will have a strictly larger index set; (iii) if S is infinite, then there are many equivalent representations using only expected utility functionals, but the closure of all such index sets is unique[23] (moreover, there is always a representations with a countable index set); and finally, (iv) if S is infinite but countable, any equivalent representation using non-expected utility functionals will have an infinite index set and thus a cardinality that is weakly greater than that of S.

In the literature on unforeseen contingencies, the set S is viewed as a subjective state space - the agent behaves as if she aggregates possible instances of future rankings, and those instances are indexed by elements of S. This may be sensible in an atemporal setting where ex-ante and ex-post choice are not formally related. Our own Theorem 3 implies

[23] Closure is defined with respect to a topology of the index set as described in Dekel, Lipman and Rustichini (2001).

an 'Ordinal EU Representation', and it is easy to show that revealed preference at date t satisfies weak order, continuity, non-triviality, monotonicity and weak independence. Is Hull($\overline{\Psi}_t$) therefore an endogenous state space? Proposition 1 states that the set of possible revealed rankings at date t is in general much larger than Hull($\overline{\Psi}_t$) - any non-decreasing (and possibly non-linear) function whose arguments are the ψ's in Ψ_t is a possible ranking. Is each 'possible' revealed preference ranking a state? If so, then one encounters a serious problem: ex-ante revealed preference only gives indication of how future Pareto frontiers are generated (i.e., given a menu, they specify the set of maximands of Hull($\overline{\Psi}_t$)), but give no additional information about how every conceivable ranking consistent with $\succeq_t^P$ ought to be weighted. In other words, it is not clear that associating future realized tastes with endogenous states is appropriate. One remedy for this is to impose additional structure on feasible $\succeq_t^*$'s. In particular, if one insists that revealed preference over D_t satisfies an Independence Axiom, then $\succeq_t^*$ *must* be an element of Hull($\overline{\Psi}_t$). There are shortcomings to this approach; for instance, it rules out 'contingencies' that induce non-expected utility behavior. Moreover, as we argue shortly, there are more general circumstances under which one is justified in viewing Hull($\overline{\Psi}_t$) as a subjective state space *without committing the* $\succeq_t^*$*'s to be expected utility functionals.*

Aside from such basic questions about the meaning of subjective states, there are additional technical concerns. Difficulties are encountered because the set of payoffs for lotteries in our theory is necessarily uncountably infinite (except, possibly, at date T). In such a setting, it is far from clear that the uniqueness result of Dekel, Lipman and Rustichini (2001) holds. Because they assume a finite set of lottery payoffs, the set of possible expected utility functions on their lottery space is a subset of a finite dimensional vector space. The dense sets of such a space are generally quite different from those of an infinite dimensional Banach space – i.e., the space we necessarily have to analyze in the inter-temporal setting. Finally, we mention that we are not completely satisfied with the cardinal 'minimality' characterization of Dekel, Lipman and Rustichini (2001) in another sense. For instance, if the state space associated with expected utility functionals is homeomorphic to the unit box $[0,1]^2$, their characterization does not guarantee the absence of a non-expected utility representation with the more parsimonious characterization homeomorphic to the unit interval $[0, 1]$. Note that the two state spaces are cardinally equivalent, and thus the expected utility state space is minimal according to Dekel, Lipman and Rustichini (2001). If it is possible to characterize a subjective state space for preferences via a topology, as opposed to a cardinality, then such a characterization would be deemed preferable (at least by us).

In this section we examine two questions: first we look for a natural definition of a unique subjective state space that does not depend on the representation used. Second, we find conditions under which $\text{Hull}(\overline{\Psi}_{t+1})$ is homeomorphic to the subjective state space. In particular, we wish to strengthen the results of Dekel, Lipman and Rustichini (2001) by giving conditions under which revealed preferences uniquely specify both a set of states and a topology. [24]

To begin, we seek a 'natural' definition for an endogenous state space. Consider the date $t+1$ menu, $x \equiv \{d_1, d_2, d_3, d_4\} \in X_{t+1}$. We ask the following question: what can be deleted from x without ex-ante welfare loss? To illustrate, suppose that fixing the date t consumption bundle at $z \in Z_t$, it so happens that the following revealed preference relations hold:

i) $(z, \{d_1, d_2\}) \succeq_t^* (z, x)$ and $(z, \{d_1, d_3\}) \succeq_t^* (z, x)$

ii) $(z, x) \succ_t^* (z, \{d_2, d_3, d_4\})$ and $(z, x) \succ_t^* (z, \{d_1, d_4\})$.

Thus d_4 can always be deleted from x without ex-ante welfare loss, while deleting d_1 will always cause ex-ante welfare loss. By contrast, one can delete one and only one of d_2 and d_3 without incurring welfare loss, but cannot delete both. Theorem 4 and its corollary can be used to demonstrate that $d_2 \succeq_{t+1}^P d_3$ and $d_3 \succeq_{t+1}^P d_2$, thus d_2 and d_3 are deemed equivalent under every realized date-$t+1$ revealed ranking (see Proposition 1). It therefore makes sense to identify d_2 and d_3 when discussing what can be deleted from x without ex-ante welfare loss. In summary: the critical 'elements' of x consist of the the equivalence classes $[d_1] \equiv \{d_1\}$ and $[d_2] \equiv \{d_2, d_3\}$ – the agent would strictly prefer to avoid their deletion. This is illustrated

[24] Again, this is important when the state space is infinite. The topology of the subjective state space is also important from a modeling point of view. In the Savage setting, where the filtration of states is exogenously given, one always specifies the states of nature as well as a topology for characterizing and measuring events.

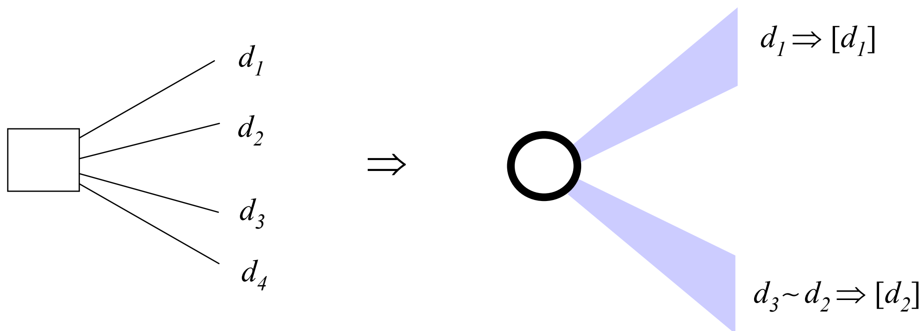

Figure 2: Turning a choice node into an endogenous-states node.

in Figure 2. We offer the following interpretation: each of the critical equivalence classes is an 'event' in an endogenous state space. By preserving critical elements at date t, the agent reserves the right to select from them (or a random mixture over them) in the next period. Note that this notion of states does not rely on the form of the representation but only on the underlying preferences. If the choice node (i.e., x) contains more elements, the number of 'events' that can be discerned may increase. In the figure each event is represented by a shaded 'branch' at the converted node: if x contained more equivalence classes that could not be deleted without welfare loss, each shaded branch would be further decomposed into a finer partition.

In generalizing the example, there are two formal difficulties:

i) For all intents and purposes, we are interested only in convex and closed choice sets of distributions. Continuity implies that deleting a measure zero set of equivalence classes will not result in welfare loss. We therefore have to refer to the removal of neighborhoods around indifference classes.

ii) In the example, the number of 'events' depends on the menu. If the menu is deformed slightly, the number of 'events' may change. At the very least, the definition of a state space should not vary within a neighborhood of a choice set. An appropriate definition of a state space must correspond to the finest 'event partition' that can be generated by slight deformations of x.

Note that the second point raises an issue that, to our knowledge, has not been discussed in the literature. Just as risk preferences can be characterized locally (see Machina (1982)), one can define an endogenous state as arising naturally from the *local* preference for flexibility of the agent.

To begin a formal development of the ideas above, let $d \in D_{t+1}$ and define, as before,

$$[d] \equiv \{d' \in D_{t+1} \mid d \succeq^P_{t+1} d' \ \&\ d' \succeq^P_{t+1} d\}$$

The agent knows at date t that she *will be* indifferent between points in $[d]$; in other words, $[d]$ is an ex-ante equivalence class corresponding to the symmetric part of $\succeq^P_{t+1}$.[25] It follows from Theorem 1 that

$$[d] = \Big(d + \ker(\Psi_{t+1})\Big) \cap D_{t+1}$$

where $\ker(\Psi_{t+1})$ is the subspace of signed measures whose expectation with respect to every element of Ψ_{t+1} is zero. $[d]$ is an element of the quotient space $D_{t+1}/\sim^P_{t+1}$ (where $\sim^P_{t+1}$ is the symmetric part of $\succeq^P_{t+1}$) that, by Theorem 1, is given by $D_{t+1}/\ker(\Psi_{t+1})$.[26]

Next, denoting the metric on D_{t+1} as ρ, define the associated ϵ-neighborhood of a set of distributions, $x \in X_{t+1}$, as $N^\rho_\epsilon(x)$.[27]

The following definition is meant to deal with the first of the two listed difficulties:

Definition 3. *Fix $x \in X_{t+1}$ and $z \in Z_t$. Define $[d] \in D_{t+1}/\sim^P_{t+1}$, where $\sim^P_{t+1}$ is the symmetric part of $\succeq^P_{t+1}$, to be an x-**critical indifference class** if and only if $\forall\ \epsilon > 0$,*

[25] For example: suppose an agent knows herself to be a mean-variance optimizer but does not know her future risk aversion coefficient. Each equivalence class would consist of all distributions with the same mean and variance.

[26] $D_{t+1}/\ker(\Psi_{t+1})$ is a subset of the larger quotient space $\mathcal{M}_{t+1}/\ker(\Psi_{t+1})$, where $\mathcal{M}$ is the space of signed measures, $\mathcal{M}_{t+1} \equiv \{\lambda d - \lambda' d' \mid \lambda, \lambda' \in \mathbb{R}_+,\ d, d' \in D_{t+1}\}$. Since $\ker(\Psi_{t+1})$ is a closed linear subspace of $\mathcal{M}_{t+1}$, $\mathcal{M}_{t+1}/\ker(\Psi_{t+1})$ is a Banach space.

[27] Here, we make use of the existence of a metric on D_{t+1} and its associated induced Hausdorff metric on X_{t+1}. Happily, our assumptions thus far guarantee that a metric that coincides with the weak* topology does in fact exist. Recall that D_{t+1} is a set of probability measures over a compact metric space, thus it itself is compact. Holmes (1975, p. 100) supplies the desired result. Note that this would not be true if D_{t+1} was not compact.

$(z, x) \succ_t^P (z, x \setminus N_\epsilon^\rho([d]))$.

Put simply, we identify 'critical' points of x up to indifference; an indifference surface is deemed critical when deleting it plus any neighborhood that contains it leads to welfare loss.

Let $R(x)$ be defined as the set of all x-critical indifference classes. To deal with the second listed difficulty, we define a state space at x as the unique (up to homeomorphism) topological space coinciding with the largest set of critical equivalence classes near x:

Definition 4. *Fix $x \in X_{t+1}$. $S(x)$ is a* **Topological Endogenous State Space** *at x if and only if there exists some $\delta > 0$ such that for every $0 < \epsilon < \delta$ the following two conditions hold*

i) $x' \in \{x'' \in X_{t+1} \mid ||x'' - x|| < \epsilon\} \Rightarrow R(x')$ is homeomorphic to a quotient of $S(x)$

ii) There exists some $x' \in \{x'' \in X_{t+1} \mid ||x'' - x|| < \epsilon\}$, such that $R(x')$ is homeomorphic to $S(x)$.

The first condition requires that the set of 'events' (i.e., critical classes) generated by any set near x corresponds to a partition of $S(x)$. The second condition requires that the finest partition that can be generated identifies $S(x)$. Note that the homeomorphism requirement guarantees topological uniqueness (which is the only sensible notion of uniqueness for a state space). There are two distinct advantages to this definition over the ones found in the literature (e.g., Kreps (1979, 1992), Nehring (1999), and Dekel, Lipman and Rustichini (2001)): first, the space is directly related to primitives as opposed to a representation; second, the definition specifies a state space along with a *topology* (the topology induced by $D_{t+1}/\ker(\Psi_{t+1})$). The latter is especially important when the state space is infinite (which is the case here).

We can now state the first main result of this section:

Theorem 5. *Let the Banach space generated by the linear span of Ψ_{t+1} (in its* sup *topology) be denoted as $\mathcal{B}_{t+1}$ and assume Axioms 1-6. Then a necessary condition for the topological endogenous state space to be homeomorphic to Hull$(\overline{\Psi}_{t+1})$ is that $\mathcal{B}_{t+1}$ has a Fréchet differentiable norm and its unit ball is weakly uniformly convex.*

Roughly speaking, the requirement on Ψ_{t+1} is that it is not too rich. Examples of candidate spaces include instances where Ψ_{t+1} spans a Hilbert space; this is true, for instance, when the generator set, Ψ_{t+1}, is finite. Ψ_{t+1}'s that are generally inadmissible include the set of all increasing continuous functions, the set of all concave functions, and the intersection of the latter two. Intuitively, if Ψ_{t+1} is too rich, then no menu, x, can generate a set of critical elements (event partition) that is fine enough to be identified with $\text{Hull}(\overline{\Psi}_{t+1})$.

Theorem 5 indicates when it is not possible for $\text{Hull}(\overline{\Psi}_{t+1})$ to be identified with an endogenous state space. It does not, however, help with the more practical question of when such an identification is valid. The next result supplies a partial answer:

Theorem 6. *Assume Axioms 1-6 and suppose $\mathcal{B}_{t+1}$ from Theorem 5 is a Hilbert Space. Then a topological endogenous state space exists, is independent of x, and is homeomorphic to Hull($\overline{\Psi}_{t+1}$).*

Although the theorem only supplies sufficient conditions, we can only conjecture that they are also necessary. The case where Ψ_t is finite dimensional is imortant. Clearly $\mathcal{B}_t$ is a finite dimensional Euclidean vector space and thus a Hilbert space, but note that $\text{Hull}(\overline{\Psi}_t)$, and thus the endogenous state space, is not finite. Moreover, If the space of outcomes is *finite* (as it might be at the terminal date, $t = T$ but cannot be at any other date), then $\mathcal{B}_t$ is necessarily a finite dimensional Euclidean vector space (a subspace of the space generated by all von Neumann-Morgenstern utility functions over the set of finite outcomes). In that case, the setting is identical to that studied by Dekel, Lipman and Rustichini (2001). In particular, Theorem 6 implies that the essentially unique subjective state space identified by them for an Ordinal EU representation coincides with ours whenever theirs is closed and convex.[28] One can therefore view our definition of a subjective state space as a refinement of theirs, motivated by explicit considerations of time-consistency for 'ex-post' choice.

Finally, note that Theorem 6 specifies a date t endogenous state space that is topologically the same across all choice sets. In other words, 'ex-post' Pareto frontiers look the same regardless of the choice node. This is a consequence of Axioms 3 and 6. Relaxing

[28]Even when it is not closed and convex, their space would be the same as ours assuming extreme tastes – see Section 2.4.

these will, in general, lead to a theory where the topology of the endogenous state space varies with the choice problem. Thus, although our definition of a topological state space is local, our axioms ensure that the topology is fixed globally. We note that under extreme tastes (Section 2.4) an endogenous state space will be a subset of $\mathrm{Hull}(\overline{\Psi}_{t+1})$ as long as $\mathcal{B}_{t+1}$ satisfies the Theorem conditions.

Remark 9. *Consider the endogenous state space implied by preferences at date* $t-1$ *(i.e.,* $\mathcal{B}_t$ *in Theorem 6). For an example where* $\mathcal{B}_t$ *is infinite dimensional, assume that every* $U_t \in \Psi_t$ *has the form*

$$U_t(z,x) = f_t\Big(z, \int \Big(\max_{d \in x} E_d[\psi]\Big) d\mu_z(\psi)\Big)$$

where $f_t : Z_t \times [0,1] \mapsto \mathbb{R}$ *is increasing in both arguments and* μ_z *is some regular Borel probability measure on Hull*$(\overline{\Psi}_{t+1})$*. Now assume that* $Z_t \subset \mathbb{R}$*, and that each of the* f_t*'s can be written as*

$$f_t(z,\omega) \equiv \int e^{-gz-\nu\omega} F_t(g,\nu)\, dg\, d\nu$$

where F_t *is square-integrable. Then the set of* F_t*'s spans a Hilbert space identifying* $\mathcal{B}_t$*.*

As a more particular example, consider functions of the form

$$f_t(z,\omega) = (1-\beta)\big(1 - z^{-\alpha_1}\big) + \beta(1 - \omega^{-\alpha_0})$$

for α_j*'s positive real constants in some bounded interval and* $\beta \in (0,1)$*. It is straight forward to demonstrate that each member of such a family is the integral (Laplace) transform of a square integrable* F_t*. Thus the space spanned by any collection of such* f_t*'s is a Hilbert space.*

Remark 10. *Suppose that the requirements of Theorem 6 are satisfied at every date* t*. In this case, one can easily augment the standard filtration (i.e., lottery tree) with subjective states in an obvious way; one simply converts each choice node at date* t *to a chance node containing branches indexed by elements of Hull*$(\overline{\Psi}_t)$*. An event is a subset of Hull*$(\overline{\Psi}_t)$ *sharing the same maximal elements from* x*, the menu at the choice node. Seen this way, the subjective states resemble standard 'Savage' states.*

3.1 Maxmin and Additive Representations

Our next result concerns the existence of Maxmin and additive representations for utility functions (i.e., elements of Ψ_t) that have the form in (6). To begin, we say that, given $z \in Z_t$, an ordering between $x, x' \in X_{t+1}$ is mixture invariant at date t with respect to the menu $x'' \in X_{t+1}$, whenever

$$U_t(z,x) \geq U_t(z,x') \;\Leftrightarrow\; U_t(z, \lambda x + (1-\lambda)x'') \geq U_t(z, \lambda x' + (1-\lambda)x'')$$

$\lambda x + (1-\lambda)x''$ denotes a probabilistic mixture of elements of x with those of elements of x''. Next, we say that the preference ordering over menus associated with U_t is convex if and only if $U_t(z,x') \geq U_t(z,x)$ and $U_t(z,x'') \geq U_t(z,x)$ imply $U_t(z, \lambda x' + (1-\lambda)x'') \geq U_t(z,x)$ for any $\lambda \in [0,1]$. Finally, we say that, given $z \in Z_t$, a menu, x'', is weakly most desirable whenever $U_t(z,x'') \geq U_t(z,x)$ for all $x \in X_{t+1}$.

Theorem 7. *Fix t, $z \in Z_t$ and $U_t \in \Psi_t$ and assume Axioms 1-6.*

i) Assume there is a weakly most desirable menu, $\bar{x}$. The ordering of every pair of menus, $x, x' \in X_{t+1}$ is has convex upper contour sets and is mixture invariant with respect to $\bar{x}$ if and only if

$$U_t(z,x) = f_t\left(z, \min_{\mu_z \in Q_z} \int \left(\max_{d \in x} E_d[\psi]\right) d\mu_z(\psi)\right)$$

where f_t is increasing in both arguments and Q_z is a closed and convex set of regular Borel measures on Hull$(\overline{\Psi}_{t+1})$.

ii) The ordering of every pair of menus, $x, x' \in X_{t+1}$ is mixture invariant with respect to any other menu in X_{t+1} if and only if

$$U_t(z,x) = f_t\left(z, \int \left(\max_{d \in x} E_d[\psi]\right) d\mu_z(\psi)\right)$$

where f_t is increasing in both arguments and μ_z is a unique positive probability measure on Hull$(\overline{\Psi}_{t+1})$.

The second part of the theorem is basically that derived by Dekel, Lipman and Rustichini (2001). Note, also, that the set Q_z in the first part of the theorem is not a set of probability measures but one of positive measures. Moreover, as noted in Maccheroni (2002), this set need not be unique.

4 Concluding Remarks

There are also practical advantages to the type of preference representation we derive here. Firstly, one can model seemingly 'inconsistent' choice without the pitfalls that such representations often entail (arbitrage in prices deduced from a time-inconsistent representative agent model, and the vulnerability to repeated manipulation that time-inconsistent agents exhibit). Moreover, the normative structure we suggest can make clear the distinction between hyperbolic discounting and time-inconsistency: the difference lies in how the agent treats commitment and flexibility. In particular, the framework easily lends itself to the investigation of behavior that is traditionally associated with time-inconsistency (e.g., procrastination). Another avenue to pursue is the existence of simple representations that easily lend themselves to solving portfolio choice problems (such as the Epstein-Zin (1989) model). Finally, an interesting question is that of aggregation: what happens to the endogenous state space in equilibrium? Can endogenous states be 'correlated' across agents? If so, to what degree can such states be priced or hedged?

APPENDIX: Proofs

Proof of Theorem 1:

See Dubra, Maccheroni and Ok (2001) and Sagi (2000). The latter has equivalent but slightly different axiomatization. $\square$

Proof of Proposition 1:

First note that for any $d, d' \in D_t$, $d \succ_t^P d' \Rightarrow d \succ_t^* d'$. Thus whenever $E_d[U] > E_{d'}[U]$ for every $U \in \Psi_t$ then $d \succ_t^P d'$ and consequently, $d \succ_t^* d'$ (or $V_t^*(d) > V_t^*(d')$. If V_t^* depends on something other than the set $\{E_d[U]\}_{U \in \Psi_t}$ then there is some d and d' in D_t such that $E_d[U] = E_{d'}[U]$ for all $U \in \Psi_t$ and yet $V_t(d') > V_t(d)$. Axiom 4 implies the existence of $p, q \in D_t$ such that $q \succ_t^P p$. Theorem 1 leads to: $(1-\epsilon)d + \epsilon q \succ_t^P (1-\epsilon)d' + \epsilon q$ for every $\epsilon \in (0,1)$, but this is not consistent with both $d' \succ_t^* d$ and continuity of $\succeq_t^*$. Thus V_t^* can only depend on the set $\{E_d[U]\}_{U \in \Psi_t}$. Finally, continuity of v_t^* over $[0,1]^{\Psi_t}$ follows from continuity of V_t^* and continuity of the projection of D_t into $[0,1]^{\Psi_t}$. $\square$

Proof of Lemma 1:

Fix t, $z \in Z_t$, convex sets $f, g \in X_{t+1}$, and suppose that for every $d \in g$ there is some $d' \in f$ such that $d' \succeq_{t+1}^P d$. Axiom 4 implies that there must exist $q, p \in D_{t+1}$ such that $q \succ_{t+1}^P p$. Construct the sequence of sets $\{f_n\}$ and $\{g_n\}$ where $f_n \equiv (1-2^{-n})f + 2^{-n}q$ (the probabilistic mixture is taken element by element) and $g_n \equiv (1-2^{-n})f + 2^{-n}p$. Theorem 1 guarantees that for every $d \in g_n$ there is some $d' \in f_n$ such that $d' \succ_{t+1}^P d$. Thus, by Axiom 6, $(z, f_n) \succ_t^P (z, g_n)$. The continuity axiom then implies the desired result. $\square$

Proof of Theorem 2:

For necessity, assume that if $d \succ_{t+1}^P q$ for some $d \in f$ and $q \in D_{t+1}$, $\underset{d \in f}{max}\, E_d[\psi_{t+1}] \geq \underset{d \in g}{max}\, E_d[\psi_{t+1}]$ for every $\psi_{t+1} \in \mathrm{Hull}(\overline{\Psi}_{t+1})$ and strict inequality holds for at least one ψ_{t+1}, then $(z, f) \succ_t^P (z, g)$. Now, suppose that f and g are convex, and that for every $d \in g$ there is some $d' \in f$ such that $d' \succeq_{t+1}^P d$, and the relation is strict for some element of g^P. Since each element in g is weakly dominated by some element in f, Theorem 1 implies that $\underset{d\in f}{max}\ E_d[\psi_{t+1}] \geq \underset{d\in g}{max}\ E_d[\psi_{t+1}]$ for every $\psi_{t+1} \in \text{Hull}(\overline{\Psi}_{t+1})$. The existence of a strict inequality for some utility function follows from the fact that g is convex, meaning that by Theorem 1 g^P contains the set of maximands of $\text{Hull}(\overline{\Psi}_{t+1})$ over g. Again, by Theorem 1, one of these maximands will be strictly dominated (with respect to all elements of Ψ_{t+1}) by an element of f. By hypothesis, it follows that $(z,f) \succ_t^P (z,g)$

To prove sufficiency, assume that Axiom 6 holds. Suppose that $\underset{d\in f}{max}\ E_d[\psi_{t+1}] \geq \underset{d\in g}{max}\ E_d[\psi_{t+1}]$ for every $\psi_{t+1} \in \text{Hull}(\overline{\Psi}_{t+1})$ with a strict inequality holding for at least one ψ_{t+1}. Also, suppose that $d_f \succ_{t+1}^P q$ for some $d_f \in f$ and $q \in D_{t+1}$. Assume, for the time being, that both f and g are convex. Consider $d \in g$ such that d is not $\succeq_{t+1}^P$-dominated by any $d' \in f$. Define a cone at d via

$$B_d \equiv \{\mu \in D_{t+1} \mid \min_{\psi\in\Psi_{t+1}} E_{\mu-d}[\psi] \geq 0\ \}$$

By Theorem 1, $B_d \cap f = \emptyset$, otherwise d would be weakly dominated by some element of f. Because both B_d and f are convex and closed[29], they can be separated by a linear functional that supports the cone B_d at d. The first Bishop-Phelps Theorem (see Holmes (1975) p. 166) implies that the separating functional is arbitrarily close to some $\psi \in \text{Hull}(\overline{\Psi}_{t+1})$. For such a ψ, $\max_{d'\in f} E_{d'}[\psi] < E_d[\psi]$. This contradicts the hypothesis; thus, it cannot be that $B_d \cap f = \emptyset$. But then Theorem 1 implies that there is some $d' \in f$ such that $d' \succeq_{t+1}^P d$.

It remains to demonstrate strict dominance over some $d \in g^P$. By hypothesis, there is a utility function $\psi_{t+1}^* \in \text{Hull}(\overline{\Psi}_{t+1})$ that is maximized at some $d \in g^P$ with $E_{d'}[\psi_{t+1}^*] > E_d[\psi_{t+1}^*]$ for some $d' \in f$. Now, Axiom 2 implies that there is some $\alpha > 0$ such that $E_{d'}[\psi_{t+1}^*] > E_{(1-\alpha)d'+\alpha q}[\psi_{t+1}^*] > E_d[\psi_{t+1}^*]$. Note that $(1-\alpha)d' + \alpha d_f \succ_{t+1}^P (1-\alpha)d' + \alpha q \succ_{t+1}^{NP} d$ and that $(1-\alpha)d' + \alpha d_f \in f$ (since f is convex). Let g' be defined as the set constructed by taking the closure of the convex hull of $g \cup (1-\alpha)d' + \alpha q$. Clearly $\underset{d\in f}{max}\ E_d[\psi_{t+1}] \geq \underset{d\in g'}{max}\ E_d[\psi_{t+1}]$ for every $\psi_{t+1} \in \text{Hull}(\overline{\Psi}_{t+1})$ with a strict inequality holding for ψ_{t+1}^*. Thus, by an argument similar to the one in the previous paragraph, $\exists d'' \in f$ such that $d'' \succeq_{t+1}^P \hat{d}$ for every $\hat{d} \in g'$. Moreover, ${g'}^P$ contains, by construction, $(1-\alpha)d' + \alpha q$ which is strictly dominated by $(1-\alpha)d' + \alpha d_f \in f$. Thus Axiom 6 implies that $(z,f) \succ_t^P (z,g')$. Moreover, since Lemma 1 implies that $(z,g') \succeq_t^P (z,g)$, it follows that $(z,f) \succ_t^P (z,g)$.

[29]The fact that B_d is closed is demonstrated in an updated version of Dubra, Maccheroni and Ok (2001).

Finally, if f and g are not convex, then the results can be derived for the closed convex hulls of f and g. Axiom 5 then gives the desired ordering. $\square$

Proof of Theorem 3:

Fix $U_t \in \Psi_t$, $z \in Z_t$, and $x, x' \in X_{t+1}$. Consider the case where $w_{t+1}(\cdot, x) = w_{t+1}(\cdot, x')$. Theorem 2 and a continuity argument as given in Lemma 1 imply that $(z, x) \succeq_t^P (z, x')$ and $(z, x') \succeq_t^P (z, x)$. From Theorem 1 it must be that $U_t(z, x) = U_t(z, x')$. This establishes that

$$U_t(z, x) \equiv u_t(z, w_{t+1}(\cdot, x))$$

Continuity follows from the fact that both $w_{t+1}(\cdot, x))$ and U_t are continuous over X_{t+1}. The rest of the properties follow directly from Lemma 1 and Theorem 2. $\square$

Proof of Theorem 4:

Fix $z \in Z_t$ and $x, x' \in X_{t+1}$.

Suppose $\exists d \in x$ such that $d \succ_{t+1}^{NP} p'$ for all $p' \in x'$. Clearly $w_{t+1}(\cdot, x \cup x') > w_{t+1}(\cdot, x')$; thus by Theorem 3, $(z, x \cup x') \succ_t^P (z, x')$. Axiom 1 therefore implies that $(z, x \cup x') \succ_t^* (z, x)$.

Now, suppose that $(z, x \cup x') \succ_t^* (z, x')$. Theorem 1 and the definition of $\succ_t^{NP}$ imply that $(z, x \cup x') \succ_t^{NP} (z, x')$. Theorem 1 guarantees that for some $U_t \in \Psi_t$ it is the case that $U_t(z, x \cup x') > U_t(z, x')$; so Theorem 3 gives that there exist $\psi \in \text{Hull}(\overline{\Psi}_{t+1})$ and $d \in x \cup x'$ such that $E_d[\psi] > \underset{p' \in x'}{max}\ E_{p'}[\psi]$. Note that it must be that $d \in x$. By Theorem 1 this establishes that there exists $d \in x$ such that $d \succ_{t+1}^{NP} p'$ for every $p' \in x'$. $\square$

Proof of Corollary to Theorem 4:

This is a direct consequence of Theorem 4 and the fact that $d' \not\succ_{t+1}^{NP} d \Leftrightarrow d \succeq_{t+1}^{P} d'$. $\square$

Proof of Theorem 5:

The proof is trivial when Ψ_{t+1} is a singleton; assume, therefore, that it is not. Let $\mathcal{M}_{t+1}$ be the linear space of signed regular Borel measures generated by $\lambda(d - d')$ for any $d, d' \in D_{t+1}$ and $\lambda \in \mathbb{R}$. Note that $\mathcal{M}_{t+1}$ is a Banach space and that $[\mathcal{M}_{t+1}] \equiv \mathcal{M}_{t+1}/\ker(\Psi_{t+1})$ is homeomorphic to $\mathcal{B}_{t+1}^*$, the dual of $\mathcal{B}_{t+1}$ (the Banach Space spanned by Ψ_{t+1} – see Theorem

16E in Holmes (1975)). Also, note that $D_{t+1}/\ker(\Psi_{t+1})$ is homeomorphic to a convex and compact subset of $\mathcal{B}^*_{t+1}$.[30]

We seek to characterize the set $R(x)$. Consider $x \in X_{t+1}$ and assume, without loss of generality (due to Axiom 5) that x is convex. The continuity axiom (Axiom 2) and the linearity of the projection of D_{t+1} onto $D_{t+1}/\ker(\Psi_{t+1})$ ensures that $[x] \equiv \{[d] \mid d \in x\}$ is closed and convex. Let $\overline{E}(x)$ be defined as the closure of the set of extremal points of $[x]$.[31] Now, we claim that $R(x) \subseteq \overline{E}(x)$. To see this, note that $\left(\operatorname*{argmax}_{[d]\in[x]} E_d[\psi]\right) \cap \overline{E}(x) \neq \emptyset$ $\forall \psi \in \Psi_{t+1}$ (any linear functional is maximized at some extremal point). Thus, thanks to Theorem 2 one can always find a neighborhood of $[d] \notin \overline{E}(x)$ that is disjoint from $\overline{E}(x)$ and can be deleted without utility loss.

Theorem 2 implies that deleting a neighborhood around an extremal point that supports an element of $\mathrm{Hull}(\overline{\Psi}_{t+1})$ will lead to utility loss. Thus $R(x)$ contains elements of $\overline{E}(x)$ that are support points of x via functionals from $\mathrm{Hull}(\overline{\Psi}_{t+1})$. Formally, $R(x)$ is the closure of all such support points. I.e., in general[32]

$$R(x) = \mathrm{Closure}\left(\left\{[d] \in [x] \ : \ d = \operatorname*{argmax}_{d'\in x} E_{d'}[\psi], \ \psi \in \mathrm{Hull}(\overline{\Psi}_{t+1})\right\}\right) \cap \overline{E}(x) \tag{A-1}$$

The objective is to demonstrate that $R(x)$ is homeomorphic to a subset of $\mathrm{Hull}(\overline{\Psi}_{t+1})$ (this is equivalent to requiring $R(x)$ to be homeomorphic to a quotient of $\mathrm{Hull}(\overline{\Psi}_{t+1})$). By Eqn. (A-1), a necessary and sufficient condition for arbitrary $[x]$ is that the set of extreme points of $[x]$ is homeomorphic to a subset of the unit ball of $\mathcal{B}_{t+1}$. In turn, this is true if and only if $\mathcal{B}^*_{t+1}$ (in its norm topology) is homeomorphic to $\mathcal{B}_{t+1}$ (in its norm topology). Necessary and sufficient conditions for the latter are that $\mathcal{B}_{t+1}$ has a Fréchet differentiable norm and its unit ball is weakly uniformly convex (see Cudia (1964) Theorem 4.18). $\square$

[30] Compactness follows from the fact that the range of elements of Ψ_{t+1} over probability measures is $[0,1]$ - recall that the elements of Ψ_{t+1} are normalized to take values in $[0,1]$.

[31] The set of extremal elements is the set of points in $[x]$ that cannot be generated as convex combinations of other points in $[x]$.

[32] Note that in general,

$$\mathrm{Closure}\left(\left\{[d] \in [x] \ : \ d = \operatorname*{argmax}_{d'\in x} E_{d'}[\psi], \ \psi \in \mathrm{Hull}(\overline{\Psi}_{t+1})\right\}\right) \not\subseteq \overline{E}(x)$$

thus Eqn. (A-1) is not redundant.

Proof of Theorem 6:

Continuing from the proof of Theorem 5, it suffices to demonstrate that in the neighborhood of any $x \in X_{t+1}$, there is some x' such that $R(x')$ is homeomorphic to $\text{Hull}(\overline{\Psi}_{t+1})$. We begin by showing that a linear function from $\text{Hull}(\overline{\Psi}_{t+1})$ optimized over a sufficiently small ball in $\mathcal{B}^*_{t+1}$ can be uniquely identified with a probability measure. Fix $\psi \in \text{Hull}(\overline{\Psi}_{t+1})$ and consider the optimization program:

$$\max_{\mu \in \mathcal{B}^*_{t+1}} \left\{ < \mu, \psi > \right\} \quad ||\mu|| \leq \epsilon$$

Since μ and ψ are in the same space ($\mathcal{B}_{t+1} \sim \mathcal{B}^*_{t+1}$), μ can be identified with some square-integrable density function, ψ_μ and the problem reduces to a simple quadratic optimization with solution $\psi_\mu(\epsilon) = \epsilon \frac{\psi - \int \psi}{||\psi - \int \psi||}$ (recall that μ has total measure of zero). Now, let $\eta \equiv \min_{\psi \in \text{Hull}(\overline{\Psi}_{t+1})} \frac{||\psi - \int \psi||}{\int \psi}$; note that $\eta > 0$ since each $\psi \in \text{Hull}(\overline{\Psi}_{t+1})$ is continuous and has range of $[0, 1]$; next, let m be the Lebesgue measure of $Z_{t+1} \times X_{t+2}$. Thus, for any $0 < \epsilon < \frac{\eta}{m}$, $\frac{1}{m} + \psi_\mu(\epsilon)$ is everywhere greater than zero and its integral is 1, thereby uniquely defining a density function for a probability measure over $Z_{t+1} \times X_{t+2}$. Denote this measure as $c(\psi, \epsilon)$.

Now fix some arbitrary $x \in X_{t+1}$ and set

$$y \equiv \bigcup_{[d] \in [x]} \{(1 - \delta)[d] + \delta(\mu + [\mathbf{1}]) - [0] \mid ||\mu|| \leq \epsilon,\ \mu \in \mathcal{B}^*_{t+1}\}$$

where $[\mathbf{1}]$ is the uniform probability measure, and let $[x'] \equiv (y + [0]) \cap D_{t+1}/\ker(\Psi_{t+1})$. Clearly, by choosing δ sufficiently small, x' can be made arbtrarily close to x. Note first that y is a smooth and solid set (it is the union of smooth sets with non-empty interiors) thus its extreme points are homeomorphic to the unit sphere in $\mathcal{B}_{t+1}$ where the homeomorphism is given by

$$H(\psi) = \operatorname*{argmax}_{\mu \in y} \ < \mu, \psi >$$

Moreover, each $\psi \in \text{Hull}(\overline{\Psi}_{t+1})$ is maximized on y at the surface of a ball about $(1 - \delta)[d] + \delta[\mathbf{1}] - [0]$, for some $[d] \in [x]$. By the argument in the previous paragraph, the maximand, and thus the associated unique extreme point, is given by $(1 - \delta)[d] + \delta[c(\psi, \epsilon)] - [0]$ where $c(\psi, \epsilon)$ is a probability measure for sufficiently small ϵ. In particular, this means that ψ is uniquely maximized by $(1 - \delta)[d] + \delta[c(\psi, \epsilon)]$ over $[x']$. By Eqn. (A-1), this establishes an isomorphism between $R(x')$ and $\text{Hull}(\overline{\Psi}_{t+1}))$ through $H(\cdot)$. Clearly this is also a homeomorphism. $\square$

Proof of Theorem 7:

Note that for any $\psi \in \text{Hull}(\overline{\Psi}_{t+1})$,

$$\max_{d \in \lambda x + (1-\lambda)x''} E_d[\psi] = \max_{d \in x, d'' \in x''} \{\lambda E_d[\psi] + (1-\lambda)E_{d''}[\psi]\} = \lambda \max_{d \in x} E_d[\psi] + (1-\lambda)\max_{d \in x''} E_d[\psi]$$

Thus $w_{t+1}(\cdot, \lambda x + (1-\lambda)x'') = w_{t+1}(\cdot, x) + (1-\lambda)w_{t+1}(\cdot, x'')$. The space of all such $w(\cdot,\cdot)$'s, namely W_{t+1}, is a mixture space and U_t induces a weak and continuous ranking on it. The first result follows directly from Maccheroni (2002) who proves a maxmin representation in terms of linear functional over a mixture space; the role of utility functions in his Theorem is played by scaled positive measures over $\text{Hull}(\overline{\Psi}_{t+1})$ in our case. The second part of the theorem is standard (see Hestein and Milnor (1952)) and follows from the fact that the ordering induced by U_t over the mixture space W_{t+1} satisfies an independence axiom and must therefore be a monotonic transformation of a linear functional on W_{t+1}. Note that in both parts, only positive measures (i.e., linear functionals) with full support are allowed so that the representation is increasing. Moreover, uniqueness of Q_z is not guaranteed (see Maccheroni (2002)) in the first part of the Theorem, whereas uniqueness up to affine equivalence of the linear representation guarantees the uniqueness of the probability measure in the second part. $\square$

References

[1] Al-Najjar N., E. Ozdenoren and R. Casadesus-Masanell, Probabilistic Representation of Complexity, *Working Paper* (2001).

[2] Aumann, R., Utility Theory Without the Completeness Axiom, *Econometrica* **30** (1962), 445-462.

[3] Bossert, W., Opportunity Sets and Individual Well-Being, *Social Choice and Welfare* **14** (1997), 97-112.

[4] Bossert, W., P.K. Pattanaik, and Y. Xu, Ranking Opportunity Sets: An Axiomatic Approach, *J. Econ. Theor.* **63** (1994), 326-345.

[5] Bewley, T.F., Knightian Decision Theory: Part I, *Cowles Foundation* Discussion Paper (1986).

[6] Bewley, T.F., Knightian Decision Theory: Part II, *Cowles Foundation* Discussion Paper (1987).

[7] Chew, S.H., and L.G. Epstein, The Structure of Preferences and Attitudes towards the Timing of the Resolution of Uncertainty, *International Economic Review* **30** (1989), 103-117.

[8] Chew, S.H., and L.G. Epstein, Non-Expected Utility Preferences in a Temporal Framework with an Application to Consumption-Savings Behaviour, *Journal of Economic Theory* **50** (1990), 54-81.

[9] Dekel, E, B. Lipman and A. Rustichini, Representing Preferences With A Unique Subjective State Space for Unforeseen Contingencies, *Econometrica* **69** (2001), 891-934.

[10] Dubra, J., F. Maccheroni, and E.A., Ok, Expected Utility Theory without the Completeness Axiom, *NYU Economics Working Paper* (2001).

[11] Dubra, J. and E.A., Ok, A Model of Procedural Decision Making in the Presence of Risk, *NYU Economics Working Paper* (2000).

[12] Epstein, L.G. and S.E. Zin, Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, *Econometrica* **57** (1989), 937-969.

[13] Grant, S., A. Kajii and B. Polak, Instrinsic Preference for Information, *Journal of Economic Theory* **83** (1998), 233-259.

[14] Grant, S., A. Kajii and B. Polak, Temporal Resolution of Uncertainty and Recursive Non-Expected Utility Models, *Econometrica* **68** (2000a), 425-434.

[15] Grant, S., A. Kajii and B. Polak, Preference for Information and Dynamic Consistency, *Theory and Decision* **48** (2000b), 263-286.

[16] Gul, F. and W. Pesendorfer, Temptation and Self-Control, *Econometrica* **69** (2001), 1403-1435.

[17] Gul, F. and W. Pesendorfer, Self-Control and the Theory of Consumption, *Princeton Economic Working Paper*, (2002a).

[18] Gul, F. and W. Pesendorfer, The Revealed Preference Theory of Changing Tastes, *Princeton Economic Working Paper*, (2002b).

[19] Holmes, R.B., *Geometric functional analysis and its applications.* (1975) New York : Springer-Verlag.

[20] Herstein, I.N. and J. Milnor, An axiomatic approach to measurable utility, *Econometrica* **21** (1953), 291297.

[21] Kreps, D.M., and E.L. Porteus., Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica* **46** (1978), 185-200.

[22] Kreps, D.M., and E.L. Porteus., Temporal von Neumann-Morgenstern and Induced Preferences. *Journal of Economic Theory* **20** (1979), 81-109.

[23] Kreps, D.M., A Representation Theorem for "Preferences for Flexibility". *Econometrica* **47** (1979), 565-577.

[24] Kreps, D.M., Static Choice in the Presence of Unforeseen Contingencies. In *Economic Analysis of Markets and Games*, ed. By P. Dasgupta, et. al. (1992) Cambridge: MIT Press, pp. 565-577.

[25] Kuratowski, C., *Topologie*, Vol. 2. (1950) Warsaw: Polska Akademia.

[26] Levi, I., *The Enterprise of Knowledge.* (1980) Cambridge: MIT Press.

[27] Machina, M.J., Expected Utility Analysis Without the Independence Axiom, *Econometrica* **50** (1982), 277-323.

[28] Machina, M.J., Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty, *Journal of Economic Literature* **27** (1989), 1622-1668.

[29] Mitra, T. and E.A., Ok, Extending Stochastic Dominance Relations via Expected Utility Theory, *NYU Economics Working Paper* (2000).

[30] Nehring, K., Preference for Flexibility in a Savage Framework, *Econometrica* **67** (1999), 101-119.

[31] Nehring, K. and C. Puppe, Continuous Extensions of an Order Set to the Power Set, *J. Econ. Theor.* **68** (1996), 456-479.

[32] Nehring, K. and C. Puppe, On the Multi-Preference Approach to Evaluating Opportunities, *Social Choice and Welfare* **16** (1999), 41-63.

[33] Ok, E.A., Utility Representation of an Incomplete Preference Relation, *NYU Economics Working Paper* (2000).

[34] Ozdenoren, E., Completing the State Space with Subjective States, *Journal of Economic Theory*, forthcoming (2002).

[35] Pattanaik, P.K. and Y. Xu, On Preference and Freedom, *Theory and Decision* **44** (1998), 173-198.

[36] Puppe, C., Freedom of Choice and Rational Decision, *Social Choice and Welfare* **12** (1995), 137-153.

[37] Puppe, C., An Axiomatic Approach to “Preference for Freedom of Choice”, *J. Econ. Theor.* **68** (1996), 174-199.

[38] Rigotti, L. and C. Shannon, Uncertainty and Risk in Financial Markets, *Berkeley Economics* Working paper (2001).

[39] Sagi, J.S., *Partial Ordering of Risky Choices: Anchoring, Preference for Flexibility and Applications to Asset Pricing*, Ph.D. Dissertation, University of British Columbia, 2000.

[40] Seidenfeld, T., M.J. Schervish and J.B. Kadane, A Representation of Partially Ordered Preferences, *Ann. Stat.* **23** (1995), 2168-2217.

[41] Shapley, L. and M. Baucells, “Multiperson Utility,” UCLA Working Paper 779, (1998).

[42] Skiadas, C., Recursive Utility and Preferences for Information, *Economic Theory* **12** (1998), 293-312.

[43] Skiadas, C., Conditioning and Aggregation of Preferences, *Econometrica* **65** (1997), 347-367.