Stochastic Dominance Analysis without the Independence Axiom *

Simone Cerreia–Vioglio, Fabio Maccheroni, Massimo Marinacci
Department of Decision Sciences and IGIER
Università Bocconi

June 23, 2015

Abstract

We characterize the consistency of a large class of nonexpected utility preferences (including mean-variance preferences and prospect theory preferences) with stochastic orders (for example, stochastic dominances of different degrees). Our characterization rests on a novel decision theoretic result that provides a behavioral interpretation of the set of all derivatives of the functional representing the decision maker's preferences.

As an illustration, we consider in some detail prospect theory and choice-acclimating preferences, two popular models of reference dependence under risk, and we show the incompatibility of loss aversion with prudence.

JEL classification: D81

Keywords: Stochastic dominance, integral stochastic orders, nonexpected utility, risk aversion, multi–utility representation, prospect theory, choice-acclimating personal equilibria

1 Introduction

When decision making under risk is considered, the assumption of expected utility maximization on the part of a single individual is descriptively controversial and several alternatives have been proposed. When the decision maker is not a single individual (for example, a firm) this assumption becomes less palatable also from a normative viewpoint (Diamond, 1967, Keeney, 1992, and Smith, 2004). At the same time, even if independence (the crucial behavioral assumption on which expected utility rests) is abandoned, consistency of the decision maker's preferences with stochastic dominances of different degrees remains of interest and practical relevance in the study of risk attitudes.[1] For example, the definition of (non-satiation and) risk aversion in terms of consistency with second degree stochastic dominance is commonly adopted also outside of the expected utility realm.

In decision analysis, also consistency with stochastic orders that do not belong to the stochastic dominance family naturally arises, as shown by the next example taken from Smith (2004).

*We thank Jim Smith (the Area Editor), an anonymous Associate Editor, three anonymous Referees, David Dillenberger, Edi Karni, and Mark Machina for very useful comments and suggestions. Simone Cerreia–Vioglio and Fabio Maccheroni gratefully acknowledge the financial support of MIUR (PRIN grant 20103S5RN3 005). Massimo Marinacci gratefully acknowledges the financial support of the AXA Research Fund. An earlier version of the paper was circulated under the title "A New Look at Local Expected Utility".

[1] See Levy (1992), for a survey of the applications of stochastic dominance in management science, operations research, and decision analysis.

Example 1 (The loyal firm – part I) *A firm has to choose among gambles that represent the payoffs of the firm's entire portfolio of investments, after paying employees, corporate taxes, debt holders, etc. Such gambles are formally described as probability distributions on an interval* $[m, M]$ *of monetary amounts, including* 0*. The firm's preferences can be represented by a continuous and smooth real-valued function* V *such that gamble* F *is weakly preferred to gamble* G*, denoted* $F \succsim G$*, if and only if* $V(F) \geq V(G)$*.*

The n *shareholders have expected utility preferences for gambles, that is, for every* $i = 1, ..., n$*, there exists a continuous function* $v_i : [m, M] \to \mathbb{R}$ *such that the expected utility*

$$V_i(F) = \int v_i(x)\, dF(x)$$

represents $\succsim_i$*.*

If the firm chooses gamble H*, then shareholder* i *gets the gamble* $H_i = H\left(s_i^{-1}\cdot\right)$*,*[2] *where* s_i *is* i*'s share, and his expected utility is*

$$V_i(H_i) = \int v_i(s_i x)\, dH(x).$$

A necessary requirement for the firm to make decisions in the best interests of the shareholders *is that whenever every shareholder prefers his share of* F *to his share of* G*, the firm also prefers* F *to* G*; this is equivalent to the consistency condition*

$$\left[\int u(x)\, dF(x) \geq \int u(x)\, dG(x) \qquad \forall u \in \mathcal{U}\right] \implies V(F) \geq V(G) \tag{1}$$

where $\mathcal{U} = \{v_i(s_i\cdot) : i = 1, ..., n\}$*. The relevant* stochastic order *at play is the unanimous judgment* $\succsim^{\mathcal{U}}$ *of the shareholders, formally defined by the left handside of (1), that is,*

$$F \succsim^{\mathcal{U}} G \text{ if and only if } \left[\int u(x)\, dF(x) \geq \int u(x)\, dG(x) \qquad \forall u \in \mathcal{U}\right].$$

With this notation, the consistency condition (1) becomes

$$F \succsim^{\mathcal{U}} G \implies V(F) \geq V(G). \tag{2}$$

In this paper, given any continuous and "smooth" functional V representing a decision maker's preferences $\succsim$ and any family $\mathcal{U}$ of continuous functions on $[m, M]$, we show that the consistency condition (2) is satisfied *if and only if* the set $\mathcal{V}$ of all derivatives of V, called *local utilities*, is included in the closed convex cone generated by $\mathcal{U}$ and all the constant functions (Proposition 1).[3] Assuming, without loss off generality, that $\mathcal{U}$ contains a strictly positive and a strictly negative constant function, this "closed convex cone condition" means that every local utility is either a weighted sum of elements of $\mathcal{U}$ or a limit of these weighted sums.[4]

Our differential characterization is made possible by a novel decision theoretic result (Theorem 1) which yields a behavioral interpretation of the set $\mathcal{V}$ of all local utilities by showing that it represents the largest subrelation $\succsim^*$ of $\succsim$ that satisfies the axioms of expected utility with the possible exception of completeness.

Next we elaborate on the two results. The seminal Machina (1982) showed that the global risk aversion analysis, classically carried out for expected utility preferences, naturally extends to a local risk aversion

[2] The distribution $H(s_i^{-1}\cdot)$ maps each x into $H(x/s_i)$.

[3] The name local utilities (introduced by Machina, 1982) is justified by the fact that these derivatives are continuous functions on $[m, M]$ such that their expectations can be used to locally approximate V. If v is a derivative of V at F, then

$$V(G) - V(F) \approx \int v(x)\, dG(x) - \int v(x)\, dF(x)$$

as G approaches F. See Section 2.1 for the formal definitions of smooth preference functional, local utilities, and so on.

[4] The assumption is without loss of generality because $\int u(x)\, dF(x) \geq \int u(x)\, dG(x)$ for all $u \in \mathcal{U}$ if and only if $\int u(x)\, dF(x) \geq \int u(x)\, dG(x)$ for all $u \in \mathcal{U} \cup \{c\}$ where c is any constant function. Therefore, for example, one can always replace $\mathcal{U}$ with $\mathcal{U} \cup \{1\} \cup \{-1\}$.

analysis when nonexpected utility preferences $\succsim$ can be represented by a smooth V. In the nonexpected utility realm the role of the "single global" utility of expected utility is taken by "multiple local" utilities. For example, they are all increasing and concave if and only if $\succsim$ is non-satiated and risk averse. Machina's results stimulated many studies of nonexpected utility preferences that heavily rely on local utilities (see Wang, 1993 and the references therein). As far as the consistency property (2) is concerned, this literature culminated with Chew and Nishimura (1992) where a sufficient, but obviously not necessary, condition is given: *if* the set $\mathcal{V}$ of all local utilities is included in $\mathcal{U}$, *then* property (2) is satisfied. At the same time, despite the recognized importance of the local approach, the global role of the set $\mathcal{V}$ of all local utilities – that is, its preferential underpinning – remained unexplained.

Our Theorem 1 explains the global role of the set $\mathcal{V}$ of all local utilities by characterizing an observable subrelation and computable $\succsim^*$ of $\succsim$ that is represented by $\mathcal{V}$ itself. In other words, the subrelation $\succsim^*$ is the counterpart of the set $\mathcal{V}$ in terms of choice behavior in the same way in which the preference $\succsim$ is the counterpart of V.

In turn, the "conceptual" Theorem 1 opens the way to the "operational" Proposition 1 that fully characterizes, through differential notions, property (2), that is, the consistency of a nonexpected utility preference with an integral stochastic order. In particular, Proposition 1 improves upon the sufficiency result of Chew and Nishimura by delivering a necessary and sufficient condition for consistency to hold.

In the expected utility case, in which $V(F) = \int v(x)\, dF(x)$ for some continuous v, Proposition 1 says that

$$F \succsim^{\mathcal{U}} G \implies \int v(x)\, dF(x) \geq \int v(x)\, dG(x)$$

if and only if v belongs to the closed convex cone generated by $\mathcal{U}$ and all the constant functions.[5] Thus, in general, Proposition 1 says that the stochastic dominance analysis carried out for expected utility preferences directly extends to smooth nonexpected utility preferences by replacing the utility function v with the set $\mathcal{V}$ of all local utilities, explaining the title of this paper.

Example 1 (The loyal firm – part II) *The discussed results imply that the firm makes decisions in the best interests of the shareholders only if the set $\mathcal{V}$ of all derivatives of V is included in the closed convex cone generated by $\mathcal{U} = \{v_i(s_i\cdot) : i = 1, ..., n\}$ and all the constant functions. The set $\mathcal{V}$ can be obtained by calculus techniques (see Remark 1), while the closed convex cone generated by $\mathcal{U}$ and all the constant functions is*

$$\langle \mathcal{U} \rangle = \{\lambda_1 v_1(s_1\cdot) + ... + \lambda_n v_n(s_n\cdot) + c : \lambda_1, ..., \lambda_n \geq 0 \text{ and } c \in \mathbb{R}\}$$

(Borwein and Moors, 2009). Therefore, the firm makes decisions in the best interests of the shareholders only if every local utility has the form

$$v(x) = \lambda_1 v_1(s_1 x) + ... + \lambda_n v_n(s_n x) + c \qquad \forall x \in [m, M]$$

for some set $\lambda_1, ..., \lambda_n$ of positive weights and $c \in \mathbb{R}$. Finally, since local utilities are unique up to an additive constant, this de facto *means that local utilities are weighted sums of the elements of $\mathcal{U}$.*

As detailed in the main text, some of our results do not require smoothness of the preference functional V and some extend to risky situations in which preferences are defined over probability distributions on metric spaces rather than on the interval $[m, M]$, thus allowing *inter alia* to deal with multiattribute consequences.

Finally, in order to illustrate the tractability of our approach, in Section 4 we discuss the case of risk aversion, whereas in Section 5 we use our results to study the relation between loss aversion and prudence in two popular models of reference dependence. Specifically, we compute the local utilities for prospect theory and we use them to show that the assumption of prudence is incompatible with loss aversion (Proposition 5);

[5] This is a known result a la Harsanyi (1955), see Fact 1 below.

moreover, when risk aversion and prudence are both assumed, prospect theory collapses to expected utility (Corollary 1). Finally, we show that the incompatibility of prudence and loss aversion also extends to the choice-acclimating personal equilibria of Koszegi and Rabin (2007) (Proposition 6).

2 Preliminaries

2.1 Derivatives and local utilities

Let $\mathcal{D} = \mathcal{D}(I)$ be the set of all cumulative distribution functions on a closed interval I of $\mathbb{R}$. When I is required to be bounded, we write $I = [m, M]$. We denote by F, G, and H generic elements of $\mathcal{D}$, and by x, y, and z generic elements of I. Given $F \in \mathcal{D}$, we denote by $\mathrm{E}(F)$ its expected value and by $\mathrm{Var}(F)$ its variance (when they exist). Given $x \in I$, we denote by G_x the distribution that yields x with probability 1. We endow $\mathcal{D}$ with the topology of weak convergence.[6]

We denote by $C_b(I)$ the set of all bounded and continuous functions on I. If I is bounded, then $C_b(I)$ coincides with the set $C(I)$ of all continuous functions on I.

Given a functional $V : \mathcal{D} \to \mathbb{R}$, we say that V is Gateaux differentiable at $F \in \mathcal{D}$ if and only if there exists a function $u_F \in C_b(I)$ such that for each $G \in \mathcal{D}$

$$\lim_{\theta \downarrow 0} \frac{V((1-\theta)F + \theta G) - V(F)}{\theta} = \int_I u_F(x)\, d(G - F)(x). \tag{3}$$

In this case, the Gateaux derivative u_F is called *local utility* for V at F. Notice that if u_F satisfies (3), so does $u_F + c$ for each $c \in \mathbb{R}$, and conversely if also w_F satisfies (3), then $u_F - w_F$ is constant. In other words, if V is Gateaux differentiable at F and u_F is a Gateaux derivative, then the set all Gateaux derivatives for V at F is $\nabla V(F) = \{u_F + c\}_{c \in \mathbb{R}}$.

The functional V is Gateaux differentiable (informally, *smooth*) if and only if it is Gateaux differentiable at each $F \in \mathcal{D}$, and we denote by $\nabla V : \mathcal{D} \rightrightarrows C_b(I)$ the derivative correspondence that maps each F into $\nabla V(F)$. Finally, we define

$$\operatorname{range} \nabla V = \bigcup_{F \in \mathcal{D}} \nabla V(F) = \{u_F : F \in \mathcal{D}\}.$$

The set $\operatorname{range} \nabla V$ is the collection of all local utilities of V, that we denoted by $\mathcal{V}$ in the introduction.

The next remark shows that although the definition of local utilities seems a bit involved, they can actually be computed as limits of difference quotients, exactly like standard calculus derivatives.

Remark 1 *The notion of Gateaux derivative that we use is due to von Mises (1947, p. 323). It has been widely used in Statistics since Hampel (1974) for the study of robustness (see, e.g., Fernholz, 1983, and Huber and Ronchetti, 2009). In Decision Theory, it was adopted by Chew, Karni, and Safra (1987).*

Specifically, Hampel and the subsequent statistical literature call the function $IC_{V,F} : I \to \mathbb{R}$ defined by

$$IC_{V,F}(x) = \lim_{\theta \downarrow 0} \frac{V((1-\theta)F + \theta G_x) - V(F)}{\theta} \qquad \forall x \in I \tag{4}$$

influence curve *of V at F. Now, if V is Gateaux differentiable at F and $u_F \in \nabla V(F)$, then $v_F = u_F - \int_I u_F(x)\, dF(x)$ is the only element of $\nabla V(F)$ with zero expectation with respect to F and by (3)*

$$\lim_{\theta \downarrow 0} \frac{V((1-\theta)F + \theta G) - V(F)}{\theta} = \int_I v_F(x)\, dG(x) \qquad \forall G \in \mathcal{D} \tag{5}$$

[6] See Appendix A for a formal definition of the topology of weak convergence and other technical details.

which for $G = G_{x_0}$ yields $IC_{V,F}(x_0) = v_F(x_0)$ for all $x_0 \in I$. In particular, (4) provides an explicit formula to compute the local utilities of a smooth functional V. For example, Hampel's computations show that, if $I = [m, M]$, the local utilities of the mean-variance preference functional

$$V(F) = \mathrm{E}(F) - \frac{\lambda}{2}\mathrm{Var}(F) \qquad \forall F \in \mathcal{D}([m, M])$$

of Markowitz (1952) and Tobin (1958), are given by

$$u_F(x) = x - \frac{\lambda}{2}(x - \mathrm{E}(F))^2 \qquad \forall x \in [m, M].$$

2.2 Preferences and stochastic orders

The object of our study is a binary relation $\succsim$ defined on $\mathcal{D}$ and describing the preferences of a decision maker. A functional $V : \mathcal{D} \to \mathbb{R}$ is said to represent $\succsim$, or to be a preference functional for $\succsim$, if and only if for every $F, G \in \mathcal{D}$

$$F \succsim G \Longleftrightarrow V(F) \geq V(G).$$

The axiomatic properties of $\succsim$ we discuss and use in this paper are few and classic. We next list them for completeness.

Preorder The relation $\succsim$ is reflexive and transitive.

Weak Order The relation $\succsim$ is complete and transitive.

Continuity For each pair of convergent sequences $\{F_n\}$ and $\{G_n\}$ in $\mathcal{D}$,

$$F_n \succsim G_n \quad \forall n \implies \lim_n F_n \succsim \lim_n G_n.$$

Independence For every $F, G, H \in \mathcal{D}$ and for each $\lambda \in (0, 1)$

$$F \succsim G \implies \lambda F + (1 - \lambda) H \succsim \lambda G + (1 - \lambda) H.$$

In analogy with the definition of linear orders in vector spaces, we call a binary relation $\succsim^{\#}$ on $\mathcal{D}$ satisfying Independence a *stochastic order*. Let us remark that the term "stochastic" here refers to the stochasticity of the elements of $\mathcal{D}$, the binary relation $\succsim^{\#}$ is itself deterministic (not stochastic like the ones arising from random choice). An important family of stochastic orders is the one of the so called *integral stochastic orders*, that is the binary relations $\succsim^{\mathcal{U}}$ of the form

$$F \succsim^{\mathcal{U}} G \iff \int_I u(x)\, dF(x) \geq \int_I u(x)\, dG(x) \qquad \forall u \in \mathcal{U}$$

where $\mathcal{U}$ is a given subset of $C_b(I)$.

Example 2 (Stochastic dominances and risk orders – part I) *Let $I = [m, M]$ and set*

$$\begin{aligned}
\mathcal{U}_1 &= \left\{u \in C^1([m, M]) : u' \geq 0\right\} \\
\mathcal{U}_2 &= \left\{u \in C^2([m, M]) : u' \geq 0 \text{ and } u'' \leq 0\right\} \\
\mathcal{U}_3 &= \left\{u \in C^3([m, M]) : u' \geq 0,\ u'' \leq 0, \text{ and } u''' \geq 0\right\} \\
&\ \ ... \\
\mathcal{U}_n &= \left\{u \in C^n([m, M]) : u^{(i)} \geq 0 \text{ if } i \text{ is odd and } u^{(i)} \leq 0 \text{ if } i \text{ is even, for } i = 1, ..., n\right\}.
\end{aligned}$$

The integral stochastic order $\succsim^{\mathcal{U}_1}$ is called first degree stochastic dominance*, $\succsim^{\mathcal{U}_2}$ is called* second degree stochastic dominance*, $\succsim^{\mathcal{U}_3}$ is called* third degree stochastic dominance*, ..., $\succsim^{\mathcal{U}_n}$ is called* n-th degree stochastic dominance*. While the integral stochastic order $\succsim^{\mathcal{R}_n}$ induced by*

$$\mathcal{R}_n = \left\{u \in C^n([m, M]) : (-1)^{n-1} u^{(n)} \geq 0\right\}$$

is called n-th degree risk order*, for every* $n \in \mathbb{N}$. *The first degree risk order coincides with first degree stochastic dominance and consistency with it amounts to* non-satiation*, the second relates to* risk aversion*, the third to* prudence*, the fourth to* temperance *(Eeckhoudt, Schlesinger, and Tselin, 2009), the fifth to* edginess *(Lajeri-Chaherli, 2004). The relation between stochastic dominances and risk orders resides in the observation that* $\mathcal{U}_n = \bigcap_{i=1}^{n} \mathcal{R}_i$.

The second degree risk order is commonly called concave order*.*

The following property of integral stochastic orders (which have been rediscovered many times, see, e.g., Muller, 1997, Castagnoli and Maccheroni, 1998, Dubra, Maccheroni, and Ok, 2004) will be very useful in the sequel. As a matter of notation, given a set $\mathcal{U} \subseteq C_b(I)$, we denote by $\langle \mathcal{U} \rangle$ the *C5 hull* of $\mathcal{U}$ (see Smith and McCardle, 2002), that is, the weak closure of the smallest convex cone containing $\mathcal{U}$ and all the constant functions.[7]

Fact 1 *Let* $\mathcal{U}, \mathcal{V} \subseteq C_b(I)$. *The following statements are equivalent:*

(i) $\succsim^{\mathcal{V}}$ *is consistent with* $\succsim^{\mathcal{U}}$, *that is,* $F \succsim^{\mathcal{U}} G$ *implies* $F \succsim^{\mathcal{V}} G$*;*

(ii) $\mathcal{V} \subseteq \langle \mathcal{U} \rangle$.

In particular, $\succsim^{\mathcal{V}}$ *coincides with* $\succsim^{\mathcal{U}}$ *if and only if* $\langle \mathcal{V} \rangle = \langle \mathcal{U} \rangle$.

In order to make this fact operational, it is important to be able to explicitly describe $\langle \mathcal{U} \rangle$. As observed in Example 1, if $I = [m, M]$ and $\mathcal{U} = \{u_i : i = 1, ..., n\}$ is finite, then $\langle \mathcal{U} \rangle = \{\lambda_1 u_1 + ... + \lambda_n u_n + c : \lambda_1, ..., \lambda_n \geq 0 \text{ and } c \in \mathbb{R}\}$. Next we provide an explicit description of the cones $\langle \mathcal{U}_n \rangle$ and $\langle \mathcal{R}_n \rangle$ that generate the n-th degree stochastic dominance and the n-th degree risk order for every $n \in \mathbb{N}$.[8]

Fact 2 *Let* $I = [m, M]$. *Then, under the usual convention that* $u^{(0)} = u$,

$$
\begin{aligned}
\langle \mathcal{R}_1 \rangle &= \{u \in C([m, M]) : u \text{ is increasing}\} \\
\langle \mathcal{R}_n \rangle &= \left\{u \in C([m, M]) : u^{(n-2)} \text{ exists and is concave on } (m, M)\right\} \quad \text{if } n \geq 2 \text{ is even} \\
\langle \mathcal{R}_n \rangle &= \left\{u \in C([m, M]) : u^{(n-2)} \text{ exists and is convex on } (m, M)\right\} \quad \text{if } n \geq 3 \text{ is odd}
\end{aligned}
$$

and

$$\langle \mathcal{U}_n \rangle = \bigcap_{i=1}^{n} \langle \mathcal{R}_i \rangle$$

for every $n \in \mathbb{N}$.

Throughout the paper, we will consider binary relations that can be represented by a continuous preference functional V. It is well known (see Debreu, 1964) that, in our setting, this is equivalent to assuming that $\succsim$ satisfies Weak Order and Continuity. Coupled with our results, Fact 2 will allow us to fully characterize the consistency of these preferences with stochastic dominances and risk orders of all degrees.

3 Main results

3.1 The expected utility core

In this section, we characterize the part of the decision maker's preferences that satisfies the expected utility hypothesis and describe its properties.

[7] When I is bounded this closure coincides with the supnorm closure.

[8] Also this fact is essentially folklore, for example, it can be derived from the pioneering works of Popoviciu (1933) and Karlin and Novikoff (1963). A proof is available upon request.

Definition 1 *The* expected utility core *of a binary relation* $\succsim$ *on* $\mathcal{D}$ *is the binary relation defined by*

$$F \succsim^* G \iff \lambda F + (1-\lambda) H \succsim \lambda G + (1-\lambda) H \qquad \forall \lambda \in (0,1], \forall H \in \mathcal{D}.$$

Clearly, $\succsim^*$ is a subrelation of $\succsim$. We interpret this derived binary relation as capturing the rankings for which the decision maker is sure. For, no matter how F and G are mixed with a third prospect H, the mixture of F with H dominates the one of G with H.

The expected utility core is the risk counterpart of the revealed unambiguous preference relation introduced by Ghirardato, Maccheroni, and Marinacci (2004) in a setting of decision making under ambiguity and for invariant biseparable preferences $\succsim$ (see also Nehring, 2009).[9] In a context of choice under risk, $\succsim^*$ was first studied by Cerreia–Vioglio (2009) for convex preferences $\succsim$ in order to reveal the subjective states of the decision maker a la Kreps (1979). It also plays a central role in Cerreia–Vioglio, Dillenberger, and Ortoleva (2015) where it characterizes the decision maker's indecisiveness. The following lemma lists some relevant properties of the expected utility core.

Lemma 1 *Let* $\succsim$ *be a binary relation represented by a continuous preference functional* V*. The following statements are true:*

(i) $\succsim^*$ *is a preorder that satisfies Continuity and Independence;*

(ii) $\succsim$ *is consistent with* $\succsim^*$*, that is,* $F \succsim^* G$ *implies* $F \succsim G$*;*

(iii) $\succsim$ *is consistent with a stochastic order* $\succsim^{\#}$ *if and only if* $\succsim^*$ *is consistent with* $\succsim^{\#}$*;*

(iv) if I *is bounded, then there exists a set* $\mathcal{U}^* \subseteq C(I)$ *such that*

$$F \succsim^* G \iff \int_I u(x)\, dF(x) \geq \int_I u(x)\, dG(x) \qquad \forall u \in \mathcal{U}^*; \tag{6}$$

(v) if I *is bounded, then* $\succsim$ *is consistent with an integral stochastic order* $\succsim^{\mathcal{U}}$ *if and only if* $\mathcal{U}^* \subseteq \langle \mathcal{U} \rangle$*.*

The first point shows that $\succsim^*$ satisfies all the expected utility axioms with the potential exception of completeness. The second point shows that $\succsim^*$ is a subrelation of $\succsim$, thus capturing a part of the rankings expressed by the decision maker. The third point implies that $\succsim^*$ is the largest subrelation of $\succsim$ that satisfies the expected utility axioms with the potential exception of completeness, thus supporting the interpretation that $\succsim^*$ summarizes the rankings for which the decision maker behaves like a standard expected utility agent. Point (iii) actually yields more, in fact it implies consistency of $\succsim^*$ with any subrelation of $\succsim$ that satisfies *just* independence. This is important in connection with the mean preserving spread relation (see Section 4). Point (iv) shows that, when I is bounded, $\succsim^*$ is an integral stochastic order (confirming its "expected utility nature"); also notice that, by Fact 1, both $\mathcal{U}^*$ and $\mathcal{W}^*$ represent $\succsim^*$ in the sense of (6) if and only if $\langle \mathcal{U}^* \rangle = \langle \mathcal{W}^* \rangle$.

But the most important point is (v) which derives a first characterization of the consistency of nonexpected utility preferences with integral stochastic orders. Such a characterization does not rely on differentiability properties of the representing functional, but rather on the possibility of identifying the set $\mathcal{U}^*$.[10]

[9] Cerreia-Vioglio, Ghirardato, Maccheroni, Marinacci, and Siniscalchi (2011) study $\succsim^*$ for the more general class of rational preferences $\succsim$. Several differential characterizations of $\succsim^*$ have been proposed. The first one can be found in Ghirardato, Maccheroni, and Marinacci (2004). A direct extension of this result appears in Ghirardato and Siniscalchi (2012), who develop in an ambiguity setup a local analysis close to the spirit of Machina (1982). Finally, Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2011) provide an alternative differential characterization and, inter alia, derive $\succsim^*$ for several ambiguity averse models.

[10] The proof of Proposition 6 will be an example of this fact.

In fact, there are cases in which a direct derivation of $\mathcal{U}^*$ is possible even when V is not smooth. A very important and popular case is the betweenness model of Dekel (1986). More recent models are the ones of Maccheroni (2002) and Cerreia–Vioglio, Dillenberger, and Ortoleva (2015) where preferences are represented as minima of expected utilities and certainty equivalents, respectively.

Example 2 (Stochastic dominances and risk orders – part II) *Let $I = [m, M]$ and $\succsim$ be a binary relation represented by a continuous preference functional V. Fact 2 and point (v) of Lemma 1 guarantee that $\succsim$ is consistent with:*

1. *first degree stochastic dominance (the first degree risk order) if and only if all elements of $\mathcal{U}^*$ are increasing;*
2. *second degree stochastic dominance if and only if all elements of $\mathcal{U}^*$ are increasing and concave;*
3. *third degree stochastic dominance if and only if all elements of $\mathcal{U}^*$ are increasing, concave, and have convex derivative on (m, M);*
4. *the n-th degree risk order (for $n \geq 2$) if and only if all elements of $\mathcal{U}^*$ have $(n-2)$-th derivative on (m, M) and such derivative is concave/convex if n is even/odd.*

The next Theorem 1 will show that, if V is smooth, one can choose $\mathcal{U}^* = \operatorname{range} \nabla V$ (even when I is unbounded).

Finally, as illustrated in Example 2 (part I) and Proposition 2 below, consistency with stochastic orders is a common way to model risk attitudes (see Eeckhoudt and Schlesinger, 2006, and Eeckhoudt, Schlesinger, and Tsetlin, 2009). Point (iii) of Lemma 1 shows that the expected utility core completely characterizes risk attitudes modelled through consistency with *any* stochastic order. In other words, if two preferences have the same expected utility core, then they must exhibit the same risk attitudes.

3.2 Differential characterizations

In the next theorem, our first main result, we show that the set of all local utilities represents the expected utility core of the decision maker's preferences.[11]

Theorem 1 *If $\succsim$ is a binary relation represented by a continuous and Gateaux differentiable preference functional V, then*

$$F \succsim^* G \iff \int_I u(x)\, dF(x) \geq \int_I u(x)\, dG(x) \qquad \forall u \in \operatorname{range} \nabla V.$$

Local utilities thus capture both local and global behavior that is consistent with expected utility. In particular, individually each of these utilities models a local expected utility behavior of $\succsim$, as Machina (1982) emphasized; while jointly they characterize a global expected utility feature of $\succsim$, as our result shows.

Theorem 1 can be restated by saying that the expected utility core is the integral stochastic order generated by all local utilities, that is,

$$\succsim^* = \succsim^{\operatorname{range} \nabla V}.$$

Together with the observation that $\succsim$ is consistent with a stochastic order if and only if $\succsim^*$ is consistent with it (Lemma 1), this leads to our second main result:

Proposition 1 *Let $\succsim$ be a binary relation represented by a continuous and Gateaux differentiable preference functional V. Then $\succsim$ is consistent with an integral stochastic order $\succsim^{\mathcal{U}}$ if and only if* $\operatorname{range} \nabla V \subseteq \langle \mathcal{U} \rangle$.

[11] Note that in Theorem 1 the interval I is not required to be bounded. Thus, we cannot rely on Dubra, Maccheroni, and Ok (2004) to represent $\succsim^*$, as shown by Evren (2008).

In particular, by taking $\mathcal{U}^* = \text{range}\,\nabla V$ in Example 2 (part II), we obtain a full differential characterization of consistency of $\succsim$ with stochastic dominances and risk orders of all degrees. Therefore this proposition extends (most of the extensions of) Theorem 1 of Machina (1982) including the very general Lemma 1 of Chew and Nishimura (1992). The latter essentially shows that if $\text{range}\,\nabla V \subseteq \mathcal{U}$, then $\succsim$ is consistent with $\succsim^{\mathcal{U}}$. The improvement of Proposition 1 is twofold. Not only our condition $\text{range}\,\nabla V \subseteq \langle \mathcal{U} \rangle$ is weaker, but more importantly it is both necessary and sufficient for consistency.[12]

4 Risk aversion

In the rest of the paper, to better compare our results with the literature and to avoid technicalities, we confine ourselves to the case $I = [m, M]$.

Outside the realm of expected utility, where they coincide, we have two competing notions of risk aversion: weak risk aversion and strong risk aversion (that is, aversion to mean preserving spreads). The second notion requires the definition of mean preserving spread (henceforth, MPS). We start by providing the more general notion of simple compensated spread, first introduced by Machina (1982), for a binary relation $\succsim$. The notion of MPS will be a particular case.

Given F and G in $\mathcal{D}$ and $\succsim$ on $\mathcal{D}$, we say that G is a simple compensated spread (henceforth, SCS) of F for $\succsim$ if and only if $G \sim F$ and there exists $z \in [m, M]$ such that

$$\begin{cases} F(x) \leq G(x) & \forall x \in [m, z] \\ F(x) \geq G(x) & \forall x \in [z, M] \end{cases}. \tag{7}$$

In particular, G is a MPS of F, written $F \succsim^{MPS} G$, if and only if $\mathrm{E}(G) = \mathrm{E}(F)$ and there exists $z \in [m, M]$ such that (7) holds.

Given a binary relation $\succsim$ on $\mathcal{D}$, we say that $\succsim$ is

1. *weakly risk averse* if and only if $G_{\mathrm{E}(F)} \succsim F$ for all $F \in \mathcal{D}$, that is, the decision maker prefers to receive the expected value of F with certainty to facing risk F.

2. *strongly risk averse* (or, briefly, *MPS averse*) if and only if $\succsim$ is consistent with $\succsim^{MPS}$, that is, the decision maker prefers F to all of its mean preserving spreads.

Proposition 2 *Let $\succsim$ be a binary relation represented by a continuous preference functional V and let I be bounded. The following statements are equivalent:*

(i) $\succsim$ is consistent with the concave order;

(ii) $\succsim$ is MPS averse;

(iii) $\succsim^*$ is MPS averse;

(iv) $\succsim^*$ is weakly risk averse;

(v) Each $u \in \mathcal{U}^*$ is concave.

If, in addition, V is Gateaux differentiable, then they are also equivalent to:

(vi) Each $u \in \text{range}\,\nabla V$ is concave.

[12] Indeed, although for concreteness we considered closed intervals, Lemma 1, Theorem 1, and Proposition 1) actually hold in metric spaces (that must be compact when the intervals are required to be bounded) and cumulative distribution functions are replaced by Borel probability measures.

Each of the previous conditions implies that $\succsim$ is weakly risk averse, but the converse is false.

Since it is not transitive, $\succsim^{MPS}$ is not an integral stochastic order, and so Proposition 2 is not an immediate corollary of our main results. Also observe that Proposition 2 shows how weak risk aversion of $\succsim^*$ is a stronger assumption than weak risk aversion of $\succsim$. Although this may seem counter-intuitive at first, recall that $G_{\mathrm{E}(F)} \succsim^* F$ not only implies $G_{\mathrm{E}(F)} \succsim F$ for all $F \in \mathcal{D}$, but also

$$\lambda G_{\mathrm{E}(F)} + (1-\lambda) H \succsim \lambda F + (1-\lambda) H$$

for all $F \in \mathcal{D}$, $H \in \mathcal{D}$, and $\lambda \in (0,1)$.

Like Chew, Karni, and Safra (1987), given two binary relations $\succsim_1$ and $\succsim_2$ on $\mathcal{D}$, we say that $\succsim_1$ is more risk averse than $\succsim_2$ if and only if whenever G is a SCS of F for $\succsim_2$, then $F \succsim_1 G$.

Proposition 3 *Let $\succsim_1$ be a binary relation represented by a continuous and Gateaux differentiable preference functional V, $\succsim_2$ be an expected utility preference with continuous and strictly increasing utility function v, and I be bounded. The following statements are equivalent:*

(i) $\succsim_1$ is more risk averse than $\succsim_2$;

(ii) Each $u \in \operatorname{range} \nabla V$ is a concave transformation of v.

If, in addition, $\succsim_1$ is consistent with first degree stochastic dominance, then they are also equivalent to:

(iii) Each $u \in \operatorname{range} \nabla V$ is an increasing and concave transformation of v.

Propositions 2 and 3 provide an alternative proof and a generalization of Theorems 3 and 4 of Machina (1982) that require the stronger notion of Frechet differentiability, Property 2 of Dekel (1986), and Theorem 3 of Chew, Epstein, and Segal (1991), see also Remark 2. The last two results apply to specific classes of nonexpected utility preferences.[13] The contribution of our results is both conceptual and technical. From a conceptual point of view, Proposition 2 provides an additional justification to the choice of MPS aversion as a definition of risk aversion outside the expected utility model. In fact, even without any differentiability hypothesis, this assumption is equivalent to require weak risk aversion *but* in terms of the expected utility core $\succsim^*$ of the decision maker's preference $\succsim$.

From a technical point of view our results are in terms of Gateaux derivatives rather than Frechet derivatives, that is, we require weaker and more natural differentiability assumptions. Moreover, we provide a unifying framework for some of the results in the literature and highlight the strict connection between integral stochastic orders and local utilities. To see this latter fact, assume $\succsim_2$ is an expected utility preference with continuous and strictly increasing utility function v (like in Proposition 3). Without loss of generality, assume that $v(m) = m$ and $v(M) = M$. If G is a SCS of F for $\succsim_2$, then

$$\int_{[m,M]} u(v(x))\, dF(x) \geq \int_{[m,M]} u(v(x))\, dG(x) \text{ for all concave } u \in C([m,M]). \tag{8}$$

In particular, if G is a MPS of F, then F and G satisfy (8) with v equal to the identity.[14]

[13] See Chew, Karni, and Safra (1987) for similar results concerning rank dependent utility.

[14] Along the lines of Machina (1982), note that, in view of the assumptions on $\succsim_2$ we are considering, G is a SCS of F for $\succsim_2$ if and only if $v^{-1}\left(\int_I v(x)\, dG(x)\right) = v^{-1}\left(\int_I v(x)\, dF(x)\right)$ and there exists $z \in [m,M]$ such that (7) holds. In the language of Hardy, Littlewood, and Polya (1952), the "arithmetic mean of H" given by $\mathrm{E}(H) = \int_I x dH(x)$ appearing in the definition of MPS is being replaced by the "v-mean of H" given by $\mathrm{E}_v(H) = v^{-1}\left(\int_I v(x)\, dH(x)\right)$. Thus one might think of Proposition 3 as a statement about aversion to v-mean preserving spreads.

5 Loss aversion and prudence

Loss aversion (Kahneman and Tversky, 1979) refers to the intuition that losses loom larger than gains, while prudence (also known as downside risk aversion) refers to the preference for additional risk on the upside (gain-side) rather than on the downside (loss-side) of a gamble (Menezes, Geiss, and Tressler, 1980, see also Eeckhoudt and Schlesinger, 2006, 2013).[15]

In this section we consider two very popular models of reference dependent preferences: prospect theory (Kahneman and Tversky, 1979, see Wakker, 2010, for a textbook introduction) and choice-acclimating preferences (Koszegi and Rabin, 2007, a special case of Chew, Epstein, and Segal, 1991) and, by means of the tools we developed so far, we show that loss aversion and prudence cannot coexist in these models.

5.1 Prospect theory

In prospect theory the preference functional $V : \mathcal{D} \to \mathbb{R}$ that represents $\succsim$ is given by

$$V(F) = \int_0^M w(1 - F(x))\, dv(x) - \int_m^0 \tilde{w}(F(x))\, dv(x) \tag{9}$$

where $I = [m, M]$ with $m \leq 0 \leq M$, $v : I \to \mathbb{R}$ is a continuous and strictly increasing function such that $v(0) = 0$, and $w, \tilde{w} : [0,1] \to [0,1]$ are strictly increasing and onto functions.

This very popular model has been proposed by Tversky and Kahneman (1992) to extend the scope of the classic analysis of Kahneman and Tversky (1979). Two special cases are noteworthy:

(i) $\tilde{w}(p) = w(p)$ for all $p \in [0,1]$, which corresponds to the original specification of Kahneman and Tversky (1979);

(ii) $\tilde{w}(p) = 1 - w(1-p)$ for all $p \in [0,1]$, which corresponds to rank dependent utility, a la Quiggin, 1982.

Chew, Karni, and Safra (1987) computed the local utilities of rank dependent utility. Next we compute them for prospect theory.[16]

Proposition 4 *If $w, \tilde{w} : [0,1] \to [0,1]$ are continuously differentiable, then the preference functional (9) is Gateaux differentiable, and, for each $F \in \mathcal{D}$,*

$$u_F(x) = \int_{[m,x]} \left[w'(1 - F(y))\, 1_{[0,M]}(y) + \tilde{w}'(F(y))\, 1_{[m,0)}(y) \right] dv(y) \qquad \forall x \in [m, M].$$

In prospect theory it is typically understood that 0 is the reference outcome, and *loss aversion* is formally defined by

$$v'_-(0) > v'_+(0)$$

under the implicit additional assumptions that $0 \in (m, M)$ and the left and right derivatives defined above exists. On the other hand, *prudence* corresponds to consistency with the third degree risk order (see again the works of Eeckhoudt, Schlesinger, and coauthors). The next proposition shows that the two notions cannot coexist.

[15] Although the idea of prudence and precautionary savings dates back to Kimball (1990), the general identification between the behavioral trait of prudence with consistency with the third degree risk order is due to Eeckhoudt and Schlesinger (2006). Prudence also implies preference for skewness and the two concepts coincide for some important special cases: see Arditti (1967), Whitmore (1970), Tsiang (1972), Kraus and Litzenberger (1976), and Chiu (2005).

[16] Recall that local utilities at a point F are unique only up to an additive constant. In Proposition 4, the local utility u_F has been computed by further imposing that $u_F(m) = 0$.

Proposition 5 *Let $w, \tilde{w} : [0,1] \to [0,1]$ be continuously differentiable and $m < 0 < M$. If the preference $\succsim$ represented by (9) is prudent, then v is continuously differentiable on (m, M).*

In this case, $v'(0) \neq 0$ implies $\tilde{w}(p) = 1 - w(1-p)$ for all $p \in [0,1]$.

Clearly, if $\succsim$ is consistent with third degree stochastic dominance *a fortiori* it is prudent, and so v is continuously differentiable on (m, M). But more is true, in this case V reduces to an expected utility preference functional.

Corollary 1 *Let $w, \tilde{w} : [0,1] \to [0,1]$ be continuously differentiable and $m < 0 < M$. The following statements are equivalent for a preference $\succsim$ represented by (9):*

(i) $\succsim$ is consistent with third degree stochastic dominance;

(ii) $w(p) = p = \tilde{w}(p)$ for all $p \in [0,1]$ and v is increasing, concave, and has convex derivative on (m, M).

From a behavioral viewpoint, this means that a prospect theory agent is risk averse and prudent if and only if he is an expected utility maximizer. See Schmidt and Zank (2008) for a characterization of risk aversion in prospect theory.

5.2 Choice-acclimating preferences

Chew, Epstein, and Segal (1991), inter alia, study a class of preferences represented by functionals of the form

$$V(F) = \int_I \int_I \phi(x,y)\, dF(y)\, dF(x) \tag{10}$$

where $I = [m, M]$ and $\phi : I \times I \to \mathbb{R}$ is a symmetric and continuous function.[17] They also show that V is continuous and Gateaux differentiable in the sense of (3) with

$$u_F(x) = \int_I 2\phi(x,y)\, dF(y) \qquad \forall x \in I \tag{11}$$

for all $F \in \mathcal{D}$.

Remark 2 *Although $\succsim^*$ can be obtained by (11) and Theorem 1, just by using the definition of $\succsim^*$, it is possible to show that*

$$F \succsim^* G \iff \int_I \phi(x,y)\, dF(x) \geq \int_I \phi(x,y)\, dG(x) \qquad \forall y \in I \tag{12}$$

(see Appendix B). Together with Fact 1, this means that $\langle \mathrm{range}\, \nabla V \rangle = \langle \phi(\cdot, y) : y \in I \rangle$, and one can indifferently set $\mathcal{U}^* = \mathrm{range}\, \nabla V$ or $\mathcal{U}^* = \{\phi(\cdot, y) : y \in I\}$.

As proved by Masatlioglu and Raymond (2014), Koszegi and Rabin (2007) consider the following specification of (10):[18]

$$V(F) = \int_I v(x)\, dF(x) + \int_I \int_I \mu(v(x) - v(y))\, dF(y)\, dF(x) \tag{13}$$

where $v : I \to \mathbb{R}$ is a continuous, strictly increasing, and continuously differentiable function and $\mu : \mathbb{R} \to \mathbb{R}$ is continuous, strictly increasing, twice differentiable on $\mathbb{R} \setminus \{0\}$ and such that

$$\mu'_-(0) > \mu'_+(0).$$

Again the latter condition captures *loss aversion* and again it is incompatible with *prudence*.

Proposition 6 *A preference $\succsim$ represented by (13) cannot be prudent.*

[17] For example, the mean-variance preference functional of Remark 1 corresponds to $2\phi(x,y) = x + y - \frac{\lambda}{2}(x-y)^2$.

[18] The specification corresponds to $2\phi(x,y) = v(x) + v(y) + \mu(v(x) - v(y)) + \mu(v(y) - v(x))$.

6 Conclusions

In this paper we considered a decision maker choosing among gambles and whose preferences $\succsim$ are represented by a continuous and smooth real-valued functional V. We proved two main results:

Theorem 1 showing that the largest subrelation $\succsim^*$ of $\succsim$ satisfying the axioms of expected utility (called expected utility core of $\succsim$), is represented by the set of all derivatives of V (called local utilities), that is,

$$F \succsim^* G \iff \int u(x)\, dF(x) \geq \int u(x)\, dG(x) \qquad \forall u \in \operatorname{range} \nabla V.$$

Proposition 1 showing that $\succsim$ is consistent with an integral stochastic order $\succsim^{\mathcal{U}}$, that is, with the unanimous judgement of the expected utility maximizers with utility in $\mathcal{U}$, if and only if each local utility is, up to a constant, either a weighted sum of elements of $\mathcal{U}$ or a limit of weighted sums plus constants, formally

$$\operatorname{range} \nabla V \subseteq \langle \mathcal{U} \rangle .$$

The first result is altogether new: it provides a global behavioral interpretation of a mathematical object, the set $\operatorname{range} \nabla V$ of all local utilities of V, that was known to be important in describing the properties of the underlying preference $\succsim$, but had no ordinal counterpart. Theorem 1 shows that this ordinal counterpart is $\succsim^*$. This finding also confirms the interpretation of $\succsim^*$ as the "expected utility essence" of $\succsim$ in that it is represented as a "multi–expected–utility" by the derivatives of V, which locally approximate V with expected utility functionals.

The second result, completes a strand of the decision theoretic literature ranging from Machina (1982) to Chew and Nishimura (1992) that studied the relations between risk attitudes (captured by consistency with stochastic orders) and local utilities. This literature either considered a specific form of V and/or a specific integral stochastic order (like Machina, 1982, and Chew, Karni, and Safra, 1987) or gave sufficient conditions rather than full characterizations (like Chew and Nishimura, 1992). In contrast, Proposition 1 provides a necessary and sufficient condition that applies to each preference functional V and each integral stochastic orders. Finally, while the differentiability assumption makes Proposition 1 starker, our results can be extended by direct characterization of the expected utility core.

As a byproduct of our analysis, we show that two popular loss aversion theories cannot account for prudence. Indeed, an inspection of the proofs shows that the same incompatibility with loss aversion holds also for higher orders of risk aversion such as temperance and edginess.

Appendices

A Distributions and integrals

We denote a closed interval by I. Let $m, M \in \mathbb{R}$ be such that $M > m$. We next formally define the set $\mathcal{D}(I)$. We have four possible cases:

1. $\mathcal{D}((-\infty, \infty)) = \left\{F \in \mathbb{R}^{\mathbb{R}} : F \text{ is increasing, right continuous}, \lim_{t \to -\infty} F(t) = 0, \lim_{t \to +\infty} F(t) = 1\right\}$;
2. $\mathcal{D}([m, \infty)) = \{F \in \mathcal{D}(-\infty, \infty) : F(y) = 0 \text{ for all } y < m\}$;
3. $\mathcal{D}((-\infty, M]) = \{F \in \mathcal{D}(-\infty, \infty) : F(y) = 1 \text{ for all } y \geq M\}$;
4. $\mathcal{D}([m, M]) = \mathcal{D}([m, \infty)) \cap \mathcal{D}((-\infty, M])$.

Next, we define two other important sets:

1. $\Delta_I(\mathbb{R})$, the set of all Borel probability measures with support I;

2. $\Delta(I)$, the set of all Borel probability measures on I.

Given $\mathcal{D}(I)$, we endow it with the topology of weak convergence: given $\{F_n\} \subseteq \mathcal{D}(I)$ and $F \in \mathcal{D}(I)$ we have that $\lim_n F_n = F$ if and only if $\lim_n F_n(x) = F(x)$ for all $x \in (-\infty, \infty)$ which is a continuity point of F (see Billingsley, 1995, p. 327).

Given I bounded and $\Delta(I)$, we endow the latter set with the weak* topology: given $\{\mu_n\} \subseteq \Delta(I)$ and $\mu \in \Delta(I)$ we have that $\lim_n \mu_n = \mu$ if and only if $\lim_n \int_I f d\mu_n = \int_I f d\mu$ for all $f \in C(I)$.

Next, we define two maps $T : \mathcal{D}(I) \to \Delta_I(\mathbb{R})$ and $P : \Delta_I(\mathbb{R}) \to \Delta(I)$. T is such that $T(F)$ is the unique measure on the real line, denoted by $\hat{\mu}_F$, such that $\hat{\mu}_F((a,b]) = F(b) - F(a)$ for all $a, b \in \mathbb{R}$. By Theorem 12.4 of Billingsley (1995), T is well defined. It is immediate to see that this map is affine. On the other hand, P is such that $\mu = P(\hat{\mu})$ is the measure $\hat{\mu}$ restricted to I, that is, $P(\hat{\mu})(B) = \hat{\mu}(B \cap I)$ for all Borel sets B of I. It is immediate to see that P is well defined and affine. Note that $P \circ T : \mathcal{D}(I) \to \Delta(I)$ is a map that associates to each distribution $F \in \mathcal{D}(I)$ a unique probability measure denoted by μ_F in $\Delta(I)$. If I is bounded, then $P \circ T$ is an affine homeomorphism.

Given $u \in C_b(I)$, we denote the Lebesgue-Stieltjes integral $\int_I u d\mu_F$ by $\int_I u(x) dF(x)$. If I is equal to $[m, M]$, then its relation with the Riemann-Stieltjes integral is such that:

$$\int_{[m,M]} u(x) dF(x) = \int_{[m,M]} u d\mu_F = u(m) F(m) + \int_m^M u(x) dF(x),$$

where the first equality is by definition and the second one is a well known fact. Note that the last integral is a Riemann-Stieltjes integral. Often, to differentiate a Lebesgue-Stieltjes integral from a Riemann-Stieltjes integral, we will denote the first one by $\int_{[m,M]} u(x) dF(x)$ and the second one by $\int_m^M u(x) dF(x)$. Finally, given $u \in C_b(I)$, we denote by $\int_I u(x) d(G - F)(x)$, or just by $\int_I u d(G - F)$, the difference $\int_I u(x) dG(x) - \int_I u(x) dF(x)$.

We recall that the weak topology on $C_b(I)$ is the weakest topology on it that declares continuous all functionals of the form $u \mapsto \int_I u(x) dF(x)$ where F is any element of $\mathcal{D}(I)$; when I is bounded this topology coincides with the weak topology induced by the supnorm.

B Proofs and related analysis

The proof of Lemma 1 is basically contained in Cerreia–Vioglio (2009), the difference being that here the setting are distribution functions rather than probability measures. We report it for the sake of completeness.

Proof of Lemma 1. (i) and (ii). Trivially, we have that $\succsim^*$ is a preorder. Next, consider $\{F_n\}, \{G_n\} \subseteq \mathcal{D}$ such that $F_n \to F \in \mathcal{D}$, $G_n \to G \in \mathcal{D}$, and $F_n \succsim^* G_n$ for all $n \in \mathbb{N}$. Fix $H \in \mathcal{D}$ and $\lambda \in (0,1]$. It follows that $\lambda F_n + (1-\lambda) H \succsim \lambda G_n + (1-\lambda) H$ for all $n \in \mathbb{N}$. Since $\succsim$ satisfies Continuity and $\lambda F_n + (1-\lambda) H \to \lambda F + (1-\lambda) H$ and $\lambda G_n + (1-\lambda) H \to \lambda G + (1-\lambda) H$, this implies that $\lambda F + (1-\lambda) H \succsim \lambda G + (1-\lambda) H$. Since $H \in \mathcal{D}$ and $\lambda \in (0,1]$ were arbitrarily chosen, we can conclude that $F \succsim^* G$. Next, consider $F, G, H \in \mathcal{D}$. Assume that $F \succsim^* G$ and $\lambda \in (0,1)$. It follows that

$$\begin{aligned}
\mu(\lambda F + (1-\lambda) H) + (1-\mu) H' &= (\mu\lambda) F + (1-\mu\lambda)\left[\frac{\mu(1-\lambda)}{1-\mu\lambda} H + \frac{1-\mu}{1-\mu\lambda} H'\right] \\
&\succsim (\mu\lambda) G + (1-\mu\lambda)\left[\frac{\mu(1-\lambda)}{1-\mu\lambda} H + \frac{1-\mu}{1-\mu\lambda} H'\right] \\
&= \mu(\lambda G + (1-\lambda) H) + (1-\mu) H' \qquad \forall \mu \in (0,1], \forall H' \in \mathcal{D},
\end{aligned}$$

proving that $\lambda F + (1-\lambda) H \succsim^* \lambda G + (1-\lambda) H$. Thus, $\succsim^*$ satisfies Independence. Finally, by definition of $\succsim^*$, (ii) trivially follows.

(iii). Let $\succsim$ be consistent with $\succsim^{\#}$ and let $\succsim^{\#}$ satisfy Independence. Assume that $F \succsim^{\#} G$. Since $\succsim^{\#}$ satisfies Independence, it follows that $\lambda F + (1-\lambda) H \succsim^{\#} \lambda G + (1-\lambda) H$ for all $\lambda \in (0,1]$ and for all $H \in \mathcal{D}$. Since $\succsim$ is consistent with $\succsim^{\#}$, it follows that $\lambda F + (1-\lambda) H \succsim \lambda G + (1-\lambda) H$ for all $\lambda \in (0,1]$ and for all $H \in \mathcal{D}$, that is, $F \succsim^{*} G$. Then $\succsim^{*}$ is consistent with $\succsim^{\#}$.

Conversely, let $\succsim^{*}$ be consistent with $\succsim^{\#}$, then $F \succsim^{\#} G$ implies $F \succsim^{*} G$ which by (ii) implies $F \succsim G$. That is, $\succsim$ is consistent with $\succsim^{\#}$.

(iv). Define $S = P \circ T$. Define also $\succsim^{\circ}$ on $\Delta(I)$ by $\mu \succsim^{\circ} \nu$ if and only if $S^{-1}(\mu) \succsim^{*} S^{-1}(\nu)$. By [17] and given the properties of S and $\succsim^{*}$, it follows that there exists a set $\mathcal{U}^{*} \subseteq C(I)$ such that

$$\mu \succsim^{\circ} \nu \iff \int_I u d\mu \geq \int_I u d\nu \qquad \forall u \in \mathcal{U}^{*}.$$

Thus, we can conclude that

$$\begin{aligned} F \succsim^{*} G &\iff S(F) \succsim^{\circ} S(G) \iff \mu_F \succsim^{\circ} \mu_G \iff \int_I u d\mu_F \geq \int_I u d\mu_G \quad \forall u \in \mathcal{U}^{*} \\ &\iff \int_I u(x)\, dF(x) \geq \int_I u(x)\, dG(x) \quad \forall u \in \mathcal{U}^{*}. \end{aligned}$$

(v). By point (iii), $\succsim$ is consistent with an integral stochastic order $\succsim^{\mathcal{U}}$ if and only if $\succsim^{*}$ is consistent with $\succsim^{\mathcal{U}}$. By point (iv), $\succsim^{*}$ is consistent with $\succsim^{\mathcal{U}}$ if and only if $\succsim^{\mathcal{U}^{*}}$ is consistent with $\succsim^{\mathcal{U}}$. By Fact 1, $\succsim^{\mathcal{U}^{*}}$ is consistent with $\succsim^{\mathcal{U}}$ if and only if $\mathcal{U}^{*} \subseteq \langle \mathcal{U} \rangle$. ■

Next, we give a version of the Mean Value Theorem. Given our framework and since the notion of differentiability we are using is a notion of Gateaux differentiability which involves just one sided derivatives and a particular domain, this result is not obvious even though the proof is fairly simple.

If $F, G \in \mathcal{D}$ and $t \in \mathbb{R}$, we set $F_t = (1-t) F + tG$ when no confusion can arise.

Proposition 7 *If $V : \mathcal{D} \to \mathbb{R}$ is continuous and Gateaux differentiable, then for every $F, G \in \mathcal{D}$ there exists $t \in (0,1)$ such that*

$$V(F) - V(G) = \int_I u_{F_t}(x)\, dF(x) - \int_I u_{F_t}(x)\, dG(x).$$

Proof. Consider $F, G \in \mathcal{D}$. Define $f : [0,1] \to \mathbb{R}$ by $f(t) = V((1-t) F + tG)$ for all $t \in [0,1]$. By routine arguments, it can be shown that f is continuous on $[0,1]$. As for differentiability of f on $(0,1)$, we follow Huber and Ronchetti (2009, pages 39-40). Note that $F_{t+h} = \left(1 - \frac{h}{1-t}\right) F_t + \frac{h}{1-t} G$ hence for each $t \in (0,1)$

$$\begin{aligned} f'_{+}(t) &= \lim_{h \downarrow 0} \frac{V(F_{t+h}) - V(F_t)}{h} = \lim_{h \downarrow 0} \frac{V\left(\left(1 - \frac{h}{1-t}\right) F_t + \frac{h}{1-t} G\right) - V(F_t)}{h} \\ &= \lim_{h \downarrow 0} \frac{1}{1-t} \frac{V\left(\left(1 - \frac{h}{1-t}\right) F_t + \frac{h}{1-t} G\right) - V(F_t)}{\frac{h}{1-t}} = \frac{1}{1-t} \int_I u_{F_t} d(G - F_t) = \int_I u_{F_t} d(G - F) \end{aligned}$$

(note that as h goes to 0^{+} eventually $F_{t+h} \in \mathcal{D}$) analogously $F_{t-h} = \left(1 - \frac{h}{t}\right) F_t + \frac{h}{t} F$ and

$$\begin{aligned} f'_{-}(t) &= \lim_{h \downarrow 0} \frac{V(F_{t-h}) - V(F_t)}{-h} = \lim_{h \downarrow 0} \frac{V\left(\left(1 - \frac{h}{t}\right) F_t + \frac{h}{t} F\right) - V(F_t)}{-h} \\ &= \lim_{h \downarrow 0} -\frac{1}{t} \frac{V\left(\left(1 - \frac{h}{t}\right) F_t + \frac{h}{t} F\right) - V(F_t)}{\frac{h}{t}} = -\frac{1}{t} \int_I u_{F_t} d(F - F_t) = \int_I u_{F_t} d(G - F) \end{aligned}$$

that is, $f'(t) = \int u_{F_t} dG - \int u_{F_t} dF$. By the Mean Value Theorem for functions of a real variable, it follows that there exists $t \in (0,1)$ such that

$$\int_I u_{F_t} dG - \int_I u_{F_t} dF = f'(t) = \frac{f(1) - f(0)}{1 - 0} = f(1) - f(0) = V(G) - V(F),$$

proving the statement. ∎

More powerful results of this kind appear in Cerreia–Vioglio, Maccheroni, Marinacci, and Montrucchio (2014), where von Mises calculus is studied in a very general setting.

Proof of Theorem 1. Setting $\mathcal{V} = \operatorname{range} \nabla V$, next we show that $\succsim^{\mathcal{V}}$ coincides with $\succsim^*$. Consider F and G in $\mathcal{D}$. Assume that $F \succsim^{\mathcal{V}} G$. By Proposition 7 and since $F \succsim^{\mathcal{V}} G$, we have that there exists $t \in (0,1)$ such that

$$V(F) - V(G) = \int_I u_{F_t}(x)\, dF(x) - \int_I u_{F_t}(x)\, dG(x) \geq 0,$$

yielding that $F \succsim G$. By Lemma 1 and since $\succsim^{\mathcal{V}}$ is a stochastic order, it follows that $F \succsim^* G$. Viceversa, assume that $F \succsim^* G$. Consider $H \in \mathcal{D}$. By definition of $\succsim^*$ and since V represents $\succsim$, we have that $V((1-\theta) H + \theta F) \geq V((1-\theta) H + \theta G)$ for all $\theta \in (0,1]$. This implies that

$$\frac{V((1-\theta) H + \theta F) - V(H)}{\theta} \geq \frac{V((1-\theta) H + \theta G) - V(H)}{\theta} \qquad \forall \theta \in (0,1].$$

By passing to the limits and since V is Gateaux differentiable, it follows that $\int_I u_H(x)\, dF(x) - \int_I u_H(x)\, dH(x) \geq \int_I u_H(x)\, dG(x) - \int_I u_H(x)\, dH(x)$. Since H was arbitrarily chosen, we can conclude that $F \succsim^{\mathcal{V}} G$. ∎

Proof of Proposition 1. By Lemma 1, $\succsim$ is consistent with $\succsim^{\mathcal{U}}$ if and only if $\succsim^*$ is consistent with $\succsim^{\mathcal{U}}$. By Theorem 1, $\succsim^* = \succsim^{\operatorname{range} \nabla V}$, then $\succsim$ is consistent with $\succsim^{\mathcal{U}}$ if and only if $\succsim^{\operatorname{range} \nabla V}$ is consistent with $\succsim^{\mathcal{U}}$. By Fact 1, $\succsim^{\operatorname{range} \nabla V}$ is consistent with $\succsim^{\mathcal{U}}$ if and only if $\operatorname{range} \nabla V \subseteq \langle \mathcal{U} \rangle$. ∎

Proof of Proposition 2. (i) implies (ii). It is well known that if $F \succsim^{MPS} G$, then $\int_I u(x)\, dF(x) \geq \int_I u(x)\, dG(x)$ for all concave $u \in C(I)$. Since $\succsim$ is consistent with the concave order, it follows that $F \succsim G$.

(ii) implies (iii). By definition of MPS and since $\succsim$ is MPS averse, note that

$$\begin{aligned} F \succsim^{MPS} G &\implies \lambda F + (1-\lambda) H \succsim^{MPS} \lambda G + (1-\lambda) H \qquad \forall \lambda \in (0,1], \forall H \in \mathcal{D} \\ &\implies \lambda F + (1-\lambda) H \succsim \lambda G + (1-\lambda) H \qquad \forall \lambda \in (0,1], \forall H \in \mathcal{D} \implies F \succsim^* G, \end{aligned}$$

proving that $\succsim^*$ is MPS averse.

(iii) implies (iv). Since $G_{\mathrm{E}(F)} \succsim^{MPS} F$ for all $F \in \mathcal{D}$, it follows that $G_{\mathrm{E}(F)} \succsim^* F$ for all $F \in \mathcal{D}$.

(iv) implies (v). Since I is bounded and by Lemma 1, we have that there exists a set $\mathcal{U}^* \subseteq C(I)$ that represents $\succsim^*$ as in (6). Pick $x, y \in I$. Consider $F = \frac{1}{2} G_x + \frac{1}{2} G_y$. By assumption, it follows that $G_{\frac{1}{2}x + \frac{1}{2}y} \succsim^* F$. We can conclude that $u\left(\frac{1}{2}x + \frac{1}{2}y\right) \geq \frac{1}{2} u(x) + \frac{1}{2} u(y)$ for all $u \in \mathcal{U}^*$, that is, each $u \in \mathcal{U}^*$ is concave.

(v) implies (i). By Lemma 1 and since each $u \in \mathcal{U}^*$ is concave, the statement follows.

We just showed that (i), (ii), (iii), (iv) and (v) are equivalent. Now assume that V is also Gateaux differentiable. By Theorem 1, it follows that $\mathcal{U}^*$ can be chosen to be $\operatorname{range} \nabla V$, making (v) equivalent to (vi). ∎

Proof of Proposition 3. Without loss of generality assume that $v \in C([m, M])$ is normalized, that is, $v(m) = m$ and $v(M) = M$. Define $\bar{V} : \mathcal{D} \to \mathbb{R}$ by

$$\bar{V}(F) = \int_{[m,M]} v(x)\, dF(x) \qquad \forall F \in \mathcal{D}.$$

Consider F and G in $\mathcal{D}$. Assume that G is a SCS of F for $\succsim_2$, we denote it by $F \succsim^{SCS} G$. Recall that $F \succsim^{SCS} G$ if and only if $\bar{V}(F) = \bar{V}(G)$ and there exists $z \in [m, M]$ such that

$$\begin{cases} F(x) \leq G(x) & \forall x \in [m, z) \\ F(x) \geq G(x) & \forall x \in [z, M] \end{cases}.$$

(i) implies (ii). By definition of SCS and since $\succsim_1$ is more risk averse than $\succsim_2$ and $\succsim_2$ is Expected Utility, note that

$$F \succsim^{SCS} G \implies \lambda F + (1-\lambda) H \succsim^{SCS} \lambda G + (1-\lambda) H \qquad \forall \lambda \in (0,1], \forall H \in \mathcal{D}$$
$$\implies \lambda F + (1-\lambda) H \succsim_1 \lambda G + (1-\lambda) H \qquad \forall \lambda \in (0,1], \forall H \in \mathcal{D} \implies F \succsim_1^* G,$$

proving that $\succsim_1^*$ is more risk averse than $\succsim_2$. Consider $F \in \mathcal{D}$. It is immediate to see that $G_{v^{-1}(\bar{V}(F))} \in \mathcal{D}$. Next, note that $G_{v^{-1}(\bar{V}(F))} \succsim^{SCS} F$ for all $F \in \mathcal{D}$. Consider $y_1, y_2 \in [m, M] = v([m, M])$. There exists $x_1, x_2 \in [m, M]$ such that $v(x_i) = y_i$ for $i \in \{1, 2\}$. Define $F = \frac{1}{2} G_{x_1} + \frac{1}{2} G_{x_2}$ and $\bar{y} = v^{-1}(\bar{V}(F)) = v^{-1}\left(\frac{1}{2} v(x_1) + \frac{1}{2} v(x_2)\right)$. We thus have that $G_{\bar{y}} \succsim^{SCS} F$ and so $G_{\bar{y}} \succsim_1^* F$. For each $u \in \operatorname{range} \nabla V$ define $f_u = u \circ v^{-1} \in C([m, M])$. By Theorem 1 and since $G_{\bar{y}} \succsim_1^* F$, we have that for each $u \in \operatorname{range} \nabla V$

$$\begin{aligned}
f_u\left(\frac{1}{2} y_1 + \frac{1}{2} y_2\right) &= u \circ v^{-1}\left(\frac{1}{2} y_1 + \frac{1}{2} y_2\right) = u\left(v^{-1}\left(\frac{1}{2} v(x_1) + \frac{1}{2} v(x_2)\right)\right) = u(\bar{y}) \\
&= \int_{[m,M]} u(x)\, dG_{\bar{y}}(x) \geq \int_{[m,M]} u(x)\, dF(x) = \frac{1}{2} u(x_1) + \frac{1}{2} u(x_2) \\
&= \frac{1}{2} u\left(v^{-1}(v(x_1))\right) + \frac{1}{2} u\left(v^{-1}(v(x_2))\right) = \frac{1}{2} u\left(v^{-1}(y_1)\right) + \frac{1}{2} u\left(v^{-1}(y_2)\right) \\
&= \frac{1}{2} f_u(y_1) + \frac{1}{2} f_u(y_2),
\end{aligned}$$

proving that f_u is concave and $u = f_u \circ v$.

(ii) implies (i). Consider $F, G \in \mathcal{D}$ and $u \in \operatorname{range} \nabla V$. By Theorem 1 and since each $u \in \operatorname{range} \nabla V$ is a concave transformation of v, if $F \succsim^{SCS} G$, then $\int_I u(x)\, dF(x) \geq \int_I u(x)\, dG(x)$ for all $u \in \operatorname{range} \nabla V$ which, in turn, implies that $F \succsim G$, proving the statement.

We just showed that (i) and (ii) are equivalent. Now assume that $\succsim_1$ is also consistent with first degree stochastic dominance. By Proposition 1, it follows that each $u \in \operatorname{range} \nabla V$ is also increasing. By the same proof of (i) implies (ii), we have that f_u is also increasing and this yields that (i) implies (iii). Trivially, (iii) implies (ii). ■

Consider V defined as in (9). We first report a simple property.

Lemma 2 $V : \mathcal{D} \to \mathbb{R}$ *is continuous.*

Proof of Proposition 4. We want to compute the Gateaux derivative of V at F in direction $G - F$, that is,

$$\lim_{\theta \downarrow 0} \frac{V((1-\theta) F + \theta G) - V(F)}{\theta} \qquad \forall F, G \in \mathcal{D}. \tag{14}$$

The computation is simplified by the observation that for each function $f : [m, M] \to \mathbb{R}$ of bounded variation, the Riemann-Stieltjes integral $\int_m^M f(x)\, dv(x)$ coincides with the Lebesgue-Stieltjes integral $\int_{[m,M]} f dv$ of f with respect to the Borel measure induced on $[m, M]$ by any continuous and increasing extension of v to $\mathbb{R}$. Set $H = G - F$, and note that, provided the limit in (14) exists, it is equal to

$$\begin{aligned}
&= \lim_{\theta \downarrow 0} \frac{V(F + \theta(G - F)) - V(F)}{\theta} \\
&= \lim_{\theta \downarrow 0} \frac{\int_{[0,M]} w(1 - F - \theta H)\, dv - \int_{[m,0]} \tilde{w}(F + \theta H)\, dv - \int_{[0,M]} w(1 - F)\, dv + \int_{[m,0]} \tilde{w}(F)\, dv}{\theta} \\
&= \lim_{\theta \downarrow 0} \int_{[0,M]} \frac{w(1 - F(x) - \theta H(x)) - w(1 - F(x))}{\theta} dv(x) - \int_{[m,0]} \frac{\tilde{w}(F(x) + \theta H(x)) - \tilde{w}(F(x))}{\theta} dv(x).
\end{aligned}$$

For each $x \in [m, M]$, we have that

- if $x \in [0, M]$ and $H(x) \neq 0$, then

$$\lim_{\theta \downarrow 0} \frac{w(1 - F(x) - \theta H(x)) - w(1 - F(x))}{\theta} = \lim_{\theta \downarrow 0} \frac{w(1 - F(x) - \theta H(x)) - w(1 - F(x))}{-\theta H(x)} (-H(x))$$
$$= -w'(1 - F(x)) H(x)$$

and the same holds when $H(x) = 0$;

- if $x \in [m, 0]$ and $H(x) \neq 0$, then

$$\lim_{\theta \downarrow 0} \frac{\tilde{w}(F(x) + \theta H(x)) - \tilde{w}(F(x))}{\theta} = \lim_{\theta \downarrow 0} \frac{\tilde{w}(F(x) + \theta H(x)) - \tilde{w}(F(x))}{\theta H(x)} H(x) = \tilde{w}'(F(x)) H(x)$$

and the same holds when $H(x) = 0$.

Continuous differentiability on $[0, 1]$ of w and $\tilde{w}$ implies their Lipschitzianity so that, for each $x \in [m, M]$ and each $\theta \in (0, 1)$,

$$\left| \frac{w(1 - F(x) - \theta H(x)) - w(1 - F(x))}{\theta} \right| \leq \frac{L_w |1 - F(x) - \theta(G(x) - F(x)) - (1 - F(x))|}{\theta} \leq L_w$$

and

$$\left| \frac{\tilde{w}(F(x) + \theta H(x)) - \tilde{w}(F(x))}{\theta} \right| \leq \frac{L_{\tilde{w}} |F(x) + \theta(G(x) - F(x)) - F(x)|}{\theta} \leq L_{\tilde{w}}.$$

Therefore the Dominated Convergence Theorem applied to each sequence $\theta_n \to 0^+$ yields that

$$\lim_{\theta \downarrow 0} \frac{V(F + \theta(G - F)) - V(F)}{\theta} = \int_{[0,M]} -w'(1 - F(x))(G(x) - F(x))\, dv(x)$$
$$- \int_{[m,0]} \tilde{w}'(F(x))(G(x) - F(x))\, dv(x).$$

Now define

$$\phi_F(x) = \begin{cases} w'(1 - F(x)) & x \in [0, M] \\ \tilde{w}'(F(x)) & x \in [m, 0) \end{cases} \tag{15}$$

and note that ϕ_F is bounded and Borel measurable on $[m, M]$ with

$$\lim_{\theta \downarrow 0} \frac{V(F + \theta(G - F)) - V(F)}{\theta} = -\int_{[m,M]} (G(x) - F(x)) \phi_F(x)\, dv(x).$$

Setting "$du_F = \phi_F dv$", or more precisely $u_F(x) = \int_{[m,x]} \phi_F dv$ for all $x \in [m, M]$, it is not difficult to show that $u_F \in C([m, M])$ and

$$\int_{[m,M]} (G(x) - F(x)) \phi_F(x)\, dv(x) = \int_m^M (G - F)\, du_F$$
$$= (G(M) - F(M)) u_F(M) - (G(m) - F(m)) u_F(m) - \left(\int_m^M u_F dG - \int_m^M u_F dF \right)$$
$$= -\left(\int_{[m,M]} u_F dG - \int_{[m,M]} u_F dF \right) = -\int_{[m,M]} u_F d(G - F)$$

where the second equality follows by integration by parts, the third by $G(M) = F(M) = 1$ and $u_F(m) = 0$, and the last one by definition, proving the statement. ■

Lemma 3 *If* $w, \tilde{w} : [0,1] \to [0,1]$ *are continuously differentiable,* $m < 0 < M$*, and all local utilities* u_F *of the preference functional (9) are differentiable (resp., continuously differentiable) on* (m, M)*, then* v *is differentiable (resp., continuously differentiable) on* (m, M)*.*

In this case, $v'(0) \neq 0$*, implies* $\tilde{w}(p) = 1 - w(1-p)$ *for all* $p \in [0,1]$*.*

Before entering the proof's details notice that if the preference $\succsim$ represented by V is consistent with any risk order or any stochastic dominance of a degree $n \geq 3$, then all local utilities u_F are continuously differentiable. In fact, each u_F admits $n-2$ continuous derivatives on (m, M) because its $(n-2)$-th derivative is either convex or concave and so it is continuous; clearly its derivatives of lower order are also continuous since they are differentiable.

Proof. Recall that, for each $F \in \mathcal{D}$,

$$u_F(x) = \int_{[m,x]} \left[w'(1 - F(y)) 1_{[0,M]}(y) + \tilde{w}'(F(y)) 1_{[m,0)}(y) \right] dv(y) \qquad \forall x \in [m, M]$$

and, for all $p \in [0,1]$, set $F_p = pG_m + (1-p) G_M$.

First we prove that for each $p \in [0,1]$,

$$w'(1-p) \frac{v(z) - v(x)}{z - x} = \frac{u_{F_p}(z) - u_{F_p}(x)}{z - x} \qquad \forall x, z \in [0, M) \tag{16}$$

$$\tilde{w}'(p) \frac{v(z) - v(x)}{z - x} = \frac{u_{F_p}(z) - u_{F_p}(x)}{z - x} \qquad \forall x, z \in (m, 0] \tag{17}$$

provided $x \neq z$. Since u_{F_p} is differentiable, for each $p \in [0,1]$,

$$w'(1-p) v'_+(x) = \left(u_{F_p}\right)'_+(x) \qquad \forall x \in [0, M) \tag{18}$$

$$w'(1-p) v'_-(x) = \left(u_{F_p}\right)'_-(x) \qquad \forall x \in (0, M) \tag{19}$$

$$\tilde{w}'(p) v'_-(x) = \left(u_{F_p}\right)'_-(x) \qquad \forall x \in (m, 0] \tag{20}$$

$$\tilde{w}'(p) v'_+(x) = \left(u_{F_p}\right)'_+(x) \qquad \forall x \in (m, 0) \tag{21}$$

and all left and right derivatives above are finite.

Distinguish the following cases

- for $x \in [0, M]$,
$$u_F(x) = \int_{[m,0)} \tilde{w}'(F(y)) \, dv(y) + \int_{[0,x]} w'(1 - F(y)) \, dv(y)$$
- in particular, for $x = 0$,
$$u_F(0) = \int_{[m,0)} \tilde{w}'(F(y)) \, dv(y)$$
- for $x \in [m, 0)$,
$$u_F(x) = \int_{[m,x]} \tilde{w}'(F(y)) \, dv(y)$$
  and this also holds for $x = 0$.

For every $x, z \in [0, M)$, if $z > x$, then

$$u_F(z) - u_F(x) = \int_{(x,z]} w'(1 - F(y)) \, dv(y)$$

and if $x > z$, then $u_F(x) - u_F(z) = \int_{(z,x]} w'(1 - F(y)) \, dv(y)$, thus

$$u_F(z) - u_F(x) = \begin{cases} \int_{(x,z]} w'(1 - F(y)) \, dv(y) & z > x \\ -\int_{(z,x]} w'(1 - F(y)) \, dv(y) & z < x \end{cases} .$$

If $F = F_p$, then $F_p = p$ on (m, M) and so

$$\frac{u_{F_p}(z) - u_{F_p}(x)}{z - x} = \begin{cases} w'(1-p)\frac{v(z)-v(x)}{z-x} & z > x \\ -w'(1-p)\frac{v(x)-v(z)}{z-x} & z < x \end{cases} = w'(1-p)\frac{v(z)-v(x)}{z-x} \qquad \forall x, z \in [0, M),$$

proving (16).

Analogously, for every $x, z \in (m, 0]$, if $z > x$, then

$$u_F(z) - u_F(x) = \int_{(x,z]} \tilde{w}'(F(y))\, dv(y)$$

and if $x > z$, then $u_F(x) - u_F(z) = \int_{(z,x]} \tilde{w}'(F(y))\, dv(y)$, thus

$$u_F(z) - u_F(x) = \begin{cases} \int_{(x,z]} \tilde{w}'(F(y))\, dv(y) & z > x \\ -\int_{(z,x]} \tilde{w}'(F(y))\, dv(y) & z < x \end{cases}.$$

If $F = F_p$, then $F_p = p$ on (m, M) and so

$$\frac{u_{F_p}(z) - u_{F_p}(x)}{z - x} = \begin{cases} \tilde{w}'(p)\frac{v(z)-v(x)}{z-x} & z > x \\ -\tilde{w}'(p)\frac{v(x)-v(z)}{z-x} & z < x \end{cases} = \tilde{w}'(p)\frac{v(z)-v(x)}{z-x} \qquad \forall x, z \in (m, 0],$$

proving (17).

Equations (18), (19), (20), (21), and finiteness of all left and right derivatives follow easily (note that there exist $p, q \in (0, 1)$ such that $\tilde{w}'(p) \neq 0$ and $w'(1-q) \neq 0$).

The latter equations imply that: since u_{F_p} is differentiable (resp., continuously differentiable) on (m, M) for all $p \in [0, 1]$, then v is differentiable (resp., continuously differentiable) on $(0, M)$ and $(m, 0)$.

But more is true, choosing $x = 0$ in (18) and (20), it follows that

$$\tilde{w}'(p)\, v'_-(0) = \left(u_{F_p}\right)'_-(0) = \left(u_{F_p}\right)'_+(0) = w'(1-p)\, v'_+(0) \qquad \forall p \in [0, 1], \tag{22}$$

then by integrating both sides of (22) over $[0, 1]$,

$$v'_-(0) = v'_-(0)\,\tilde{w}(1) = \int_0^1 \tilde{w}'(p)\, v'_-(0)\, dp = \int_0^1 w'(1-p)\, v'_+(0)\, dp = v'_+(0)$$

so that v is differentiable also at 0. If u_F is continuously differentiable on (m, M) for all $F \in \mathcal{D}$, then choosing p such that $\tilde{w}'(p) \neq 0$, and taking $x_n \to 0^-$ it follows that

$$\tilde{w}'(p)\, v'(x_n) = \left(u_{F_p}\right)'(x_n) \to \left(u_{F_p}\right)'(0) = \tilde{w}'(p)\, v'(0)$$

and so $v'(x_n) \to v'(0)$. Analogously, choosing q such that $w'(1-q) \neq 0$, and taking $y_n \to 0^+$ it follows that

$$w'(1-q)\, v'(y_n) = \left(u_{F_q}\right)'(y_n) \to \left(u_{F_q}\right)'(0) = w'(1-q)\, v'(0)$$

and so $v'(y_n) \to v'(0)$, so that v is continuously differentiable also at 0.

Finally, if $v'(0) \neq 0$, then (22) becomes

$$\tilde{w}'(p) = w'(1-p) \qquad \forall p \in [0, 1]$$

so that $\tilde{w}(q) = \int_0^q \tilde{w}'(p)\, dp = \int_0^q w'(1-p)\, dp = 1 - w(1-q)$ for all $q \in [0, 1]$. ■

Proof of Proposition 5. As observed immediately after the statement of Lemma 3, consistency with the third degree risk order guarantees that all local utilities u_F are continuously differentiable (with convex derivative), hence the same lemma yields both continuous differentiability of v on (m, M) and the second part of the statement. ■

Proof of Corollary 1. (i) implies (ii). By Proposition 5 and as observed immediately after the statement of Lemma 3, note that if $\succsim$ is consistent with third degree stochastic dominance, then v and all local derivatives are continuously differentiable on (m, M). Moreover, $\succsim$ is consistent with second degree stochastic dominance, and this implies that u_F is concave for all $F \in \mathcal{D}$. But then v' is decreasing on $(m, 0]$, by (20), and on $[0, M)$, by (18), thus it is decreasing on (m, M), so that v is concave. Since it is strictly increasing, then $v'(0) \neq 0$ and, by Lemma 3, $\tilde{w}(p) = 1 - w(1-p)$ for all $p \in [0, 1]$.

Arbitrarily choose $p < q$ in $(0, 1)$ and $x < t$ in $(m, 0)$. For each $F \in \mathcal{D}$,

$$\begin{aligned} u_F(z) - u_F(x) &= \int_{(x,z]} \tilde{w}'(F(y))\, dv(y) \qquad \forall z \in (x, 0) \\ u_F(z) - u_F(t) &= \int_{(t,z]} \tilde{w}'(F(y))\, dv(y) \qquad \forall z \in (t, 0) \end{aligned}$$

hence choosing $H_{pq} \in \mathcal{D}$ such that $H_{pq} = p$ in a neighborhood U_x of x and $H_{pq} = q$ in a neighborhood U_t of t, then

$$\begin{aligned} \frac{u_{H_{pq}}(z) - u_{H_{pq}}(x)}{z - x} &= \tilde{w}'(p) \frac{v(z) - v(x)}{z - x} \qquad \forall z \in U_x,\ z > x \\ \frac{u_{H_{pq}}(z) - u_{H_{pq}}(t)}{z - t} &= \tilde{w}'(q) \frac{v(z) - v(t)}{z - t} \qquad \forall z \in U_t,\ z > t \end{aligned}$$

so that

$$v'(x)\, \tilde{w}'(p) = u'_{H_{pq}}(x) \geq u'_{H_{pq}}(t) = v'(t)\, \tilde{w}'(q)$$

for all $0 < p < q < 1$ and all $m < x < t < 0$. But then letting $t \to x$, by continuity of v', we obtain $v'(x)\, \tilde{w}'(p) \geq v'(x)\, \tilde{w}'(q)$ and $v'(x) > 0$ yields $\tilde{w}'(p) \geq \tilde{w}'(q)$, in turn, this implies that $\tilde{w}$ is concave. As a consequence, $w(p) = 1 - \tilde{w}(1-p)$ is convex.

Now consider G_0 and observe that for every $x, z \in [0, M)$, with $z > x$,

$$u_{G_0}(z) - u_{G_0}(x) = \int_{(x,z]} w'(1 - G_0(y))\, dv(y) = w'(0)\,(v(z) - v(x))$$

and so

$$u'_{G_0}(x) = w'(0)\, v'(x).$$

Analogously, for every $t, z \in (m, 0]$, with $z < t$, then

$$u_{G_0}(z) - u_{G_0}(t) = -\int_{(z,t]} \tilde{w}'(G_0(y))\, dv(y) = -\int_{(z,t]} \tilde{w}'(0)\, dv(y)$$

because v is continuous at 0 and $\tilde{w}'(G_0(y)) = \tilde{w}'(0)$ on $(m, 0)$, thus

$$u'_{G_0}(t) = \tilde{w}'(0)\, v'(t).$$

For $x = t = 0$, we have $w'(0)\, v'(0) = \tilde{w}'(0)\, v'(0)$ and

$$w'(0) = \tilde{w}'(0)$$

but $\tilde{w}'(p) = w'(1-p)$ for all $p \in (0, 1)$ and by continuity $\tilde{w}'(0) = w'(1)$, that is, $w'(0) = w'(1)$ and w' being increasing must be constant on $[0, 1]$. Therefore $w(p) = p = \tilde{w}(p)$ for all $p \in [m, M]$.

Finally, this implies

$$V(F) = \int_{[m,M]} v(y)\, dF(y) \qquad \forall F \in \mathcal{D}$$

and consistency with third degree stochastic dominance also implies that v' is convex.

(ii) implies (i). It is a well known fact. ■

Derivation of Equation 12. Define $W : \mathcal{D} \times \mathcal{D} \to \mathbb{R}$ by

$$W(F,G) = \int_{[m,M]} \int_{[m,M]} \phi(x,y)\, dF(x)\, dG(y) \qquad \forall (F,G) \in \mathcal{D} \times \mathcal{D}.$$

It is immediate to check that W is affine in both components, $W(F,G) = W(G,F)$ for all $(F,G) \in \mathcal{D} \times \mathcal{D}$, and $V(F) = W(F,F)$ for all $F \in \mathcal{D}$. We start by observing two facts:

(a) Fix $F, H \in \mathcal{D}$. If we define $F_\gamma = \gamma F + (1-\gamma) H$ for all $\gamma \in (0,1]$, then

$$\begin{aligned} V(F_\gamma) &= W(F_\gamma, F_\gamma) = \gamma W(F, F_\gamma) + (1-\gamma) W(H, F_\gamma) \\ &= \gamma^2 W(F,F) + \gamma(1-\gamma) W(F,H) + (1-\gamma)\gamma W(H,F) + (1-\gamma)^2 W(H,H) \\ &= \gamma^2 W(F,F) + 2(1-\gamma)\gamma W(F,H) + (1-\gamma)^2 W(H,H). \end{aligned}$$

(b) Fix $F, G \in \mathcal{D}$. If $\int_{[m,M]} \phi(x,y)\, dF(x) \geq \int_{[m,M]} \phi(x,y)\, dG(x)$ for all $y \in [m,M]$, then for each $H \in \mathcal{D}$

$$W(F,H) = \int_{[m,M]} \int_{[m,M]} \phi(x,y)\, dF(x)\, dH(y) \geq \int_{[m,M]} \int_{[m,M]} \phi(x,y)\, dG(x)\, dH(y) = W(G,H).$$

In particular, since H was arbitrarily chosen, we have that

$$V(F) = W(F,F) \geq W(G,F) = W(F,G) \geq W(G,G) = V(G).$$

Next, by facts (a) and (b), observe that

$$\begin{aligned} F \succsim^* G &\iff \lambda F + (1-\lambda) H \succsim \lambda G + (1-\lambda) H \qquad \forall \lambda \in (0,1], \forall H \in \mathcal{D} \\ &\iff V(\lambda F + (1-\lambda) H) - V(\lambda G + (1-\lambda) H) \geq 0 \qquad \forall \lambda \in (0,1], \forall H \in \mathcal{D} \\ &\iff \lambda^2 (V(F) - V(G)) + 2\lambda(1-\lambda)(W(F,H) - W(G,H)) \geq 0 \qquad \forall \lambda \in (0,1], \forall H \in \mathcal{D} \\ &\iff \lambda (V(F) - V(G)) + 2(1-\lambda)(W(F,H) - W(G,H)) \geq 0 \qquad \forall \lambda \in (0,1], \forall H \in \mathcal{D} \\ &\iff V(F) \geq V(G) \text{ and } W(F,H) - W(G,H) \geq 0 \qquad \forall H \in \mathcal{D} \\ &\iff V(F) \geq V(G) \text{ and } \int_{[m,M]} \phi(x,y)\, dF(x) \geq \int_{[m,M]} \phi(x,y)\, dG(x) \qquad \forall y \in [m,M] \\ &\iff \int_{[m,M]} \phi(x,y)\, dF(x) \geq \int_{[m,M]} \phi(x,y)\, dG(x) \qquad \forall y \in [m,M], \end{aligned}$$

proving the statement. ■

Proof of Proposition 6. Before starting, note that since v is strictly increasing and continuously differentiable, we have that there exists $\bar{y} \in (m,M)$ such that $v'(\bar{y}) > 0$ and $v'(x) \geq 0$ for all $x \in [m,M]$. Wlog, we can assume that $v(\bar{y}) = 0$. By contradiction, assume that $\succsim$ is prudent. By point (v) of Lemma 1 and Remark 2, this means that the set $\mathcal{U}^* = \{\phi(\cdot, y)\}_{y \in [m,M]}$ is included in $\langle \mathcal{R}_3 \rangle = \{u \in C([m,M]) : u' \text{ exists}$ and is convex on $(m,M)\}$. Among all the elements of $\mathcal{U}^*$, consider $\phi(\cdot, \bar{y}) : I \to \mathbb{R}$. By Masatlioglu and Raymond (2014), note that

$$\begin{aligned} \phi(x, \bar{y}) &= \frac{v(x) + v(\bar{y}) + \mu(v(x) - v(\bar{y})) + \mu(v(\bar{y}) - v(x))}{2} \\ &= \frac{1}{2}(v(x) + \mu(v(x)) + \mu(-v(x))) \qquad \forall x \in I. \end{aligned}$$

Since $\phi(\cdot, \bar{y})$ is differentiable on $(m,M) \ni \bar{y}$, observe also that

$$\phi'_\pm(\bar{y}, \bar{y}) = \frac{1}{2}\left(v'(\bar{y}) + \mu_\pm(0) v'(\bar{y}) - \mu_\mp(0) v'(\bar{y})\right)$$

and $\phi'_+(\bar{y},\bar{y}) = \phi'_-(\bar{y},\bar{y})$. Since $v'(\bar{y}) > 0$, this implies that

$$v'(\bar{y}) + \mu_+(0)\, v'(\bar{y}) - \mu_-(0)\, v'(\bar{y}) = v'(\bar{y}) + \mu_-(0)\, v'(\bar{y}) - \mu_+(0)\, v'(\bar{y}),$$

that is, $\mu_+(0) = \mu_-(0)$, a contradiction with $\mu'_-(0) > \mu'_+(0)$. ■

References

[1] F. D. Arditti, Risk and the required return on equity, *Journal of Finance*, 22, 19–36, 1967.

[2] P. Billingsley, *Probability and measure*, 3rd ed., Wiley, New York, 1995.

[3] J. M. Borwein and W. B. Moors, Stability of closedness of convex cones under linear mappings, Journal of Convex Analysis, 16, 699–705, 2009.

[4] E. Castagnoli and F. Maccheroni, Generalized stochastic dominance and unanimous preferences, in G. Giorgi and F. Rossi (Eds.), *Generalized convexity and optimization for economic and financial decisions*, Pitagora, Bologna, 1998.

[5] S. Cerreia–Vioglio, Maxmin expected utility on a subjective state space: convex preferences under risk, mimeo, 2009.

[6] S. Cerreia–Vioglio, D. Dillenberger, and P. Ortoleva, Cautious expected utility and the certainty effect, *Econometrica*, forthcoming, 2015.

[7] S. Cerreia–Vioglio, P. Ghirardato, F. Maccheroni, M. Marinacci, and M. Siniscalchi, Rational preferences under ambiguity, *Economic Theory*, 48, 341–375, 2011.

[8] S. Cerreia–Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio, Uncertainty averse preferences, *Journal of Economic Theory*, 146, 1275–1330, 2011.

[9] S. Cerreia–Vioglio, F. Maccheroni, M. Marinacci, L. Montrucchio, Affine Gateaux differentials and the von Mises statistical calculus, mimeo, 2014.

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

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

[12] S. H. Chew and N. Nishimura, Differentiability, comparative statics, and non-expected utility preferences, *Journal of Economic Theory*, 56, 294–312, 1992.

[13] W. H. Chiu, Skewness preference, risk aversion, and the precedence relation on stochastic changes, *Management Science*, 51, 1816–1828, 2005.

[14] G. Debreu, Continuity properties of Paretian utility, *International Economic Review*, 5, 285–293, 1964.

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

[16] P. A. Diamond, Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: Comment, *The Journal of Political Economy*, 75, 765–766, 1967.

[17] J. Dubra, F. Maccheroni, and E. A. Ok, Expected utility theory without the completeness axiom, *Journal of Economic Theory*, 115, 118–133, 2004.

[18] L. Eeckhoudt and H. Schlesinger, Putting risk in its proper place, *The American Economic Review*, 96, 280–289, 2006.

[19] L. Eeckhoudt, H. Schlesinger, and I. Tsetlin, Apportioning of risks via stochastic dominance, *Journal of Economic Theory*, 144, 994–1003, 2009.

[20] L. Eeckhoudt and H. Schlesinger, Higher-order risk attitudes, in G. Dionne (Ed.), *Handbook of insurance*, Springer, New York, 2013

[21] O. Evren, On the existence of expected multi-utility representations, *Economic Theory*, 35, 575–592, 2008.

[22] L. T. Fernholz, *Von Mises calculus for statistical functionals*, Springer, New York, 1983.

[23] P. Ghirardato, F. Maccheroni, and M. Marinacci, Differentiating ambiguity and ambiguity attitude, *Journal of Economic Theory*, 118, 133–173, 2004.

[24] P. Ghirardato and M. Siniscalchi, Ambiguity in the small and in the large, *Econometrica*, 80, 2827–2847, 2012.

[25] F. R. Hampel, The influence curve and its role in robust estimation, *Journal of the American Statistical Association*, 69, 383–393, 1974.

[26] G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities*, Cambridge University Press, Cambridge, 1952.

[27] J. C. Harsanyi, Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility, *The Journal of Political Economy*, 63, 309–321, 1955.

[28] P. J. Huber and E. M. Ronchetti, *Robust statistics*, 2^{nd} ed., Wiley, New York, 2009.

[29] D. Kahneman and A. Tversky, Prospect theory: an analysis of decision under risk, *Econometrica*, 47, 263–291, 1979.

[30] S. Karlin and A. Novikoff, Generalized convex inequalities, Pacific Journal of Mathematics, 13, 1251–1279, 1963.

[31] R. L. Keeney, On the foundations of prescriptive decision analysis, in W. Edwards (Ed.), *Utility theories: measurements and applications*, Kluwer, Boston, 1992.

[32] M. S. Kimball, Precautionary saving in the small and in the large, *Econometrica*, 58, 53–73, 1990.

[33] B. Koszegi and M. Rabin, Reference-dependent risk attitudes, *American Economic Review*, 97, 1047–1073, 2007.

[34] A. Kraus and R. H. Litzenberger, Skewness preference and the valuation of risk assets, *Journal of Finance*, 31, 1085–1100, 1976.

[35] D. M. Kreps, A representation theorem for "preference for flexibility", *Econometrica*, 47, 565–577, 1979.

[36] F. Lajeri–Chaherli, Proper prudence, standard prudence and precautionary vulnerability, *Economics Letters*, 82, 29–34, 2004.

[37] H. Levy, Stochastic dominance and expected utility: survey and analysis, *Management Science*, 38, 555–593, 1992.

[38] F. Maccheroni, Maxmin under risk, *Economic Theory*, 19, 823–831, 2002.

[39] M. J. Machina, "Expected utility" analysis without the independence axiom, *Econometrica*, 50, 277–323, 1982.

[40] H. M. Markowitz, Portfolio selection, *Journal of Finance*, 7, 77–91, 1952.

[41] Y. Masatlioglu and C. Raymond, A behavioral analysis of stochastic reference dependence, mimeo, 2014.

[42] C. Menezes, C. Geiss, and J. Tressler, Increasing downside risk, *American Economic Review*, 70, 921–932, 1980.

[43] R. von Mises, On the asymptotic distribution of differentiable statistical functions, *Annals of Mathematical Statistics,* 18, 309–348, 1947.

[44] A. Muller, Stochastic orders generated by integrals: a unified study, *Advances in Applied Probability*, 29, 414–428, 1997.

[45] K. Nehring, Imprecise probabilistic beliefs as a context for decision-making under ambiguity, *Journal of Economic Theory*, 144, 1054–1091, 2009.

[46] T. Popoviciu, Sur quelques proprietes des fonctions d'une ou de deux variables reelles, Ph.D. thesis, Ecole Normale Superieure, Paris, 1933.

[47] J. Quiggin, A theory of anticipated utility, *Journal of Economic Behavior and Organization,* 3, 323–343, 1982.

[48] U. Schmidt and H. Zank, Risk aversion in cumulative prospect theory, *Management Science*, 54, 208–216, 2008.

[49] J. E. Smith, Risk sharing, fiduciary duty, and corporate risk attitudes, *Decision Analysis*, 1, 114–127, 2004.

[50] J. E. Smith and K. F. McCardle, Structural properties of stochastic dynamic programs, *Operations Research*, 50, 796–809, 2002.

[51] J. Tobin, Liquidity preference as behavior towards risk, *Review of Economic Studies*, 25, 65–86, 1958.

[52] S. C. Tsiang, The rationale of the mean-standard deviation analysis, skewness preference, and the demand for money, *American Economic Review*, 62, 354–371, 1972.

[53] A. Tversky and D. Kahneman, Advances in prospect theory: cumulative representation of uncertainty, *Journal of Risk and uncertainty,* 5, 297–323, 1992.

[54] P. Wakker, *Prospect theory: for risk and ambiguity*, Cambridge University Press, Cambridge, 2010.

[55] T. Wang, L_p–Frechet differentiable preference and "local utility" analysis, *Journal of Economic Theory*, 61, 139–159, 1993.

[56] G. A. Whitmore, Third-degree stochastic dominance, *American Economic Review*, 60, 457–459, 1970.