DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES
CALIFORNIA INSTITUTE OF TECHNOLOGY
PASADENA, CALIFORNIA 91125

Dynamic Consistency Implies Approximately Expected Utility Preferences

Kim C. Border

Uzi Segal

SOCIAL SCIENCE WORKING PAPER 821

November 1992

Dynamic Consistency Implies Approximately Expected Utility Preferences

Kim C. Border Uzi Segal

Abstract

Machina has proposed a definition of dynamic consistency which admits non-expected utility functionals. We show that even under this new definition, a dynamically consistent preference relation that is differentiable becomes arbitrarily close to an expected utility preference after the realization of a low probability event.

Keywords: nonexpected utility; dynamic consitency; decision theory; expected utility

JEL Classification numbers: 026

Dynamic Consistency Implies Approximately Expected Utility Preferences*

Kim C. Border† Uzi Segal‡

1 Introduction

One of the strongest normative objections to violations of the independence axiom is that such behavior involves dynamically inconsistent decisions in compound lotteries. A compound lottery is a two-stage lottery in which the first stage uses a randomizing device to select a lottery to be played at the second stage. We use the notation $(X_1, p_1; \ldots; X_m, p_m)$ to denote the compound lottery that yields the simple lottery X_i with probability p_i. It is widely argued that if a decision maker's preferences violate the independence axiom, then he is vulnerable to Dutch books. For example, suppose $X \succ Y$, but for some Z and $\alpha \in (0,1)$, $(Y, \alpha; Z, 1-\alpha) \succ (X, \alpha; Z, 1-\alpha)$. Suppose the decision maker holds a ticket for the compound lottery $(X, \alpha; Z, 1-\alpha)$. Offer him the opportunity to switch, for a small amount of money (say ε), to the compound lottery $(Y, \alpha; Z, 1-\alpha)$, which he will accept. If the Z-event happens, let him play Z. However, if the Y-event happens, offer him, for another small amount of money (say ε'), the opportunity to trade Y for X. Since X is preferred to Y, he will accept this offer too. Eventually, he plays the compound lottery $(X - \varepsilon' - \varepsilon, \alpha; Z - \varepsilon, 1-\alpha)$, which is stochastically dominated by his original holding. (See Raiffa [26, Chapter 4.9], Green [11], and Machina [21].)

A similar argument against violations of expected utility theory may be found in Border [2] or Fishburn [10]. There a decision maker is asked to make contingent choices. If theses choices are not consistent with expected utility maximization of an increasing utility, then there is a compound lottery for which the decision maker's contingent choices are stochastically dominated. Thus, for this compound lottery, the decision maker is not willing to stand by his announced choices. His ex ante preferences are not the same as his ex post preferences.

*We thank Larry Epstein, Mark Machina, and Joe Ostroy for fruitful and stimulating discussions.

†Division of the Humanities and Social Sciences, 228-77, California Institute of Technology, Pasadena CA 91125

‡Department of Economics, University of Toronto and Division of the Humanities and Social Sciences, 228-77, California Institute of Technology, Pasadena CA 91125

Machina [21] (see also McClennen [23, 24]) argues that dynamic consistency should not require that the ex post preference relation has to be the same as the decision maker's ex ante preference. Rather, the ex post preference relation should be the same as the one the decision maker ex ante *plans* to use ex post. Only under expected utility do these two definitions coincide. Formally, the proposed definition of dynamic consistency is this. Suppose that ex ante the decision maker faces the compound lottery $(X_1, p_1; \ldots; X_m, p_m)$. If the outcome of the compound lottery is X_i, then the decision maker ex post weakly prefers Y to Z if and only if $(X_1, p_1; \ldots; Y, p_i; \ldots; X_m, p_m) \succcurlyeq (X_1, p_1; \ldots; Z, p_i; \ldots; X_m, p_m)$ ex ante.

In this framework, it is not a well posed question to ask a non-expected utility maximizer to make contingent choices without specifying the probabilities used, as the experience of the risk at the first stage affects the preferences of the decision maker in the second stage. The above definition allows ex post preferences to depend on the ex ante distribution of lotteries and the actual outcome of the first stage. We analyze below some of the implications of this definition, and reach some unexpected conclusions.

To illustrate our major result, consider a preference relation $\succcurlyeq$ that is *quadratic* in the probabilities (see Machina [20] and Chew, Epstein, and Segal [5]). That is, it can be represented by

$$Q(F_X) = \int\int \varphi(x, y)\, dF_X(x)\, dF_X(y)$$

for some symmetric function $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$, where F_X is the cumulative distribution function of X. For a finite lottery $X = (x_1, p_1; \ldots; x_n, p_n)$ we have

$$Q(F_X) = \sum_i \sum_j p_i p_j \varphi(x_i, x_j).$$

Let $Y = (y_1, q_1; \ldots; y_m, q_m)$ and consider the compound lottery $(Y, 1-\varepsilon; Z, \varepsilon)$. If the Z-event happens, then the updated preference relation can be represented by the function

$$W(\cdot) = Q\Big(\varepsilon F_{(\cdot)} + (1-\varepsilon) F_Y\Big).$$

For $X = (x_1, p_1; \ldots; x_n, p_n)$ we obtain

$$\begin{aligned} W(X) &= \varepsilon^2 Q(F_X) + (1-\varepsilon)^2 Q(F_Y) + 2\varepsilon(1-\varepsilon) \sum_i \sum_j p_i q_j \varphi(x_i, y_j) \\ &= (1-\varepsilon)^2 Q(F_Y) + \varepsilon\Big[\varepsilon Q(F_X) + 2(1-\varepsilon) \sum_i \sum_j p_i q_j \varphi(x_i, y_j)\Big]. \end{aligned}$$

This function is ordinally equivalent to

$$W'(X) = \frac{\varepsilon}{2} Q(F_X) + (1-\varepsilon) \sum_i \sum_j p_i q_j \varphi(x_i, y_j).$$

Now consider what happens to this as ε becomes small. That is, what happens to ex post preferences after a low probability event has been realized. It is easy to see that for our quadratic case,

$$W'(X) \xrightarrow[\varepsilon \to 0]{} \sum_i \sum_j p_i q_j \varphi(x_i, y_j). \tag{1}$$

Let $w(x) = \sum_j q_j \varphi(x, y_j)$ and the expression in (1) becomes

$$W'(X) = \sum_i p_i w(x_i),$$

which is an expected utility functional.

This is not a peculiarity of the quadratic preference functional. It happens for all non-expected preferences that are representable by a differentiable functional, and that satisfy a very weak regularity condition, which we introduce in Section 3. Nonetheless, most alternatives to expected utility described in the literature satisfy our hypotheses. We show that if ex ante preferences are continuously differentiable, then dynamic consistency implies that ex post preferences after a low probability event occurs are close to expected utility preferences.

Figure 1 gives a hint of this last claim. Consider a set of lotteries represented as a subset of $\mathbb{R}^2$. Let X, Y, and Z be three lotteries. Let $\succcurlyeq$ be represented by V and define the "local" preference relation $\succcurlyeq_{Z,\varepsilon}$ by $X \succcurlyeq_{Z,\varepsilon} Y$ if and only if $V((1-\varepsilon)Z + \varepsilon X) \geq V((1-\varepsilon)Z + \varepsilon Y)$. For ε sufficiently close to 0, $X \succ_{Z,\varepsilon} Y$ if and only if the segment $\overline{cd}$ joining $(1-\varepsilon)Z+\varepsilon X$ and $(1-\varepsilon)Z+\varepsilon Y$ is steeper than the tangent to the indifference curve at Z. For any T and $\alpha \in (0,1]$, the segment $\overline{ab}$ connecting $\alpha X+(1-\alpha)T$ and $\alpha Y+(1-\alpha)T$

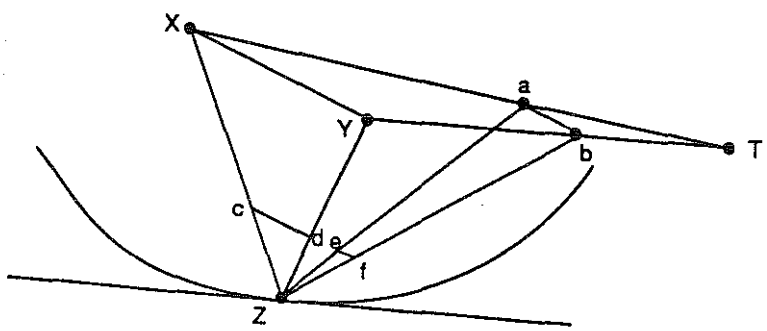

Figure 1: Why local preferences obey the independence axiom.

is parallel to $\overline{XY}$. Likewise, the segment $\overline{ef}$ connecting $(1-\varepsilon)Z + \varepsilon[\alpha X + (1-\alpha)T]$ and $(1-\varepsilon)Z + \varepsilon[\alpha Y + (1-\alpha)T]$ is parallel to $\overline{ab}$. Therefore, defining $\succcurlyeq_Z$ to be the limit of the orders $\succcurlyeq_{Z,\varepsilon}$ it follows that $X \succcurlyeq_Z Y$ if and only if $\alpha X + (1-\alpha)T \succcurlyeq_Z \alpha Y + (1-\alpha)T$. In other words, the order $\succcurlyeq_Z$ satisfies the Independence Axiom, and can therefore be represented by an expected utility functional.

Algebraically what is happening is this. If ex ante preferences are represented by a numerical function V, then the ex post preferences $\succsim_{Z,\varepsilon}$ are represented by $U_\varepsilon(X) = V((1-\varepsilon)Z + \varepsilon X)$. This is ordinally equivalent to

$$W_\varepsilon(X) = \frac{V((1-\varepsilon)Z + \varepsilon X) - V(Z)}{\varepsilon},$$

which converges to the derivative of V at Z, as $\varepsilon \to 0$. The derivative is a linear functional, and linearity corresponds to expected utility preferences. Note that the U_ε's themselves need not converge to the derivative.

The reason that this argument is not a complete proof is that there is a difference between convergence of utilities and the preferences that they represent. Consider, for example, the family of utility functions over the unit interval $[0,1]$ given by $u_n(x) = x/n$. All these functions represent the same order on $[0,1]$. The limit of the sequence of the utility functions is $u_0(x) = 0$, which represents a different relation. In order for the preferences to converge, the preference relation represented by the limit of the W_ε functions must be *locally strict* (see Section 3).

We also propose an extension to allow zero probability events. If the first stage induces a nonatomic distribution of lotteries, so that each outcome has zero probability, then the definition in [21] induces ex post preferences that are completely indifferent among all lotteries. Indeed, the decision maker does not care ex ante what outcome he receives if a zero-probability event happens. (Expected utility too has difficulty conditioning on zero probability events.) By the above definition of dynamic consistency, this implies that ex post he is indifferent between all possible outcomes if this zero-probability event actually happens. If all possible events have zero probability, then ex post the decision maker will be indifferent between all simple lotteries. We suggest to remedy this by defining ex post preferences after a zero probability event as the limit of ex post preferences after low probability events. It follows that under this definition, ex post preferences after a zero probability event are always linear.

Besides implying approximate linearity of preferences, dynamic consistency may imply discontinuity with respect to history. In Section 5, we discuss three formal definitions designed to formalize the notion of dynamic consistency and show that the only one that is immune to manipulation implies that ex post preferences must be even closer to expected utility.

2 Compound lotteries and dynamic consistency

The set of *prizes* is a compact interval $[a,b]$ of the real line, with $a < b$. The set $\mathcal{L}$ of *lotteries* is the set of (countably additive) Borel probabilities on $[a,b]$. It is endowed with the topology of weak convergence. That is, associating to each probability X its cumulative distribution function F_X, where $F_X(t) = X([a,t])$, $X_n \to X$ if and only if

F_{X_n} converges to F_X at every point of continuity of F_X. This topology is metrizable and compact.[1]

A *compound lottery* is a Borel measurable mapping π from $[0,1]$ into $\mathcal{L}$. We shall write π_s rather than $\pi(s)$ to simplify notation. The idea is that first an s is chosen at random from $[0,1]$ according to Lebesgue measure m. Then the lottery π_s is played. This corresponds to the tree diagrams familiar in the literature. The branches of a tree correspond to subsets of $[0,1]$ where π is constant.

Every compound lottery π can be *reduced* to a lottery Z by the reduction formula

$$Z(A) = \int_{[0,1]} \pi_s(A)\, dm(s).$$

This is well defined for every Borel set $A \subset [a,b]$ and does indeed define a countably additive probability. We can write Z as the vector integral $\int_{[0,1]} \pi_s\, dm(s)$ and refer to Z as the *reduction* of π. We shall call π a *finite compound lottery* if π takes on only finitely many values in $\mathcal{L}$ and if each of these values is assigned strictly positive probability by m. Note that this does *not* imply that the reduction of π has finite support. For instance, π_s may be the uniform probability on $[a,b]$ for every s. In this case, π assumes only one value, the uniform probability, but its reduction has full support on $[a,b]$.

We now assume that a function $V: \mathcal{L} \to \mathbb{R}$ represents the ex ante preferences over lotteries, and further that compound lotteries are ranked ex ante according to the value of their reduction under V. This assumption is known as the Reduction of Compound Lotteries Axiom. Thus we shall abuse notation and write $V(\pi)$ for the value V assigns to the reduction of π. Given a compound lottery π, once a state s has been drawn, preferences over $\mathcal{L}$ are represented by an ex post utility function $W: \mathcal{L} \to \mathbb{R}$.

Machina [21] argues that these ex post preferences can depend on the risks born and the outcome at the first stage. That is, W must be indexed by π to represent the risks born and by π_s to denote the outcome of the first stage. Thus we shall use the somewhat cumbersome notation $W(\cdot\,|\pi, X)$ to denote the ex post function used to evaluate lotteries in the second stage of the compound lottery π, after drawing an s at the first stage with $\pi_s = X$.

There may be different ways in which the decision maker can use information regarding risk previously born. For example, he may condition his behavior on what event happened, what state of the world happened, what is the outcome of previous stages (i.e., what actual lottery he now holds), or any combination of the above. We start with one such possible definition, and discuss its comparative merits, together with some other definitions, in Section 5 below.

Given a compound lottery π and a measure X in the range of π, let $R(\pi \mid X \to Z)$

[1] See, e.g., Parthasarathy [25], for these standard results.

denote the function from $[0,1]$ into $\mathcal{L}$ defined by replacing X everywhere by Z. That is,

$$R(\pi \mid X \rightarrow Z)_s = \begin{cases} \pi_s & \pi_s \neq X \\ Z & \pi_s = X. \end{cases}$$

In the spirit of Machina we make the following definition. Its relationship to Machina's actual definition is discussed in Section 5.

1 Definition *A pair (V, W) of ex ante-ex post utility functions is* outcomewise dynamically consistent *if for every finite compound lottery π, and every X in the range of π, $W(Y \mid \pi, X) \geq W(Y' \mid \pi, X)$ if and only if $V(R(\pi \mid X \rightarrow Y)) \geq V(R(\pi \mid X \rightarrow Y'))$ for every $Y, Y' \in \mathcal{L}$.*

That is, the ex post utility of a lottery Y, after drawing X at the first stage of the compound lottery π, agrees with the ex ante utility of the compound lottery derived from π by replacing X everywhere with Y.

3 Differentiable preferences

In our discussion above, we took the limit of a sequence of preference orders. The standard way of taking such a limit is to use the Hausdorff metric[2] on the graphs of the preference relations. This notion of convergence was introduced by Kannai [13] and Mas-Colell [22] and it has the property that if $X_n \rightarrow X$, $Y_n \rightarrow Y$, $\succcurlyeq_n \rightarrow \succcurlyeq$, and $X_n \succcurlyeq_n Y_n$, then $X \succcurlyeq Y$. Further, if $\succcurlyeq_n \rightarrow \succcurlyeq$ and $X \succ Y$, then for large enough n, $X \succ_n Y$. If utility representations converge uniformly, and if the limiting utility is locally strict (we define this in the next paragraph), then the preferences they represent converge in this sense. See Appendix B for details.

Define a preference $\succcurlyeq$ to be *locally strict* if for every (x, y) with $x \succcurlyeq y$, every neighborhood of (x, y) contains a pair (x', y') with $x' \succ y'$. This condition rules out thick indifference sets, but is weaker than local nonsatiation, as it allows for satiation points. This is important because $\mathcal{L}$ is compact, so every continuous preference ordering has a satiation point. We shall say that a utility is locally strict if it represents a locally strict preference. Strict monotonicity with respect to stochastic dominance ensures that preferences are locally strict.

We shall restrict our attention to preferences having a differentiable utility. To define differentiability we shall embed $\mathcal{L}$ in a normed vector space. Again identifying each

[2]The Hausdorff metric on closed subsets of a compact metric space is defined by

$$\rho(E, F) = \inf\{\varepsilon > 0 : E \subset N_\varepsilon(F) \text{ and } F \subset N_\varepsilon(E)\},$$

where $N_\varepsilon(E) = \{x : \exists y_{y \in E}\ d(x, y) < \varepsilon\}$, the ε-neighborhood of E. See Appendix B for more on this topology.

probability X with its cumulative distribution function F_X, we can embed $\mathcal{L}$ in the vector space $L_p[a,b]$ of p-integrable functions on $[a,b]$ for some $1 \leq p < \infty$. We point out here that this identification preserves the mixture space operations on $\mathcal{L}$. That is, $\lambda F_X + (1-\lambda) F_Z$ is the cumulative distribution function of $\lambda X + (1-\lambda) Z$. Wang [29] shows that the L_p norm on all of these spaces induces the weak topology on $\mathcal{L}$. Machina [20] uses the L_1-norm, and Allen [1] advocates the L_2-norm. Some preferences, in particular, anticipated utility preferences, have L_2-differentiable utilities, but not L_1-differentiable utilities [6]. Our results only depend on differentiability with respect to *some* norm, the actual norm does not matter.

A function V is differentiable on $\mathcal{L}$ if at each point $Z \in \mathcal{L}$ there is a linear function $DV_Z: B \to \mathbb{R}$ with the following property.

For every $\varepsilon > 0$, there is a $\delta > 0$ such that if $\|X\| < \delta$, then

$$|V(Z+X) - V(Z) - DV_Z(X)| < \varepsilon \|X\|.$$

This is the standard definition of (Fréchet) differentiability, written in a more convenient form. If such a linear function exists, then it is unique and continuous. It is important to remember that the derivative is a linear function from L_p to $\mathbb{R}$, not an element of $\mathbb{R}$.

Machina [20] shows that L_1-differentiable utilities possess *local utilities*. That is, the derivative DV_Z of V at Z has a representation of the form $DV_Z(X) = \int_{[a,b]} u_Z(t)\, dX(t)$ for a function $u_Z: [a,b] \to \mathbb{R}$. The function u_Z is called the local utility at Z. If u_Z is strictly increasing, then the preferences represented by DV_Z are locally strict. Wang [29] generalizes the results of Machina and Allen [1] to all the L_p-spaces, $1 \leq p < \infty$. The choice of norm does not affect our proofs, all we require is differentiability in some norm for which the derivative can be represented as a local utility. The preferences represented by the derivative DV are locally strict if the local utility representing DV is strictly increasing.

Our major result is that under outcomewise dynamic consistency, after a low probability event, ex post preferences must be close to expected utility preferences. The following theorem makes this result precise.

2 Theorem *Suppose V is continuously differentiable with DV locally strict everywhere, and suppose (V, W) is dynamically consistent. Let $\varepsilon > 0$ be given. Then there is some $\delta > 0$ such that for any X, if π is any compound lottery assigning X probability less than δ, then there is a linear preference $\succcurlyeq$ such that the Hausdorff distance between $\succcurlyeq$ and the preferences represented by $W(\cdot \mid \pi, X)$ is less than ε.*

Proof: Let $\succcurlyeq_{Z,\lambda}$ denote the preferences represented by the utility $U(X) = V((1-\varepsilon)Z + \varepsilon X)$, and let $\succcurlyeq_X$ denote the (linear) preferences represented by the local utility at X.

Lemma 6 and Theorem 8 in Appendix A show that if $Z \to X$ and $\lambda \downarrow 0$, then $\succcurlyeq_{Z,\lambda} \to \succcurlyeq_X$ in the Hausdorff metric.

Now let $\varepsilon > 0$ be given. For every $X \in \mathcal{L}$, there is a $\delta(X) > 0$ such that if $\|Z - X\| < \delta(X)$ and $0 < \lambda < \delta(X)$, then the Hausdorff distance $\rho(\succcurlyeq_{Z,\lambda}, \succcurlyeq_X) < \varepsilon$. The collection $\{N_{\delta(X)}(X) : X \in \mathcal{L}\}$ of balls of radius $\delta(X)$ around X is an open cover of $\mathcal{L}$. Since $\mathcal{L}$ is compact, there is a subcover corresponding to the finite set $\{X_1, \ldots, X_n\}$. Set $\delta = \min_i \delta(X_i)$. Then for every $Z \in \mathcal{L}$, there is some $X_i \in \{X_1, \ldots, X_n\}$ with $\|Z - X_i\| < \delta$. Furthermore, if for $0 < \lambda < \delta$, we have $\rho(\succcurlyeq_{Z,\lambda}, \succcurlyeq_X) < \varepsilon$. ∎

This theorem says that for any degree of approximation, there is a low enough probability so that ex post preferences after the realization of an event of this low probability are uniformly approximately linear. Lotteries with low probability outcomes are not unusual. Indeed the case where each of the prizes occurs with small probability is "typical" in a topological sense. Let $\mathcal{L}_\varepsilon = \left\{X \in \mathcal{L} : \exists c_{c \in [a,b]}\ X(\{c\}) \geq \varepsilon\right\}$, the set of lotteries where some prize receives probability at least ε. Then $\mathcal{L}_\varepsilon$ is closed and has an empty interior (cf. Parthasarathy [25, Theorem 8.1, p. 53]). That is, its complement is open and dense. This means that $\mathcal{L}_\varepsilon$ is topologically a small subset of $\mathcal{L}$. In this sense "most" lotteries involve small probability events.

4 Zero probability events

Dynamic consistency runs into difficulty when π is not finite. If X occurs with zero probability, dynamic consistency implies that $W(\cdot \mid \pi, X)$ must be indifferent among all lotteries. It is quite easy to construct compound lotteries in which each elementary lottery occurs with probability zero. In this case, ex post preferences are flat, no matter what happens at the first stage. In order to circumvent this difficulty, we propose to extend the notion of outcomewise dynamic consistency to more general compound lotteries along the following lines. If X in the range of π is assigned zero probability by m, we want to look at the ex post preferences induced by a compound lottery derived from π, but which assigns probability ε to X. We then let ε decrease to zero, and take the limit of these preferences (in the topology of closed convergence). We view this as being in the same spirit as Kreps and Wilson's [17] consistency condition for beliefs in sequential equilibrium.

3 Definition *A pair (V, W) of ex ante-ex post utilities is* outcomewise strongly dynamically consistent *if it is dynamically consistent in the sense of Definition 1 and if X in the range of compound lottery π has probability zero under m, then $W(\cdot \mid \pi, X)$ represents $\succcurlyeq_Z$, which is defined as follows. Let Z be the reduction of π, and let the preference $\succcurlyeq_{Z,\varepsilon}$ on $\mathcal{L}$ to be represented by the utility function*

$$U_{Z,\varepsilon}(\cdot) = V((1-\varepsilon)Z + \varepsilon\cdot).$$

If the limit as $\varepsilon \downarrow 0$ of $\succcurlyeq_{Z,\varepsilon}$ exists, call it $\succcurlyeq_Z$.

We first show that the limit in Definition 3 exists for differentiable utilities.

4 Theorem *Suppose π has reduction Z, lottery X in the range of π has probability zero under m, the utility V is differentiable at Z and DV_Z is locally strict, and (V, W) is outcomewise strongly dynamically consistent. Then the limit as $\varepsilon \downarrow 0$ of $\succcurlyeq_{Z,\varepsilon}$ exists and is represented by the derivative DV_Z of V at Z.*

Proof: Note that $\succcurlyeq_{Z,\varepsilon}$ is represented by

$$\frac{V((1-\varepsilon)Z + \varepsilon \cdot) - V(Z)}{\varepsilon},$$

which converges uniformly on $\mathcal{L}$ to $DV_Z(\cdot - Z)$. Theorem 8 implies that the limit $\succcurlyeq_Z$ exists and is represented by the utility $W(Y) = DV_Z(Y - Z)$. By linearity, $DV_Z(Y)$ is a positive affine transformation of $DV_Z(Y - Z)$ and so represents $\succcurlyeq_Z$. ∎

Note well what this says. The derivative DV_Z is a linear function. That is, it satisfies the independence axiom, and so is an expected utility ordering. Thus for nonatomic compound lotteries, those assigning probability zero to every lottery in their range, all the ex post preferences, as defined in Definition 3 are expected utility orderings.

5 Manipulation and continuity with respect to history

The idea that dynamically consistent preferences depend on previously born risk lends itself to (at least) three formal definitions. These definitions differ in the information that the preferences may be conditioned on: With respect to outcomes (Definition 1 above), or with respect to arbitrary events (Definition 5 below), or with respect to "branch" of the decision tree (also decribed below). In this section we show that under the first definition, the ex post preferences may not be continuous with respect to previous uncertainty.[3] It turns out that the alternative definition may expose the decision maker to manipulations. The third definition (Machina's [21] definition) solves both the discontinuity and Dutch books problems. However, in view of our major result (Theorem 2), this last definition implies that next period's preferences must be even closer to expected utility than under Definition 1.

[3]There are two ways we can make precise the notion of convergence for compound lotteries. The first is just convergence in distribution. That is, $\pi_n \to \pi$ if and only if $\int_{\mathcal{L}} \varphi(X)\, d(m\pi_n^{-1})(X) \to \int_{\mathcal{L}} \varphi(X)\, d(m\pi^{-1})(X)$ for every continuous real function φ on $\mathcal{L}$. The second notion of convergence is that of convergence in measure. That is, $\pi_n \to \pi$ if and only if for every $\varepsilon > 0$, $\lim_{n\to\infty} m\{s : \|\pi_{n_s} - \pi_s\| > \varepsilon\} = 0$. Our analysis can use either notion.

Consider the sequence of compound lotteries $\pi_n = (X, p; Y_n, q; Z, 1-p-q)$, $n = 0, \ldots \infty$, where $Y_n \to Y_0$. Suppose first that $Y_0 \neq X$. The ex post preferences, given that the decision maker wins X, is given by

$$\begin{aligned} W(X' \mid \pi_n, X) &= V(R(\pi_n \mid X \to X')) \\ &= V(X', p; Y_n, q; Z, 1-p-q). \end{aligned}$$

Since V is continuous,

$$\begin{aligned} V(X', p; Y_n, q; Z, 1-p-q) &\to V(X', p; Y_0, q; Z, 1-p-q) \\ &= W(X' \mid \pi_0, X). \end{aligned}$$

In other words, $W(X' \mid \pi_n, X) \to W(X' \mid \pi_0, X)$.

Consider however the case $Y_0 = X$ where for every $n \geq 1$, $Y_n \neq X$. Then

$$\begin{aligned} W(X' \mid \pi_n, X) &= V(X', p; Y_n, q; Z, 1-p-q) \\ &\to V(X', p; X, q; Z, 1-p-q) \\ &\neq V(X', p+q; Z, 1-p-q) \\ &= W(X' \mid \pi_0, X). \end{aligned}$$

Not only are $V(\cdot, p; X, q; Z, 1-p-q)$ and $V(\cdot, p+q; Z, 1-p-q)$ different in their values, but if the decision maker does not maximize a betweenness[4] (see [4, 7]) functional, then it may well happen that these two functions are not even *ordinally* equivalent. For example, it may happen that

$$(X', p+q; Z, 1-p-q) \succ (X, p+q; Z, 1-p-q) \succ (X', p; X, q; Z, 1-p-q),$$

hence, $W(X' \mid \pi_0, X) > W(X \mid \pi_0, X)$, but for every $n \geq 1$, $W(X \mid \pi_n, X) > W(X' \mid \pi_n, X)$. It thus follows that W (and the ex post preference relation) are not continuous in the ex ante options.

There is of course nothing wrong with the possibility of having preferences that are not continuous. However, it is obvious from the above analysis that non-betweenness preferences *must* have points of discontinuity with respect to born risk.

Some may argue that offers should be made conditional on events rather than outcomes. In other words, when the decision maker is offered the chance to replace a lottery X by Y, he assumes that this offer is made conditional on some past event (which led to his holding of X), and not on the fact that he holds X. In particular, it may be that this choice would have been offered to him even if he had received other lotteries in the first stage, or would not have been offered if some other event leading to X had occurred, even if the outcome in this case were X. Formally, for a compound lottery π, event E, and a lottery Z, let

$$R(\pi \mid E \to Z) = \begin{cases} Z & s \in E \\ \pi_s & s \notin E \end{cases}$$

[4]The preference relation $\succeq$ satisfies the betweenness axiom if $[X \succeq Y] \Rightarrow [\forall \alpha \in [0,1]\, X \succeq \alpha X + (1-\alpha) Y \succeq Y]$.

In other words, $R(\pi \mid E \to Z)$ is the compound lottery obtained from π by replacing the outcomes of π in case E happens by Z.

5 Definition *A pair* (V, W) *of ex ante-ex post utility functions is* eventwise dynamically consistent *if for every finite compound lottery* π*, lottery* Z *and event* E*,*

$$W(Z \mid \pi, E) = V(R(\pi \mid E \to Z)).$$

That is, the ex post utility of a lottery Z, after $s \in E$ happens, agrees with the ex ante utility of the compound lottery derived from π by replacing π_s by Z on E.

This kind of dynamic consistency exposes the decision maker to Dutch books. Suppose that the decision maker's non-expected utility preferences satisfy

$$(X, T; X, \neg T) \succ (Y, T; Y, \neg T) \succ (X, T; Y, \neg T)$$

where $\neg T$ is the event "not T" and $\Pr(T) \in (0, 1)$. Suppose the decision maker possesses the compound lottery $(X, T; Y, \neg T)$. In each state $s \in T$ he is offered the chance to replace the outcome conditional on T by Y. By the definition of eventwise dynamic consistency he is willing to pay a positive amount of money for this option. Next, offer him a chance to replace the outcome Y conditional on $T \cup \neg T$ by X. Again, he is willing to pay a positive sum for this offer. Eventually, he will play X, his state s outcome, but will pay twice for it. Note that this analysis does not require that anyone will be trying to cheat the decision maker. However, since we know that $s \in T$ happened, there is no way to check whether these offers are genuine or not.

It is important to note that this analysis does not depend in any way on "hidden nodes". In fact, even if the decision maker knows in advance what options will be offered to him after the first stage of the uncertainty is resolved, it still follows from Definition 5 that he will trade X for Y and then again for X.

There is one way in which Definition 5 does not lead to such Dutch books. If all offers must be restricted to the actual state s that happened, then such a manipulation as described above is impossible. But then, as we showed in Section 4, the updated preferences must *always* be expected utility.

This brings us to Machina's definition of dynamic consistency. We assume that the decision maker is endowed with a partition of states of the world that is at least as fine as the partition induced by π. The elements of this partition are called *branches*. Suppose the compound lottery π yielded lottery X, but the decision maker knows that branch E in the partition actually occurred. Then he should evaluate lotteries ex post by substituting only on E and not on all of $\pi^{-1}(X)$. This feature eliminates the discontinuity and manipulation discussed above, but makes the probability of the realized event even smaller, and so strengthens the conclusions about approximate linearity of ex post preferences.

6 Concluding remarks

In this paper we discussed some aspects of dynamic consistency. Our major claim is that if dynamic consistency means that the preference relation (over lotteries) in the next period is the same as the conditional preference relation today, then it will almost always be arbitrarily close to expected utility. One possible claim against our analysis is that although it is true that *next* period preferences are going to be close to linear, this puts no restrictions whatsoever on *today's* preferences. This is of course true, but it ignores the fact that today is yesterday's tomorrow. In other words, we cannot neglect problems in the future, because at each point of time we already are in the future of previous periods.

This observation is also relevant to another possible objection to our analysis. Theorem 2 shows that topologically, next period's preferences are almost always as close as we wish to expected utility. One may argue that this topological sense is irrelevant, because it assumes a certain kind of uniformity over this period's compound lotteries, whereas in the real world, most compound lotteries involve a small number of possible outcomes (an outcome by itself being a ticket for a lottery in the next period), each with a sufficiently large probability. In other words, even if they are topologically rare, compound lotteries where at least one of the outcomes is received with probability greater then ε (we denoted the set of these lotteries by $\mathcal{L}_\varepsilon$) are the typical case the decision maker faces today. This may be so. However, it is unrealistic to assume that there are only two periods in life. If we consider compound lotteries over n periods, then dynamic consistency is defined inductively, where the compound lotteries the decision maker faces at period $i+1$ are reduced by the reduction of compound lotteries axiom, and the preference relation at this period is required to be dynamically consistent with the preference relation at period i. Even if at each period there are few possible outcomes, each with sufficiently large probability (say larger than ε), the probability of each final outcome in an n-stage lottery tends to zero as n increases, unless from a certain stage on there is almost no uncertainty, and a certain branch of the tree is received with probability arbitrarily close to one.

It may seem as though Theorem 2 is relevant not for the preference relation over the whole set $\mathcal{L}$, but only for a small subset of it. More specifically, we are interested in the relation between X and Y given that probability $1-\lambda$ is assigned to the outcome Z. As λ decreases to zero, the domain of our discussion also shrinks to zero. This is technically true, but irrelevant. It is true that the *ex ante* domain goes down to zero. However, we are interested in the *ex post* relation (representable by W). The domain of this preference relation is $\mathcal{L}$, no matter how small $\lambda > 0$ is.

It would be wrong to conclude from this paper that dynamic consistency must imply expected utility behavior. An alternative approach to dynamic consistency under uncertainty is discussed in Karni and Safra [14, 15]. Another approach is suggested by Segal [27], where the preference relation is assumed to satisfy a compound independence axiom, but to violate the reduction of compound lotteries axiom. This approach is useful

in proving the existence of Nash equilibrium with non-expected utility preferences [8]. One may of course argue that violations of the reduction of compound lotteries axiom expose the decision maker to another kind of Dutch book. We discuss the validity of such arguments in [3] (see also [28]). If we are willing to assume both consequentialism and the reduction of compound lotteries axiom, Karni and Schmeidler [16] show that dynamic consistency is equivalent to expected utility. Finally, Epstein and Le Breton [9] discuss dynamic consistency in a Savage-like framework and raise question about the appeal of models of preferences that feature a separation of tastes and beliefs.

Appendix A: Results on Differentiability

Given a function V, define the difference quotient

$$\Delta_{Z,\lambda}(X) = \frac{V(Z+\lambda X) - V(Z)}{\lambda}.$$

An equivalent definition of differentiability of V at Z, is that as $\lambda \to 0$ the difference quotient $\Delta_{Z,\lambda}$ converges uniformly to DV_Z on norm bounded sets.

The function V is *continuously differentiable* if the mapping $Z \mapsto DV_Z$ is continuous, where the space of continuous linear functionals on L_p is topologized by the operator norm. That is, if V is continuously differentiable and if $Z_n \to Z$, then the linear function DV_{Z_n} converges uniformly to the linear function DV_Z on norm bounded subsets of L_p.

For continuously differentiable functions we have the following stronger result.

6 Lemma *Suppose $V: G \to \mathbb{R}$ is continuously differentiable on an open convex set G. Then the difference quotient $\Delta_{Z,\lambda}(h)$ converges to $DV_X(h)$ uniformly on norm bounded sets as $Z \to X$ and $|\lambda| \downarrow 0$. That is, for every $\varepsilon > 0$ and $M > 0$, for every $X \in G$, there is $\delta > 0$, such that*

$$|\Delta_{Z,\lambda}(h) - DV_X(h)| < \varepsilon,$$

whenever $\|h\| < M$, $0 < |\lambda| < \delta$, and $\|Z - X\| < \delta$.

Proof: This follows from [19, Corollary to 7.4, p. 149]. ∎

Appendix B: Utilities and preferences

It is usually more convenient to work with utilities than with preferences when discussing convergence. In this section we outline the general relation between convergence of utilities and convergence of preferences.

It is well known that if utilities converge uniformly on compact sets, the corresponding preferences converge in the topology of closed convergence, provided the limit preference is locally nonsatiated. Mas-Colell [22, 1.18, p. 313] proves this result for the case of convex preferences on Euclidean spaces, but does not use convexity in the proof. We cannot use this theorem however. The problem is that our preferences are not locally nonsatiated, as the lottery yielding b with probability one is a satiation point.

It is for this reason that we introduce the notion of locally strict preferences. Recall that $\succcurlyeq$ is locally strict if for every (x, y) with $x \succcurlyeq y$, in every neighborhood of (x, y) there is a pair (x', y') with $x' \succ y'$. This condition is weaker than local nonsatiation, as it allows for satiation points. It implies that the graph of $\succcurlyeq$ has no isolated points, but is stronger: The preference where all lotteries are indifferent has no isolated points, but is not locally strict. Local strictness exactly characterizes the continuity points of the mapping from utilities to preferences. For example, $U_n(x) = \frac{1}{n}x$ on the reals converges uniformly on compacta to zero, but the preferences remain strict until the limit. The preferences represented by the zero function make everything indifferent and are not locally strict. On the other hand, $U(x) = -|x|$ is locally strict, and if $U_n \to U$ uniformly on compacta, then the preferences do converge, even though U has a satiation point.

We only deal with preferences on $\mathcal{L}$, a compact metric space, but we state the next theorem more generally, since the proof is no harder. Recall that a topological space is *locally compact* if every point has a compact neighborhood. Every compact space is locally compact. We will need the following fact [18, Theorem 41.8, p. 44].

7 Theorem *If M is locally compact separable metrizable space, then there is a sequence $K_1, K_2, \ldots$ of compact sets satisfying $M = \bigcup_{n=1}^{\infty} K_n$ and $K_n \subset \operatorname{int} K_{n+1}$ for all n.*

Furthermore M is hemicompact, *that is, every compact subset K of M is contained in some K_n.*

Thus let (M, d) be a locally compact separable metric space, and let $K_1, K_2, \ldots$ be a sequence of compact subsets of M given by Theorem 7. The topology of *uniform convergence on compacta* on $C(M)$ is generated by the metric

$$d(f, g) = \sum_{n=1}^{\infty} \min\left\{\frac{1}{2^n}, \sup_{x \in K_n} d(f(x), g(x))\right\}.$$

(This topology is the same as the *compact-open* topology on $C(M)$.) A sequence $f_n \to f$ in this topology if and only if $f_n|_K \to f|_K$ uniformly on every compact subset K of M. If M is compact, then this is just the topology of uniform convergence on M. See Willard [30, Section 43] for these results.

For locally compact separable metric spaces, there is a generalization of the Hausdorff metric topology on the collection of closed subsets. It is called the *topology of closed convergence*. It is a compact metrizable topology and a sequence F_n of nonempty closed

sets converges to F in the topology of closed convergence if and only if $\mathrm{L}iF_n = F = \mathrm{L}sF_n$, where

$$\mathrm{L}sF_n = \Big\{x : \textit{for every neighborhood } G \textit{ of } x,\ \bigcap_{m=1}^{\infty} \bigcup_{n\geq m} (F_n \cap G) \neq \varnothing\Big\},$$

and

$$\mathrm{L}iF_n = \Big\{x : \textit{for every neighborhood } G \textit{ of } x,\ \bigcup_{m=1}^{\infty} \bigcap_{n\geq m} (F_n \cap G) \neq \varnothing\Big\}.$$

If M is compact, this topology is generated by the Hausdorff metric. See Hildenbrand [12, Section B.II, pp. 15–21] for these results. Clearly $\mathrm{L}iF_n \subset \mathrm{L}sF_n$ for any sequence. Thus to show convergence, we need only show that $\mathrm{L}sF_n \subset F \subset \mathrm{L}iF_n$. We mention here that since we will be considering subsets of $M \times M$, that if M is a locally compact space, so is $M \times M$ [30, Theorem 18.6, p. 131].

For a continuous function $U \in C(M)$, let $\succcurlyeq_U$ denote the preference induced by U. That is, $x \succcurlyeq_U y$ if and only if $U(x) \geq U(y)$. Define $\Psi\colon C(M) \to \mathcal{P}$ by $\Psi\colon U \mapsto \succcurlyeq_U$.

8 Theorem *Let M be a locally compact separable metric space. Let $\mathcal{P}$ denote the set of continuous preferences on M, topologized with the topology of closed convergence. Let $C(M)$ be topologized with the topology of uniform convergence on compacta. If $\succcurlyeq_U$ is locally strict, then Ψ is continuous at U. If M has no isolated points, and Ψ is continuous at U, then $\succcurlyeq_U$ is locally strict.*

Proof: We first show that if $\succcurlyeq_U$ is locally strict, then Ψ is continuous at U. So suppose $U_n \to U$, and write $\succcurlyeq_n$ for $\Psi(U_n)$. We wish to show that $\mathrm{L}s \succcurlyeq_n \subset \succcurlyeq_U \subset \mathrm{L}i \succcurlyeq_n$.

If $(x,y) \in \mathrm{L}s \succcurlyeq_n$. Then there is a sequence $\{(x_n, y_n)\}$ converging to (x,y), with $x_n \succcurlyeq_n y$. Thus $U_n(x_n) \geq U_n(y_n)$ for each n. Since U_n converges uniformly to U, Thus $U_n(x_n) \to U(x)$ and $U_n(y) \to U(y)$. Consequently $x \succcurlyeq_U y$, so $\mathrm{L}s \succcurlyeq_n \subset \succcurlyeq_U$.

Next suppose $x \succcurlyeq_U y$ and let G be a neighborhood of (x,y). Since $\succcurlyeq_U$ is locally strict, there is some $(x', y') \in G$ with $x' \succ y'$. Thus $U(x') > U(y')$. Since $U_n \to U$, for large enough n we have $U_n(x') > U_n(y')$. Thus $\succcurlyeq_n \cap G \neq \varnothing$ for large n. Therefore $(x,y) \in \mathrm{L}i \succcurlyeq_n$. That is, $\succcurlyeq_U \subset \mathrm{L}i \succcurlyeq_n$.

Therefore, $\succcurlyeq_n \to \succcurlyeq_U$ in the topology of closed convergence, and Ψ is continuous at U.

Now suppose M has no isolated points and $\succcurlyeq_U$ is not locally strict. We shall construct a sequence U_n converging to U that satisfies $\succcurlyeq_{U_n} \not\to \succcurlyeq_U$.

So suppose $x \succcurlyeq_U y$, but there is a neighborhood $N_x \times N_y$ of (x,y) with the property that, for all $y' \in N_y$ and all $x' \in N_x$, we have $y' \succcurlyeq_U x'$. Replacing y by $y' \neq y$ if necessary, we may assume $x \neq y$. (We can do this since y is not isolated.) We may also take then take N_x and N_y to be disjoint compact sets. By Urysohn's Lemma [30,

Theorem 15.6, p. 102], there is a continuous function f on M taking on values in $[0,1]$ with $f(N_x) = 1$ and $f(N_y) = 0$. Now set $U_n = U - \frac{1}{n}f$. Then U_n converges uniformly to U, and $U_n(x') < U_n(y')$ for each $x' \in N_x$ and $y' \in N_y$. Thus $N_x \times N_y$ contains no points in the graph of $\succcurlyeq_{U_n}$, and so (x,y) does not belong to $\mathrm{L}s \succcurlyeq_{U_n}$. That is, Ψ is not continuous at U. ∎

References

[1] Allen, B. 1987. Smooth preferences and the approximate expected utility hypothesis. *Journal of Economic Theory* 41(2):340–355.

[2] Border, K. C. 1992. Revealed preference, stochastic dominance, and the expected utility hypothesis. *Journal of Economic Theory* 56:20–42.

[3] Border, K. C. and U. Segal. 1991. Dutch book arguments and subjective probability. manuscript.

[4] Chew, S. H. 1983. A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox. *Econometrica* 51:1065–1092.

[5] Chew, S. H., L. Epstein, and U. Segal. 1991. Mixture symmetry utility functions. *Econometrica* 59:139–163.

[6] Chew, S. H., E. Karni, and Z. Safra. 1987. Risk aversion in the theory of expected utility with rank dependent preferences. *Journal of Economic Theory* 42:370–381.

[7] Dekel, E. 1986. An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom. *Journal of Economic Theory* 40:304–318.

[8] Dekel, E., Z. Safra, and U. Segal. 1991. Existence and dynamic consistency of Nash equilibrium with non-expected utility preferences. *Journal of Economic Theory* 55:229–246.

[9] Epstein, L. and M. Le Breton. 1992. Dynamically consistent beliefs must be Bayesian. manuscript.

[10] Fishburn, P. C. 1975. Separation theorems and expected utility. *Journal of Economic Theory* 11:16–34.

[11] Green, J. R. 1987. 'Making book against oneself,' the independence axiom and non-linear utility theory. *Quarterly Journal of Economics* 102(4):785–796.

[12] Hildenbrand, W. 1974. *Core and equilibria of a large economy*. Princeton: Princeton University Press.

[13] Kannai, Y. 1970. Continuity properties of the core of a market. *Econometrica* 38:791–815.

[14] Karni, E. and Z. Safra. 1989. Ascending bid auctions with behaviorally consistent bidders. *Annals of Operations Research* 19:435–446.

[15] Karni, E. and Z. Safra. 1989. Dynamic consistency, revelation in auctions, and the structure of preferences. *Review of Economic Studies* 56:421–434.

[16] Karni, E. and D. Schmeidler. 1991. Atemporal dynamic consistency and expected utility theory. *Journal of Economic Theory* 54:401–408.

[17] Kreps, D. M. and R. Wilson. 1982. Sequential equilibria. *Econometrica* 50:863–894.

[18] Kuratowski, K. 1968. *Topology*, volume 2. New York: Academic Press.

[19] Loomis, L. H. and S. Sternberg. 1968. *Advanced calculus.* Reading, Massachusetts: Addison–Wesley.

[20] Machina, M. J. 1982. 'Expected utility' analysis without the independence axiom. *Econometrica* 50(2):277–323.

[21] Machina, M. J. 1989. Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature* 27:1622–1668.

[22] Mas-Colell, A. 1974. Continuous and smooth consumers: Approximation theorems. *Journal of Economic Theory* 8:305–336.

[23] McClennen, E. F. 1988. Dynamic choice and rationality. In Munier, B., ed., *Risk, Decision, and Rationality.* Reidel, Dordrecht, Holland.

[24] McClennen, E. F. 1990. *Rationality and dynamic choice.* Cambridge: Cambridge University Press.

[25] Parthasarathy, K. R. 1967. *Probability measures on metric spaces.* Probability and Mathematical Statistics. New York: Academic Press.

[26] Raiffa, H. 1968. *Decision analysis: Introductory lectures on choice under uncertainty.* Reading, Massachusetts: Addison Wesley.

[27] Segal, U. 1990. Two-stage lotteries without the reduction axiom. *Econometrica* 58(2):349–377.

[28] Segal, U. 1992. The independence axiom versus the reduction axiom: Must we have both? In Edwards, W., ed., *Utility Theories: Measurements and Applications.* Kluwer, Boston.

[29] Wang, T. 1990. L_p-Fréchet differentiable preference and 'local utility' analysis. Manuscript, Department of Economics, Univerity of Toronto.

[30] Willard, S. 1970. *General topology.* Reading, Massachusetts: Addison Wesley.