Stochastic dominance for law invariant preferences: The happy story of elliptical distributions

M. Del Vigna*

October 27, 2012

Abstract

We study the connections between stochastic dominance and law invariant preferences. Whenever the functional that represents preferences depends only on the law of the random variable, we shall look for conditions that imply a ranking of distributions. In analogy with the Expected Utility paradigm, we prove that functional dominance leads to first order stochastic dominance. We analyze in details the case of Dual Theory of Choice and Cumulative Prospect Theory, including all its distinctive features such as S-shaped value function, reversed S-shaped probability distortions and loss aversion. These cases can be seen as special examples of a more general scheme. We find necessary and sufficient conditions that imply preferences to depend only on the mean and variance of the lottery. Our main result is a characterization of such distributions that imply Mean-Variance preferences, namely the elliptical ones. In particular, we prove that under mild assumptions over the reference wealth, the prospect value of a portfolio depends only on its mean and variance if and only if the random assets' return are elliptically distributed. The analysis is of particular relevance for optimal portfolio choice, mutual fund separation and Capital Asset Pricing equilibria.

Keywords: Stochastic dominance - Cumulative Prospect Theory - Elliptical distributions - Mean-Variance analysis.
AMS classification (2000): 91B06, 91G10.
JEL classification: D03, D81, G11.

1 Introduction

Decisions among risky ventures need a way to rank them. While complete orderings are extremely rare to achieve, partial orderings are quite easy to obtain. When facing multiple alternatives, a decision maker can at first discard those which result inefficient. After this removal process, the remaining alternatives form an efficient set to her.

In economic contexts, a risky venture is often called a lottery.[1] To reach an ordering, we can attach to each lottery its probability distribution. Thus, ranking prospects can be interpreted as ranking distributions and efficient lotteries translate into efficient distributions. Notably, the roots of this procedure can be traced back to statistical hypotheses tests. One of the earlier results slightly departing from this field of study is Lehmann [26], which undoubtedly inspired subsequent works in pure and applied decision theory. Systematic studies in financial frameworks appeared in the early 60's, culminating with the fundamental contributions of Hanoch and Levy [17] and Hadar and Russell [16]. Quirk and Saposnik [34] introduced the terms "stochastic dominance" (SD hereafter). In its original framing, SD is a weak partial ordering over a set of distributions. Influenced by the von Neumann-Morgenstern's Expected Utility (EU) paradigm, subsequent literature refers to SD as a conditions over a set of utility functions that imply ranking of distributions. However, it makes sense to speak about SD only if one is able to provide such rankings.

*Dipartimento di Matematica per le Decisioni, Università degli Studi di Firenze, matteo.delvigna@unifi.it.

[1] To us, the words lottery, prospect and random variable will be synonymous.

Connections between SD and EU has been extensively studied and understood; for a comprehensive survey, see Levy [28]. The most relevant cases are those of First Stochastic Dominance (FSD) to include increasing utility functions and Second Stochastic Dominance (SSD) to incorporate concave utility functions. After that, a number of adjustments to SD has been implemented in order to accommodate various classes of utility functions. However, SD for alternative preference paradigms lacks a systematic treatment.

The first aim of the present paper is to give a unified framework for SD relations for law invariant preferences. Using familiar language, law invariance means that the decision maker only cares about probabilities, totally disregarding events. In turn, the functional that represents preferences depends exclusively on the distribution of the prospect. Law invariant preferences are not new to the financial literature and they have a wide range of applications. Just to mention a few, Kusuoka [25] applies this idea to coherent risk measures. Carlier and Dana [5] do the same in the context of portfolio optimization problems for concave utility functions. In Campi and Del Vigna [4] law invariance naturally appears in a number of continuous-time portfolio optimization problem. Finally, law invariance is closely related to the quantile approach. In this approach, the preference functional is framed in terms of the inverse of the distribution. This technique has often been employed in the SD literature, but also in portfolio choice problems. See He and Zhou [18] for a modern general treatment.

The family of law invariant preferences includes EU as well as Prospect Theory (PT) by Kahneman and Tversky [21], the Dual Theory of Choice (DT) by Yaari [39] and Tversky and Kahneman' Cumulative Prospect Theory (CPT) [37]. In this paper, we will focus on DT and the more challenging CPT. As it will appear clear, in our setting DT and CPT can be seen as involved examples of a more general picture. Basic results concerning SD for the aforementioned paradigms are well-known. Interestingly, PT even violates FSD due to the distortion applied to perceived probabilities. On the contrary, DT has a plain analytical tractability which makes it suitable for ranking distributions. Various SD relations for DT are studied in [15], where the authors also adapt DT to multivariate risks. We obtain analogous results using a quite different language and we provide substantially different proofs.

CPT deserves a special treatment. On one hand, it is not difficult to prove that the CPT functional value (also called prospect value) is decreasing in the distribution of the lottery, a concept to be made rigorous later. However, many authors circumvent technical difficulties assuming specific forms for the value functions and/or the probability distortions. On the contrary, we do not set such limitations and we prove generalized versions of this results using a less sophisticated machinery. On the other hand, proving the converse is not immediate at all. To the best of our knowledge, we are proving for the first time the following fact: If the prospect value of some lottery dominates that of another lottery for any CPT decision maker in a suitable class, then the same ordering is preserved over the distributions of such lotteries; see Proposition 6. In other words, the prospect value ranking translates to SD. Note that the same result arises quite naturally in the EU paradigm. As a consequence, rankings of distributions can be compared even across different preference paradigms.

In the literature, other types of stochastic dominance have been investigated. Here is a list of works related to ours. Concerning CPT, Baucells and Heukamp [2] provide various extensions to SSD. Notably, they are able to incorporate all the basic CPT features, ranging from S-shaped value function to reversed S-shaped weighting functions and loss aversion. A different perspective is that of Barberis and Huang [1], where SSD is shown starting from normal distributions for the prospects. Levy and Wiener [31] introduce SSD* and Prospect Stochastic Dominance. The first one is just the mirror of SSD for risk-seeking decision makers while the latter is an extension of SD that allows for probability distortions. Their study on Prospect SD is based on the quantile approach and it leads to results similar to ours on DT for risk averse agents. Notably, Levy and Wiener [31] anticipate some of our main findings. In particular, they explain how the distortion of probabilities by different subjects can affect the respective distribution rankings. However, if it is not the case then even substantially different preferences paradigms lead to the same efficient sets. Leshno and Levy [27] considere Almost Stochastic Dominance in order to shed light on some financial puzzles. In words, Almost SD originates whenever "most" decision makers agree on the set of lotteries to be excluded from an efficient set. In practice, it is the classical SD analysis discarding wild or extreme utility functions. In Levy and Levy [30] Prospect SD undergoes a deeper analysis together with Markowitz

Stochastic Dominance, i.e. SD for reversed S-shaped utility functions. Experimental studies confirm the latter form for the utility function, rejecting the hypothesis of a S-shaped utility function as in CPT. Prospect SD and Markowitz SD are extended in Wong and Chan [38] to accommodate for various shapes of the utility functions as well as for the probability weightings.

Our analysis proceeds with a review of some classes of distributions that play a relevant role in portfolio choice theory. We shall focus on spherical symmetric, elliptical and location-scale distributions. At first sight, they may seem unrelated to SD as well as to preference paradigms other than EU. Actually, we will show that they provide the natural link between law invariant functionals that rank distributions and Mean-Variance (MV) preferences. The strength of MV undoubtedly lies in its parsimonious description of investors' preferences. Nevertheless, MV leads to a number of non trivial implications for optimal investment rules, the most appealing being mutual fund separation and Capital Asset Pricing.

Regarding elliptical distributions, they were first introduced in the financial literature in order to generalize the normal distribution. Remarkably, they fit better for portfolio applications since this class includes random variables with heavy tails densities. A list of econometric works that aim at calibrating elliptical distributions to real data is an overly laborious task and is out of our scope. Levy and Duchin [29] briefly review the history of heavy-tailed distributions fittings, starting in the early 60's and still proceeding farther. They add their own estimates about goodness of fit for a number of distributions. They cover the normal, the extreme value, the log-normal, the Student-t, the skew-normal and the stable distribution, just to mention a few. However, the logistic distribution turns out to be the most promising independently on the considered time horizon. Notably, the logistic is elliptical and our findings show that it leads to MV preference functionals. Consequently, the assumption of MV decision makers has to be seriously considered a realistic one.

In the last part of the paper we focus on a conjecture by Pirvu and Schulze [33] concerning a characterization of the distributions that imply MV prospect value functionals. In the aforementioned paper, the authors prove that in the presence of a riskless asset and multivariate elliptically distributed assets' return, the prospect value of a portfolio depends only on its mean and variance. However, they conjecture (but are not able to prove) that the converse holds too. Briefly, our Corollary 4 gives the answer. Under suitable assumptions over the benchmark wealth, it is true that MV prospect value implies elliptically distributed stocks' returns. Our findings thus bridge the gap between CPT and MV analysis.

Before proceeding to our contributions, we remark that in the SD literature it is usual to prove results by contradiction, that is to exhibit a utility function which induces a violation of the assumptions. As we consider general classes of preferences, we shall provide particular functionals that lead to contradiction. Notably, most of our proofs will combine this technique with a limit argument and we preferred to give explicit expressions whenever possible. In contrast to the literature on CPT that often affords on numerical simulations, we are always able to provide analytical proofs that make our results sound from a theoretical viewpoint.

This is the outline of the paper. Section 2 introduces law invariant preferences, stochastic dominance and the main facts about distribution rankings. We proceed by analyzing SD for the DT and the CPT. Section 3 concerns spherical symmetric, elliptical and location-scale distributions. We point out the equivalence of some definitions and we highlight their useful properties. The presentation relies on Kelker [22] and Owen and Rabinovitch [32]. Section 4 shows the connections between distribution rankings, mean-variance analysis and elliptical distributions of assets' returns. The CPT case is treated in details. Section 5 concludes, suggesting various lines for future research. Involved proofs and examples are relegated to the appendix.

2 Stochastic dominance and law invariant preferences

Throughout the paper, $\mathbb{E}$ and $\mathbb{V}$ denote the mean and the covariance operator of a random element respectively, the prime denotes the transpose of a matrix or the derivative of a function and $X \stackrel{d}{=} Y$ means that the random elements X and Y have the same distribution. Integral are to be intended in the Lebesgue-Stieltjes sense and they are assumed to exists, possibly taking infinite values. We denote $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}$, $\mathbb{R}^+ := [0, +\infty)$ and $\mathbb{R}^- := (-\infty, 0]$. Finally, $\mathcal{C}^k(U;V)$ is the set of functions $f : U \to V$ continuously differentiable up to the order k.

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a non-atomic probability space and $\mathcal{X} \subseteq L^0(\Omega, \mathscr{F}, \mathbb{P})$ a given set of measurable real-valued random variables. When modeling risky choices, preferences are defined as binary relations over a set of random variables, i.e. $\mathcal{P} \subset \mathcal{X} \times \mathcal{X}$. Preferences must satisfy given axioms and under suitable assumptions it turns out that they can be represented by a real-valued functional ρ. In turn, this functional will be characterized by a set of parameters lying in a specific class S. We think of S as a product of subsets of function spaces. Hence we will write $\rho : S \times \mathcal{X} \to \overline{\mathbb{R}}$. However, in most cases preferences can be stated over a set of distribution functions instead of over $\mathcal{X}$. Thus, we are led to the definition of law invariant preferences.

Definition 1 (Law invariant and monotone functional). *Assume that preferences $\mathcal{P}$ can be represented through $\rho : S \times \mathcal{X} \to \overline{\mathbb{R}}$.*

- *ρ is law invariant in S iff for every $s \in S$, $\rho(s, X) = \rho(s, Y)$ whenever $X \overset{d}{=} Y$.*
- *For $s \in S$, $\rho(s, \cdot)$ is monotone decreasing iff $F_1(x) \leq F_2(x)$ for every $x \in \mathbb{R}$ implies $\rho(s, F_1) \geq \rho(s, F_2)$.*

The previous definition is in the same spirit to that of law invariant coherent risk measures in Kusuoka [25]. From now on, we shall deal exclusively with law invariant ρ. In words, law invariant preferences means that our agent is only interested in probabilities and not in events. Therefore, we will focus on distributions instead of random variables. Since we identify elements in $\mathcal{X}$ whenever they have the same probability law, we shall indifferently replace $\rho(s, X)$ with $\rho(s, F_X)$, where F_X is the distribution of X.

Example 1. If we consider an EU agent then S is a set of utility functions. We can choose the class S to be $\mathcal{U}_0 := \{u \in \mathcal{C}^0(\mathbb{R};\mathbb{R}) \,|\, \text{u is non decreasing}\}$ or $\mathcal{U}_1 := \{u \in \mathcal{C}^\infty(\mathbb{R};\mathbb{R}) \,|\, u' > 0\}$. The representing functional will be $\rho^{EU}(u, X) := \mathbb{E}[u(X)] = \int_{\mathbb{R}} u(x) dF_X(x)$. Another type of preferences is described by the Dual Theory of Choice as axiomatized in Yaari [39]. In this case, S will be a set of probability distortions. We could choose $\mathcal{T}_0 := \{T : [0,1] \to [0,1] \,|\, T(0) = 0,\, T(1) = 1,\, T \text{ is non decreasing}\}$, with representing functional $\rho^Y(T, X) := \int_0^{+\infty} T(1 - F_X(x)) dx$.

The functional $\rho(s, \cdot)$ describes an ordering induced by the preferences $\mathcal{P}$ of a specific decision maker. Such ordering is arbitrary in the sense that nobody can assess whether it is good or not. But if the preference axioms are reasonable, i.e. they make sense economically, then the alternatives for S will be somewhat restricted and $\rho(s, \cdot)$ will satisfy some basic properties independently of $s \in S$. Even if not so evident, this is the idea that lies behind Stochastic Dominance criteria. Formally, we have the following definition.

Definition 2 (Stochastic dominance). *Let $X_1, X_2 \in \mathcal{X}$ and let F_1, F_2 be the respective distribution functions.*

- *F_1 stochastically dominates F_2 iff $F_1(x) \leq F_2(x)$ for every $x \in \mathbb{R}$.*
- *X_1 dominates X_2 with respect to ρ and S ($X_1 \succeq_{(\rho,S)} X_2$ or $F_1 \succeq_{(\rho,S)} F_2$) iff for every $s \in S$, it holds $\rho(s, F_1) \geq \rho(s, F_2)$.*
- *(ρ, S) ranks the distributions iff $F_1 \succeq_{(\rho,S)} F_2$ implies F_1 stochastically dominates F_2.*

The intuition behind (ρ, S) ranking the distributions is to obtain a converse of the monotonicity of ρ as in Definition 1. Now, one of the main application of SD concerns efficient sets. Whenever we are able to rank two distributions using SD, then we can exclude the dominated alternative from the set of optimal choices. Proceeding this way, the remaining lotteries will be the efficient ones. In order to deal with efficient sets, it would be convenient to find conditions that ensure F_2 to be dominated by some F_1. A look at Definition 2 shows that this approach heavily depends on the preference functional and a relative class of parameters. In fact, it would be sufficient to find a pair (ρ, S) that ranks the distributions to come to an efficient set.

Remark 1. When originally defining SD, Quirk and Saposnik [34] also required $F_1(x) < F_2(x)$ for some $x \in \mathbb{R}$. Notwithstanding, the weak partial ordering induced over a given set of distributions is the same as ours. The authors also proved that such ordering is in fact equivalent to the one induced by EU with the class of increasing functions; see Proposition 1. From that moment on, SD is often used as a synonymous to $\succeq_{(\rho^{EU},\mathcal{U}_0)}$. We prefer to keep the two concepts separately. The reason will appear clear later. Similarly, in Hanoch and Levy [17] the authors say that X dominates Y $(X\,D\,Y)$ if $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for every $u \in \mathcal{U}_0$ and the equality holds strictly for some u_0. This criterion induces an ordering, denoted D by the authors, which can be intended as a rule $\succneqq_{(\rho^{EU},\mathcal{U}_0)}$. However, we prefer to use the weak orderings induced by $\succeq_{(\rho,S)}$ since they allow more concise proofs and lead to clearer results without altering the economic meaning. Finally, observe that $\succeq_{(\rho,S)}$ is an equivalence relation. Hence, if we define equivalence classes of distributions by identifying those random variables such that $\rho(s,X) = \rho(s,Y)$ for every $s \in S$, then $\succneqq$ and $\succeq$ lead to the same efficient sets.

In order to have a sound economic criterion to rank risky prospects, it is necessary to make this criterion independent on the initial wealth of the decision maker. In other words, whether to accept or not a gamble should be exclusively a matter of preferences. This idea lies in the original spirit of SD and in its further developments. Indeed, in Hanoch and Levy [17], the authors write: "The variables X and Y are defined here as the money payoffs of a given venture, which are additions (or reductions) to the individual's (constant) wealth ..."(p. 336). We highlight the fact that the property of independence from the initial endowment is easy to prove in the EU case and consequently it is not always explicitly mentioned. As a matter of fact, the following general result is crucial and it shows that the presence of a constant initial wealth is neutral to the ranking.

Lemma 1. *Assume (ρ, S) ranks the distributions and $\rho(s,\cdot)$ is monotone decreasing for every $s \in S$. Then $X_1 \succeq_{(\rho,S)} X_2$ iff $(X_1 + w) \succeq_{(\rho,S)} (X_2 + w)$ for every $w \in \mathbb{R}$.*

Proof. By Definition 2 and the monotonicity of ρ, $\rho(s,X_1) \geq \rho(s,X_2)$ for every $s \in S$ if and only if F_1 stochastically dominates F_2. Now, adding a constant w to X_i is equivalent to shift the distribution F_i by the same quantity w. Therefore, F_1 dominates F_2 iff the distribution of $X_1 + w$ is always equal of smaller than the distribution of $X_2 + w$. This is in turn equivalent to $\rho(s,X_1+w) \geq \rho(s,X_2+w)$ for every $s \in S$. $\square$

We remark that both the monotonicity and the ranking property are necessary in order to guarantee the desired equivalence. We now review some basic results of SD in the context of von Neumann-Morgenstern's Expected Utility paradigm. Apparently, the EU functional is law invariant in the sense of Definition 1. Next, we partially recover the notation from Example 1, so that we introduce the classes

$$\begin{aligned}
&\mathcal{U}_0 := \{u \in \mathcal{C}^0(\mathbb{R};\mathbb{R}) \,|\, u \text{ is non decreasing}\},\\
&\mathcal{U}_1 := \{u \in \mathcal{C}^1(\mathbb{R};\mathbb{R}) \,|\, u' > 0\},\\
&\mathcal{U}_2 := \{u \in \mathcal{C}^0(\mathbb{R};\mathbb{R}) \,|\, u \text{ is concave and non decreasing}\}.
\end{aligned}$$

The EU functional is defined as

$$\rho^{EU}(u, F_X) := \int_{\mathbb{R}} u(x)dF_X(x). \tag{1}$$

Given two distributions F_1, F_2, financial jargon usually refers to First Stochastic Dominance.[2] Using our notations, FSD is just SD as in Definition 2. It is well-known that the EU functional is monotone decreasing and $(\rho^{EU},\mathcal{U}_0)$ ranks the distributions. We slightly generalize this result to the class $\mathcal{U}_1$ proving the following proposition.[3]

Proposition 1. *$F_1 \succeq_{(\rho^{EU},\mathcal{U}_1)} F_2$ iff $F_1(x) \leq F_2(x)$ for every $x \in \mathbb{R}$.*

[2] This terminology was introduced by Hadar and Russell [16].

[3] When restricting to a smaller class of functions, the proof for the sufficient condition is straightforward. On the other hand, proving the necessary part may become mathematically overwhelming.

Proof. ($\Leftarrow$) The sufficient part readily follows by the property $u' > 0$. ($\Rightarrow$) For the necessary part, suppose $F_1(\bar{x}) > F_2(\bar{x})$ for some $\bar{x}$. By the right continuity of the distributions, there exists $n \in \mathbb{N}$ such that $F_1(x) > F_2(x)$ for every $x \in [\bar{x}, \bar{x} + \frac{1}{n}]$. Define a sequence $\{u_n\}_n \subset \mathcal{U}_1$ by

$$u_n(x) = \mathcal{N}\left(2n^2\left[x - (\bar{x} + \tfrac{1}{2n})\right]\right)$$

where $\mathcal{N}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$ is the distribution function of a standard Gaussian random variable. Clearly, similar sequences $\{u_n\}_n$ work well too. To compute $\rho^{EU}(u_n, F_1)$ we split the integral in three parts and estimate

$$\begin{aligned}\rho^{EU}(u_n, F_1) &= \int_{-\infty}^{\bar{x}} u_n(x)dF_1(x) + \int_{\bar{x}}^{\bar{x}+1/n} u_n(x)dF_1(x) + \int_{\bar{x}+1/n}^{+\infty} u_n(x)dF_1(x) \\ &\leq \mathcal{N}(-n)F_1(\bar{x}) + \mathcal{N}(n)[F_1(\bar{x} + \tfrac{1}{n}) - F_1(\bar{x})] + 1 - F_1(\bar{x} + \tfrac{1}{n}).\end{aligned}$$

Similarly, for $\rho^{EU}(u_n, F_2)$ we have

$$\rho^{EU}(u_n, F_2) \geq \mathcal{N}(-n)[F_2(\bar{x} + \tfrac{1}{n}) - F_2(\bar{x})] + \mathcal{N}(n)[1 - F_2(\bar{x} + \tfrac{1}{n})].$$

For the proof to be completed, it is sufficient to find $\hat{n} \in \mathbb{N}$ such that $\rho^{EU}(u_{\hat{n}}, F_1) - \rho^{EU}(u_{\hat{n}}, F_2) < 0$. Therefore, even if $\lim_n u_n \notin \mathcal{U}_1$ we can let n go to infinity and see what happens. Recalling the right continuity of F_i and using a standard dominated convergence argument, we obtain

$$\lim_{n\to\infty} \left[\rho^{EU}(u_n, F_1) - \rho^{EU}(u_n, F_2)\right] = F_2(\bar{x}) - F_1(\bar{x}) < 0.$$

This means that for every $\epsilon > 0$ there exists $\bar{n} \in \mathbb{N}$ such that for every $n \geq \bar{n}$ it holds

$$\rho^{EU}(u_n, F_1) - \rho^{EU}(u_n, F_2) < \epsilon + (F_2(\bar{x}) - F_1(\bar{x})).$$

Choosing $\hat{n} = \bar{n}$ and $\epsilon = \frac{1}{2}(F_1(\bar{x}) - F_2(\bar{x}))$ gives the result. □

Even if the previous finding is well-known, we preferred to give here a detailed proof since it illustrates the general scheme of the proofs concerning SD. In the same vein, the financial literature often refers to Second Stochastic Dominance as to a ranking among the integrals of the distributions. Formally, F_1 SSD F_2 if and only if $\int_{-\infty}^{x} F_1(t)dt \geq \int_{-\infty}^{x} F_2(t)dt$ holds for every $x \in \mathbb{R}$. Restricting the attention to concave functions, we see that such ordering is indeed obtained through a specific functional dominance. For a proof of the next result in a general setting, see Hanoch and Levy [17].

Proposition 2. *$F_1 \succeq_{(\rho^{EU}, \mathcal{U}_2)} F_2$ iff $\int_{-\infty}^{x}[F_2(t) - F_1(t)]dt \geq 0$ for every $x \in \mathbb{R}$.*

Remember that in the EU paradigm, a risk averse agent is characterized by a concave and non decreasing utility function. Hence, the previous result says that risk averse decision makers dislike mean-preserving spreads. We now proceed to derive analogues for Yaari' [39] Dual Theory of Choice and for Tversky and Kahneman' [37] Cumulative Prospect Theory.

2.1 Stochastic Dominance in the Dual Theory of Choice

In 1987, Yaari [39] axiomatized what he called "Dual Theory of Choice" in order to explain some of the well-known paradoxes originating from the EU paradigm. By imposing a series of dual axioms, he explicitly derived a functional representation of dual preferences. In the original setting, Yaari [39] considered only prospects whose support was included in $[0, 1]$. An immediate extension concerns lotteries with non negative compact support. In the same spirit of He and Zhou [18] we consider a generalized framework which includes non negative random variables $X : \Omega \to \mathbb{R}^+$. As in Example 1, given a distribution F_X the mathematical framing of Yaari's preference functional will be

$$\rho^Y(T, F_X) := \int_0^{+\infty} T(1 - F_X(x))dx, \tag{2}$$

where $T : [0, 1] \to [0, 1]$ is the probability distortion or weighting function.

Some remarks are in order. In his Theorem 1, Yaari [39] proved that for bounded lotteries, T must be continuous and non decreasing to preserve the order of the probabilities. As a by product, T fixes the end points, i.e. $T(0) = 0$ and $T(1) = 1$. From a mathematical viewpoint ρ^Y is a Choquet expectation, that is an integral with respect to the non additive measure $T \circ \mathbb{P}$. Observe that ρ^Y is well defined, possibly taking an infinite value. Furthermore, the integrand is non increasing and if the support of X is bounded then $\rho^Y < +\infty$. Note that the distortion T affects the decumulative probabilities and no utility function appears in (2). In fact, it can be shown that T is nothing but the inverse of a suitable utility function u. Not surprisingly, the preferences of a risk-averse agent are characterized by a convex T. To intuitively see this, observe that the inverse of a concave utility function u becomes a convex weighting function T and vice versa. Motivated by this argument, we introduce the following classes of distortions.

$$\begin{aligned}
\mathcal{T}_0 &:= \{T \in \mathcal{C}^0([0,1];[0,1]) \,|\, T(0) = 0,\, T(1) = 1,\, T \text{ is non decreasing}\},\\
\mathcal{T}_1 &:= \{T \in \mathcal{C}^1([0,1];[0,1]) \,|\, T(0) = 0,\, T(1) = 1,\, T' > 0\},\\
\mathcal{T}_2 &:= \{T \in \mathcal{C}^0([0,1];[0,1]) \,|\, T(0) = 0,\, T(1) = 1,\, T \text{ is non decreasing and convex}\}.
\end{aligned} \tag{3}$$

Since we consider only non negative prospects, in this section the distributions F_1 and F_2 are assumed to be null for $x < 0$. We also need the generalized inverse of a distribution F, namely

$$F^{-1}(t) := \sup\{x \in \mathbb{R} \,|\, F(x) \le t\}, \quad t \in [0,1], \tag{4}$$

with the convention $F^{-1}(1) := +\infty$. Observe that F^{-1} is non decreasing and *càdlàg* (right continuous with limit on the left) as well as F. Here is the first result concerning Stochastic Dominance in the DT paradigm.

Proposition 3. *$F_1 \succeq_{(\rho^Y,\, \mathcal{T}_1)} F_2$ iff $F_1(x) \le F_2(x)$ for every $x \in \mathbb{R}^+$.*

Proof. We prove here the result for the class $\mathcal{T}_0$, leaving to the Appendix the involved case of $\mathcal{T}_1$. ($\Leftarrow$) The sufficient part is a consequence of T being non decreasing, so that

$$\rho^Y(T, F_1) = \int_0^{+\infty} T(1 - F_1(x))dx \ge \int_0^{+\infty} T(1 - F_2(x))dx = \rho^Y(T, F_2).$$

($\Rightarrow$) For the necessary part, suppose $F_1(\bar{x}) > F_2(\bar{x})$ for some $\bar{x} \ge 0$. To show the contradiction, define the weighting function

$$T(x) = \begin{cases} 0 & x \in [0, 1 - F_1(\bar{x})] \\ \dfrac{x - (1 - F_1(\bar{x}))}{F_1(\bar{x}) - F_2(\bar{x})} & x \in (1 - F_1(\bar{x}), 1 - F_2(\bar{x})) \\ 1 & x \in [1 - F_2(\bar{x}), 1]. \end{cases}$$

To compute ρ^Y, we split the integral and using the generalized inverse in (4) we estimate

$$\rho^Y(T, F_1) = \int_0^{F_1^{-1}(F_2(\bar{x}))} 1\, dx + \int_{F_1^{-1}(F_2(\bar{x}))}^{\bar{x}} \frac{F_1(\bar{x}) - F_1(x)}{F_1(\bar{x}) - F_2(\bar{x})} dx \le \bar{x},$$

since the second integrand is uniformly lower than 1. For the other distribution we have

$$\rho^Y(T, F_2) = \int_0^{\bar{x}} 1\, dx + \int_{\bar{x}}^{F_2^{-1}(F_1(\bar{x}))} \frac{F_1(\bar{x}) - F_2(x)}{F_1(\bar{x}) - F_2(\bar{x})} dx.$$

Recalling that F_i are right continuous, we deduce the existence of some $\beta > 0$ such that F_1 and F_2 are continuous over $[\bar{x}, \bar{x} + \beta]$ and $F_1(\bar{x}) > F_2(x)$ for every $x \in [\bar{x}, \bar{x} + \beta]$. Observe that β does not intervene explicitly in the distortion T. Hence, defining

$$\eta := \frac{F_1(\bar{x}) - F_2(\bar{x} + \beta)}{F_1(\bar{x}) - F_2(\bar{x})} > 0, \tag{5}$$

and noticing $F_2^{-1}(F_1(\bar{x})) \ge \bar{x} + \beta$, we find

$$\rho^Y(T, F_1) - \rho^Y(T, F_2) \le - \int_{\bar{x}}^{F_2^{-1}(F_1(\bar{x}))} \frac{F_1(\bar{x}) - F_2(x)}{F_1(\bar{x}) - F_2(\bar{x})} dx \le -\beta\eta < 0. \qquad \square$$

Proposition 3 shows that ρ^Y is monotone decreasing and it ranks the distributions of non negative lotteries. Hence, the partial ordering induced by $(\rho^Y, \mathcal{T}_1)$ over the distributions coincides with that induced by $(\rho^{EU}, \mathcal{U}_1)$ for non negative random variables. Suggestively, we formalize this link as follows.

Corollary 1. $F_1 \succeq_{(\rho^Y, \mathcal{T}_1)} F_2$ *iff* $F_1 \succeq_{(\rho^{EU}, \mathcal{U}_1)} F_2$.

We now ask whether it is possible to state the equivalence of the SSD and a functional dominance relation in the DT for risk averse decision makers. If we restrict our attention to essentially bounded random variables, it turns out that the answer is positive. To prove it, we need the following technical result.

Theorem 1 (Integral Majorization Theorem). *Let $f, g : [a,b] \to \mathbb{R}$ be two given non increasing functions. The inequality*

$$\int_a^b \phi(f(x))dx \leq \int_a^b \phi(g(x))dx$$

holds for every function ϕ continuous and convex in $[a,b]$ such that the integrals exists iff

$$\int_a^s f(x)dx \leq \int_a^s g(x)dx \quad \text{for } s \in [a,b) \qquad \text{and} \qquad \int_a^b f(x)dx = \int_a^b g(x)dx.$$

The previous theorem is an extension to the continuous case of a famous inequality due to Hardy, Littlewood and Polya, also known as Karamata inequality. For a proof, see Theorem 1 in Fan and Lorentz [13], where a further generalization is shown.[4] Without loss of generality, we consider lotteries taking values over $[0, 1]$.

Proposition 4. *Assume X_1 and X_2 have support contained in $[0,1]$. Then $F_1 \succeq_{(\rho^Y, \mathcal{T}_2)} F_2$ iff $\int_0^x [F_2(t) - F_1(t)]dt \geq 0$ for every $x \in [0,1]$, with equality for $x = 1$.*

Proof. ($\Leftarrow$) Apply Theorem 1 setting $a = 0$, $b = 1$, $f(x) = 1 - F_2(x)$ and $g(x) = 1 - F_1(x)$. We deduce that for every continuous convex function $T : [0,1] \to [0,1]$ it holds $\int_0^1 T(1 - F_2(x))dx \leq \int_0^1 T(1 - F_1(x))dx$. Since we require $T(0) = 0$ and $T(1) = 1$, it follows that T must be non decreasing. In other words, $F_1 \succeq_{(\rho^Y, \mathcal{T}_2)} F_2$. ($\Rightarrow$) It is sufficient to repeat the previous argument backward. □

Perhaps the previous result is not so surprising recalling our digression concerning the properties of ρ^Y. Combining Proposition 2 and 4 we have

Corollary 2. *Assume X_1 and X_2 have support contained in $[0,1]$. Then $F_1 \succeq_{(\rho^Y, \mathcal{T}_2)} F_2$ iff $F_1 \succeq_{(\rho^{EU}, \mathcal{U}_2)} F_2$.*

Resuming our analysis we conclude that the Expected Utility paradigm and the Dual Theory of Choice are quite the same from a SD point of view. Corollary 1 and 2 tell us what distributions can be left out from the efficient sets, no matter if the agent's preferences are represented by (1) or (2). Importantly, the intuition completely lies in turning the utility function u into a probability distortion via computing the inverse u^{-1}. We now proceed to study the CPT case.

2.2 Stochastic Dominance in the Cumulative Prospect Theory

In Tversky and Kahneman [37], the authors provided the axiomatization of CPT, an improved version of the earlier PT. The main reason for the success of CPT is that it provides a sound alternative to EU. Undoubtedly, CPT is more appropriate from a descriptive viewpoint since it has been developed from experimental observations. At the same time, it retains a good analytical tractability for economic and financial modeling. This are the cornerstone of CPT.

[4] We are not aware of a similar Integral Majorization Theorem holding for unbounded intervals. Clearly, in the latter case our arguments can be readily extended.

(i) A decision maker evaluates gains and losses, i.e. deviations of the terminal wealth W^1 with respect to a benchmark, or reference wealth W^{ref}. In this setting, a random variable $Y \in \mathcal{X}$ is intended as a prospect representing possible gains or losses. Clearly, $Y(\omega) \geq 0$ will be considered a gain whereas $Y(\omega) < 0$ will represent a loss. By definition, we have $Y = W^1 - W^{ref}$.

(ii) Evaluation takes place via two distinct utility or value functions, where $u^+ : \mathbb{R}^+ \to \mathbb{R}^+$ serves for the gains and $u^- : \mathbb{R}^- \to \mathbb{R}^-$ evaluates losses. Conventionally, we set $u^+(0) = u^-(0) = 0$. Laboratory experiments suggest u^+ to be concave and u^- to be convex. Thus globally S-shaped utility functions are observed; see Figure 1. Furthermore, most of the surveyed subjects exhibit loss aversion, to be discussed later.

(iii) The decision makers assess actual probabilities correctly, but these probabilities are distorted during the decision process, similarly to what happens in the DT. Moreover, the distortions on gains are generally different from those on losses. We take T^+, T^- : $[0,1] \to [0,1]$ to be the weighting functions for the probabilities of the gains and the losses respectively. Empirical evidence suggests that people tend to overweight relatively large gains and losses of small probabilities. This results in reversed S-shaped weighting functions; see Figure 2.

Tversky and Kahneman [37] introduced CPT and the subsequent computation of the prospect value in the discrete case, i.e. when the lotteries take only a finite number of values. A continuous time functional representation of CPT preferences is as follows. Define a vector of parameters $v := (u^+, u^-, T^+, T^-)$. For a lottery Y with distribution F_Y, the CPT preference functional turns out to be

$$\rho^{CPT}(v, F_Y) := \rho_+^{CPT}(v, F_Y) + \rho_-^{CPT}(v, F_Y), \qquad \text{where} \tag{6}$$

$$\rho_+^{CPT}(v, F_Y) := \int_0^{+\infty} u^+(x) d[-T^+(1 - F_Y(x))],$$
$$\rho_-^{CPT}(v, F_Y) := \int_{-\infty}^0 u^-(x) d[T^-(F_Y(x))].$$

Let us discuss over (6). ρ^{CPT} is defined as the algebraic sum of the utility from the gains and the disutility from the losses. In turn, they are computed as Choquet integrals where u^+, u^- distort payoffs and T^+, T^- distort probabilities. Moreover, T^+ acts over the decumulative distribution whereas T^- affects the cumulative one. In this paradigm, the value functions and the distortions are complementary in determining the behavior toward risk.

Observe that in (6) the presence of the reference wealth is not explicitly seen. Actually, it is incorporated in the prospect Y and it evidently influences F_Y. Therefore, the benchmark is an element that originates from the agent's preferences but it changes the way the distributions of the lotteries are perceived. For the moment, we are not interested in an explicit form for W^{ref}, nor for W^1. Henceforth, we do not include the reference wealth in the vector of parameters v and we consider Y as a given net wealth. Moreover, the reference wealth should not be confused with the initial wealth. As we proved in Lemma 1, the initial wealth does not affect the ordering of the distributions whenever ρ is monotone decreasing and ranks the distributions in a suitable class. Since ρ^{CPT} satisfies these requirements (we will prove it in a moment), we can think of an arbitrary initial wealth W^0 to be added to Y. The presence of W^0 amounts to a shift in F_Y but it leaves unchanged the ranking of the distributions and the resulting efficient set.

Loss aversion deserves a special treatment. Loss aversion is easy to explain using familiar language by saying that "a loss hurts more than an equivalent gain". A straightforward formulation would be $u^+(x) < u^-(-x)$ for every $x > 0$. The extra condition $\lambda := \lim_{x \to 0^+} (u^+)'(x) / \lim_{x \to 0^-} (u^-)'(x) > 1$ in sometimes employed to model loss aversion against small stakes. A stronger condition that encompasses the previous two is $(u^+)'(x) < (u^-)'(-x)$ for every $x > 0$, as Tversky and Kahneman already assumed in the original version of CPT [37]. The meaning is quite clear: At any wealth level, the marginal disutility suffered through a loss is greater than the marginal utility obtained through an equivalent gain. We remark that there are many other ways to frame loss aversion. For an overview, we refer the interested reader to Schmidt and Zank [36] and Köbberling and Wakker [23].

Remark 2. Unfortunately, equation (6) can be ill-defined. Even thought we are using Lebesgue integrals, it can happen $\rho_+^{CPT}(v, F_Y) = +\infty$ and $\rho_-^{CPT}(v, F_Y) = -\infty$ for some choices of (v, F_Y). When writing $\rho^{CPT}(v, F_Y)$ to study SD relations, we will assume that it is well-defined. Notably, He and Zhou [18] and Pirvu and Schulze [33] give sufficient conditions over (v, F_Y) for (6) to be well-posed. Nonetheless, if u^+ and u^- are bounded then we can integrate by parts and obtain

$$\rho^{CPT}(v, F_Y) = \int_0^{+\infty} T^+(1 - F_Y(x))du^+(x) - \int_{-\infty}^0 T^-(F_Y(x))du^-(x),$$

as Barberis and Huang do in [1]. Moreover, if u^+ and u^- are strictly increasing, via a change of variable we get

$$\rho^{CPT}(v, F_Y) = \int_0^{+\infty} T^+(\mathbb{P}(u^+(Y) > x))dx - \int_{-\infty}^0 T^-(\mathbb{P}(u^-(Y) \leq x))dx,$$

as in Jin and Zhou [20]. We prefer to use (6) because it is the natural extension to the continuous case of the original formulation of CPT.

We now see how SD relates to CPT. To this end, we introduce the following classes.

$$\begin{aligned} \mathcal{U}_0^+ &:= \{u^+ \in \mathcal{C}^0(\mathbb{R}^+; \mathbb{R}^+) \,|\, u^+(0) = 0,\ u^+ \text{ is non decreasing}\}, \\ \mathcal{U}_0^- &:= \{u^- \in \mathcal{C}^0(\mathbb{R}^-; \mathbb{R}^-) \,|\, u^-(0) = 0,\ u^- \text{ is non decreasing}\}, \end{aligned} \tag{7}$$

and we set $\mathcal{V}_0 := \mathcal{U}_0^+ \times \mathcal{U}_0^- \times \mathcal{T}_0 \times \mathcal{T}_0$, where $\mathcal{T}_0$ is defined in (3). It is clear that the class $\mathcal{V}_0$ is too large to give a realistic description of CPT. For example, identically null value functions are permitted as well as weighting functions different from the observed reversed S-shaped ones. In such cases, the economic meaning of CPT is completely lost. However, we introduce $\mathcal{V}_0$ because it is the largest class where one can give a sufficient condition for ρ^{CPT} to be monotone decreasing. Henceforth, the same condition will be sufficient for any $\mathcal{V} \subset \mathcal{V}_0$. We formalize this fact in the following proposition.

Proposition 5. *$F_1 \succeq_{(\rho^{CPT}, \mathcal{V}_0)} F_2$ iff $F_1(x) \leq F_2(x)$ for every $x \in \mathbb{R}$.*

Proof. ($\Leftarrow$) For the gain part, note that if $\rho_+^{CPT}(v, F_1) = +\infty$ then there is nothing to prove. Hence, assume $\rho_+^{CPT}(v, F_1) < +\infty$ and for $N \in \mathbb{N}$ define

$$u_N^+(x) = \begin{cases} u^+(x) & x \in [0, N], \\ u^+(N) & x \in (N, +\infty). \end{cases}$$

Noticing that u_N^+ is non decreasing and bounded, we can integrate by parts and obtain

$$\begin{aligned} &\int_0^{+\infty} u_N^+(x) d[T^+(1 - F_2(x)) - T^+(1 - F_1(x))] \\ &\qquad = u_N^+(x)[T^+(1 - F_2(x)) - T^+(1 - F_1(x))]_{x=0}^{x=+\infty} \\ &\qquad\quad + \int_0^{+\infty} [T^+(1 - F_2(x)) - T^+(1 - F_1(x))] du_N^+(x) \\ &\qquad = \int_0^N [T^+(1 - F_2(x)) - T^+(1 - F_1(x))] du_N^+(x) \geq 0, \quad N \in \mathbb{N}. \end{aligned}$$

By a standard monotone convergence argument, we find $\rho_+^{CPT}(v, F_1) - \rho_+^{CPT}(v, F_2) \geq 0$. On the loss side, we can do a similar analysis whenever $\rho_-^{CPT}(v, F_2) < +\infty$, setting

$$u_N^-(x) = \begin{cases} u^-(x) & x \in [-N, 0], \\ u^-(-N) & x \in (-\infty, -N). \end{cases}$$

If follows

$$\begin{aligned}
&\int_{-\infty}^{0} u_N^-(x)d[T^-(F_1(x)) - T^-(F_2(x))] \\
&\quad = u_N^-(x)[T^-(F_1(x)) - T^-(F_2(x))]_{x=-\infty}^{x=0} + \int_{-\infty}^{0} [T^-(F_2(x)) - T^-(F_1(x))]du_N^-(x) \\
&\quad = \int_{-N}^{0} [T^-(F_2(x)) - T^-(F_1(x))]du_N^-(x) \geq 0, \quad N \in \mathbb{N}.
\end{aligned}$$

Again, by monotone convergence we have $\rho_-^{CPT}(v, F_1) - \rho_-^{CPT}(v, F_2) \geq 0$ and we conclude.

($\Rightarrow$) Assume $F_1(\bar{x}) > F_2(\bar{x})$ for some $\bar{x} \geq 0$. So there is a $\beta > 0$ such that F_1, F_2 are continuous and $F_1(x) > F_2(x)$ for $x \in [\bar{x}, \bar{x}+\beta]$. Observe that choosing $T^+(x) = x$, $u^-(x) \equiv 0$ and any $T^- \in \mathcal{T}_0$ puts us in a situation analogous to the EU paradigm. Defining

$$u^+(x) = \begin{cases} 0 & x \in [0, \bar{x}], \\ \frac{x-\bar{x}}{\beta} & x \in (\bar{x}, \bar{x}+\beta), \\ 1 & x \in [\bar{x}+\beta, +\infty), \end{cases}$$

we find a contradiction since

$$\rho^{CPT}(v, F_1) - \rho^{CPT}(v, F_2) = \frac{1}{\beta}\int_{\bar{x}}^{\bar{x}+\beta} [F_2(x) - F_1(x)]dx < 0.$$

Similarly, if $\bar{x} < 0$ there exists $\beta > 0$ such that $\bar{x}+\beta < 0$, F_1, F_2 are continuous and $F_1(x) > F_2(x)$ for $x \in [\bar{x}, \bar{x}+\beta]$. Then we take $u^+(x) \equiv 0$, $T^+ \in \mathcal{T}_0$, $T^-(x) = x$ and

$$u^-(x) = \begin{cases} -1 & x \in (-\infty, \bar{x}], \\ \frac{x-(\bar{x}+\beta)}{\beta} & x \in (\bar{x}, \bar{x}+\beta), \\ 0 & x \in [\bar{x}+\beta, 0]. \end{cases}$$

This leads to the estimate $\rho^{CPT}(v, F_1) - \rho^{CPT}(v, F_2) < 0$ as before. $\square$

Clearly, we are not the firsts in proving that CPT preferences are strongly related to SD. Using a completely different language, Proposition A1 in Barberis and Huang [1] shows in a similar setting that ρ^{CPT} is monotone decreasing. However, they assume u^+, u^-, T^+, T^- to be strictly increasing and continuous and that integration by parts can be applied as illustrated in Remark 2. We thus provide a further extension of this result taking a larger class of preference functionals. Moreover, we do not need to apply integration by parts and we claim that an analogous result holds even relaxing the continuity assumption over the parameters.

Now, the interesting question is: Can we obtain the same ordering over the distributions if we select a class smaller than $\mathcal{V}_0$? If possible, we would like to select such class in order to satisfy the salient features of CPT: S-shaped utility function, reversed S-shaped weighting functions and loss aversion. To this end, we introduce the following classes of value functions.

$$\begin{aligned}
\mathcal{U}_2^+ &:= \{u \in \mathcal{C}^1(\mathbb{R}^+; \mathbb{R}^+) \,|\, u^+(0) = 0,\ (u^+)' > 0 \text{ and } u^+ \text{ is concave}\}, \\
\mathcal{U}_2^- &:= \{u \in \mathcal{C}^1(\mathbb{R}^-; \mathbb{R}^-) \,|\, u^-(0) = 0,\ (u^-)' > 0 \text{ and } u^- \text{ is convex}\}.
\end{aligned} \tag{8}$$

Next, we specify the class of reversed S-shaped probability distortions.

$$\mathcal{T}_{RS} := \left\{ T \in \mathcal{C}^1([0,1];[0,1]) \;\middle|\; \begin{array}{c} T(0) = 0,\ T(1) = 1,\ T' \geq 0, \\ T(x) > x \text{ in a neighborhood of } x=0, \\ T(x) < x \text{ in a neighborhood of } x=1, \\ T \text{ crosses once the 45-degree line.} \end{array} \right\} \tag{9}$$

Hence, a function $T \in \mathcal{T}_{RS}$ is strictly concave near the origin and strictly convex near $x = 1$. The monotonicity and the single crossing property ensure that T look like those in Figure 2. We now explicitly model loss aversion via the strong requirement

$$(u^+)'(x) < (u^-)'(-x) \quad \text{for every } x > 0. \tag{10}$$

In this way, the class of parameters that we consider is

$$\mathcal{V}_1 := \mathcal{U}_2^+ \times \mathcal{U}_2^- \times \mathcal{T}_{RS} \times \mathcal{T}_{RS},$$

where u^+, u^- satisfy (10). We are now in a position to state the main result for CPT preferences.

Proposition 6. $F_1 \succeq_{(\rho^{CPT},\mathcal{V}_1)} F_2$ *iff* $F_1(x) \leq F_2(x)$ for every $x \in \mathbb{R}$.

Proof. See the Appendix. □

Apparently, Proposition 6 states that $(\rho^{CPT}, \mathcal{V}_1)$ effectively ranks the distributions. To the best of our knowledge, the previous result is completely new to the literature. We remark that it exploits the CPT paradigm in its full significance. Moreover, it will naturally suggest the family of distributions that cause ρ^{CPT} to depend only one the mean and the variance of the prospect. More important, we see that the same ordering over the distributions is induced in the EU paradigm with the class $\mathcal{U}_1$. Consequently, the two criteria identify the same efficient sets, confirming that EU and CPT decision makers are not so different as it could seem at first sight. We resume this analysis combining Proposition 1 and 6 to obtain

Corollary 3. $F_1 \succeq_{(\rho^{CPT},\mathcal{V}_1)} F_2$ *iff* $F_1 \succeq_{(\rho^{EU},\mathcal{U}_1)} F_2$.

Before focusing on the link between SD and mean-variance analysis, we need to review the basic theory of some classes of distributions that deserve a special place in the portfolio choice literature.

3 Spherical, elliptical and location-scale distributions

We start considering the following well-known definition.

Definition 3 (Spherical distribution)**.** *The n-dimensional random vector X is spherically distributed about the origin iff $X \stackrel{d}{=} RX$ for every orthogonal matrix $R \in \mathbb{R}^{n\times n}$.*

In words, X is spherically distributed about the origin if its distribution is preserved when undergoing invariant orthogonal linear transformations that leave the origin fixed. It follows from the definition that the characteristic function of X depends only on the norm of t, $t \in \mathbb{R}^n$. If X has a density, then such density will depend only on the norm of X.

Definition 4 (Elliptical distribution)**.** *Let $\Delta \in \mathbb{R}^n$ be a given vector and $\Sigma \in \mathbb{R}^{n\times n}$ be a given symmetric and positive definite matrix. The n-dimensional random vector X is elliptically distributed, $X \in E_n(\Delta, \Sigma; \Psi)$, iff the characteristic function of X has the form $\mathbb{E}[\exp(it'X)] \equiv C_X(t) = \Psi(t'\Sigma t)\exp(it'\Delta)$ for some real function Ψ.*

The pair (Δ, Σ) is the parametric part of the distribution, whereas Ψ is the non-parametric part, also called the characteristic generator. These two parts are uncoupled in the sense that the same parameters can lead to different distributions since they have distinct generators. Here is a list of the main facts concerning elliptical distributions. Let $X \in E_n(\Delta, \Sigma; \Psi)$; for the proofs, we refer the reader to Kelker [22].

(F1) (Density) If X has a density f_X, then $f_X(x) = c_n|\Sigma|^{-1/2}\phi((x-\Delta)'\Sigma^{-1}(x-\Delta))$ for some function ϕ which does not depend on n. ϕ is called the density generator. Observe that f_X is symmetric since Σ is positive definite.

(F2) (Moments) If X has finite first moment, then $\mathbb{E}[X] = \Delta$. If X has finite second moment, then $\mathbb{V}(X) = \gamma\Sigma$, where $\gamma \equiv -2\Psi'(0) \geq 0$ does not depend on the parametric part. If $\Delta = 0$ and $\Sigma = I_n$ is the identity matrix, then γ is the variance of every univariate marginal distribution.

(F3) (Linear combinations) For any matrix $T \in \mathbb{R}^{m\times n}$ of rank m, $m \leq n$, the m-dimensional random vector TX is elliptical and $TX \in E_m(T\Delta, T\Sigma T'; \Psi)$. In other words, linear combinations of the elements of X are still elliptical. Notice that the non-parametric part remains unchanged. As a particular case, we find that every subset $(X_{i_1}, \ldots, X_{i_k})$ of X, $k \leq n$, is elliptical too.

(F4) (Regression) Set $X = (X_1, X_2)$, with X_1 a m-dimensional random vector. Identify partitions of Δ and Σ as (Δ_1, Δ_2) and $(\Sigma_{11}, \Sigma_{12}, \Sigma_{21}, \Sigma_{22})$ respectively. Then, if the conditional mean of X_1 given $X_2 = x_2$ exists, it is given by $\mathbb{E}[X_1|X_2 = x_2] = \Delta_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \Delta_2)$. Hence, regression is linear for the elliptical distributions and conditional means depend only on the parametric part. If X_1 and X_2 are uncorrelated, then $\mathbb{E}[X_1|X_2] = \Delta_1$ almost surely.

Following Owen and Rabinovitch [32], we introduce the class of multivariate location-scale distributions.

Definition 5 (Location-Scale distributions). *Let $\Delta \in \mathbb{R}^n$ be a given vector and $\Sigma \in \mathbb{R}^{n\times n}$ be a given symmetric and positive definite matrix. Let X be a n-dimensional random vector such that $\Delta = \mathbb{E}[X]$ and the random variable $Z := (\alpha' X - \mathbb{E}[\alpha' X])/(k\sqrt{\alpha'\Sigma\alpha})$ has a density $f(z, \alpha)$ for any vector $\alpha \in \mathbb{R}^n$ and for any positive scalar k. Then X has a location-scale distribution iff $f(z, \alpha)$ does not depend on α.*

We now explain the deep links between the families of distributions previously introduced.

Lemma 2. *For a n-dimensional random vector X, the following relationships hold.*

(i) *$X \in E_n(\Delta, \Sigma; \Psi)$ for some Σ and Ψ iff it exists a non-singular matrix $T \in \mathbb{R}^{n\times n}$ such that $Z := T(X - \Delta)$ is spherically distributed about the origin.*

(ii) *$X \in E_n(0, \Sigma; \Psi)$ for some Σ and Ψ iff $\alpha_1' X \overset{d}{=} \alpha_2' X$ for every $\alpha_1, \alpha_2 \in A$, where* $A = \{\alpha \in \mathbb{R}^n - \{0\} \,|\, \mathbb{V}(\alpha' X) \text{ is constant}\}$.[5]

(iii) *Let X admit a density. Then $X \in E_n(\Delta, \Sigma; \Psi)$ iff X has a location-scale distribution.*

Proof. *(i)* ($\Rightarrow$) If $X \in E_n(\Delta, \Sigma; \Psi)$, then $C_{X-\Delta}(t) = \Psi(t'\Sigma t)$ and $C_Z(t) = \Psi(t'T\Sigma T't)$. Hence, for Z to be spherically distributed about the origin it suffices to find T such that $T\Sigma T' = I_n$. We can check that $T = D^{-1/2}Q^{-1}$ does the job, where Q is the orthogonal matrix (made by eigenvectors of Σ) such that $\Sigma = QDQ'$, D diagonal, and $D^{-1/2}$ is a square-root of D, i.e. $D^{-1/2}$ satisfies $D^{-1/2}D^{-1/2} = D$.

($\Leftarrow$) Conversely, assume Z is spherically distributed about the origin. By Definition 3, it follows $C_Z(t) = C_{RZ}(t) = C_Z(R't)$ for every orthogonal matrix R. Hence, $C_Z(t)$ must be of the form $\Psi(t'Mt)$ for some matrix M and some Ψ. Now, $C_X(t) = C_{T^{-1}Z+\Delta}(t) = \exp(it'\Delta)\Psi(t'\Sigma t)$ for some Σ.

(ii) Suppose $X \in E_n(0, \Sigma; \Psi)$ and fix $\alpha_1, \alpha_2 \in A$. Then we have $C_{\alpha_1' X}(t) = C_X(t\alpha_1) = \Psi(t\alpha_1'\Sigma\alpha_1 t) = \Psi(t\alpha_2'\Sigma\alpha_2 t) = C_X(t\alpha_2) = C_{\alpha_2' X}(t)$, for every $t \in \mathbb{R}$. Vice versa, if $C_{\alpha X}(t)$ is the same for every $\alpha \in A$, then the previous chain of equalities shows that $C_X(t)$ must be a function solely of the quadratic form $t'\Sigma t$ for some Σ.

(iii) See Owen and Rabinovitch [32], Proposition 2. □

Without loss of generality, one can assume that X has a density. In fact, Lemma 1 in Kelker [22] states that unless the parent spherical distribution has an atom of positive weight at the origin, all the marginal distributions of X will have densities. Elliptical distributions are also linked to stable distributions. In particular, symmetric stable distributions are elliptical. However, there are stable distributions which are not elliptical (e.g. the Lévy distribution). Similarly, there are elliptical distributions which are not stable (e.g. the logistic distribution). Thus, the two classes properly overlap. See Owen and Rabinovitch [32] for more details.

The families of distributions that we introduced, together with the stable one, are closely connected to the issue of stability of distributions with respect to their linear combinations. Generally, distributions are not closed in this sense. However, (F3) shows that elliptical distributions are and Lemma 2 strengthens this fact. Surprisingly, if X has second moment the following results hold.

Theorem 2 (Chamberlain [7], Theorem 1). *The distribution of $\alpha' X + c$ is determined by its mean and variance for every α, c iff there is a non-singular matrix T such that $z = T(x - \mathbb{E}[X])$ is spherically distributed about the origin.*

[5] In Owen and Rabinovitch [32], this statement is actually a definition of $E_n(\Delta, \Sigma; \Psi)$. In such case, one should point out that the vector Δ can be seen as an additional degree of freedom.

Theorem 3 (Chamberlain [7], Theorem 2)**.** *If $\mathbb{E}[X] \neq 0$, then the distribution of $\alpha' X$ is determined by its mean and variance for every α iff there is a non-singular matrix T such that $TX = (m, Z)$, where, conditional on m, Z is spherically distributed about the origin (here m is a scalar random variable and Z is a $(n-1)$-dimensional random vector).*

If $\mathbb{E}[X] = 0$, then the necessary and sufficient condition is that $z = TX$ be spherically distributed about the origin for some non-singular T.

Chamberlain [7] underlines that in Theorem 3 the marginal distribution of m is unrestricted, so that one has an additional degree of freedom. In fact, Theorem 2 is somewhat more powerful than Theorem 3 because the necessary condition in Theorem 2 implies the sufficient condition in Theorem 3. E.g., one can use regression as explained in (F4) to find (m, Z).

Observe that throughout this section we never referred to the preferences of some decision maker nor to utility functions or whatever. This means that previously stated facts can be applied to any preference paradigm. Especially Theorem 2 and 3 by Chamberlain [7], which are reminiscent of MV analysis in the EU case, can be applied in more general contexts. This will be the issue of the following section.

4 Mean-Variance preferences and elliptical distributions

Throughout this section, we will assume that every random variable has finite second moment, that is $\mathcal{X} \subset L^2(\Omega, \mathscr{F}, \mathbb{P})$. Hence it makes sense to conduct a Mean-Variance analysis concerning the preference functionals. We begin with a standard and intuitive definition.

Definition 6 (Mean-Variance preference functional)**.** *Let ρ represent some preferences over $\mathcal{X}$.*

- *$\rho(s, \cdot)$ is a function of mean and variance iff $\rho(s, X_1) = \rho(s, X_2)$ whenever $\mathbb{E}[X_1] = \mathbb{E}[X_2]$ and $\mathbb{V}[X_1] = \mathbb{V}[X_2]$.*
- *ρ is a function of mean and variance in the class S (ρ is MV in S) iff $\rho(s, \cdot)$ is a function of mean and variance for every $s \in S$.*

By the previous definition, it follows that if ρ is MV in S, then ρ is MV in S' for every $S' \subset S$. Evidently, if ρ is MV in S, then ρ is law invariant in S but not vice versa. Furthermore, for a given distribution F_X we can write $\rho(s, F_X) = h(s, \mathbb{E}[X], \mathbb{V}[X])$ for some $h : S \times \mathbb{R} \times \mathbb{R}^+$. To see a case where the previous definition turns out to be useful, consider EU. We can apply Proposition 1 to conclude that ρ^{EU} is MV in $\mathcal{U}_1$ iff $X_1 \overset{d}{=} X_2$ whenever $\mathbb{E}[X_1] = \mathbb{E}[X_2]$ and $\mathbb{V}[X_1] = \mathbb{V}[X_2]$.

We also remark that ρ being MV does not imply that it is increasing in the mean and decreasing in the variance, as often assumed in portfolio choice literature. Chamberlain [7] shows that if $\rho^{EU}(u, F_X) \equiv h(u, \mathbb{E}[X], \mathbb{V}[X])$ and u is concave, then h is a decreasing function of $\mathbb{V}[X]$. Conversely, he provides an example where h neither is increasing in the mean nor is quasi-concave. For a similar result in the CPT paradigm, see Example 2 in the Appendix. For the more intuitive case of ρ^{CPT} increasing in the mean and decreasing in the variance, see Del Vigna [11] where a list of various examples is given.

Another way for ρ to be MV in some class S is to suppose that the distribution of X is of a particular type. For example, if X is normally distributed then ρ^{EU} is MV in $\mathcal{U}_0$. Notably, in this case we do not have to impose specific assumptions on S. Everything we need is $u(X) \in L^2(\Omega)$ for some $u \in \mathcal{U}_0$. In the EU paradigm, these well-known facts are the starting point in the characterization of those distributions that imply MV utility functions, namely the elliptical ones. We now extend this argument further. As usual, we take the random variables X_i to have distribution F_i. The following lemma will be crucial.

Lemma 3. *Assume (ρ, S) ranks the distributions.*

- *If ρ is law invariant in S and $X_1 \succeq_{(\rho,S)} X_2$ then ρ is MV in S.*
- *If ρ is MV in S then ρ is law invariant in S and $X_1 \overset{d}{=} X_2$ whenever $\mathbb{E}[X_1] = \mathbb{E}[X_2]$ and $\mathbb{V}[X_1] = \mathbb{V}[X_2]$.*

Proof. *(i)* First recall that (ρ, S) ranking the distributions and $X_1 \succeq_{(\rho,S)} X_2$ imply that F_1 stochastically dominates F_2. If the mean and the variance of X_1 and X_2 coincide then $X_1 \stackrel{d}{=} X_2$. Next, the law invariance of ρ leads to $\rho(s, X_1) = \rho(s, X_2)$ for every $s \in S$.
(ii) If $\mathbb{E}[X_1] = \mathbb{E}[X_2]$, $\mathbb{V}[X_1] = \mathbb{V}[X_2]$ and ρ is MV in S, then $\rho(s, X_1) = \rho(s, X_2)$ for every $s \in S$. In particular, we have $X_1 \succeq_{(\rho,S)} X_2$ which implies $X_1 \stackrel{d}{=} X_2$ since $\mathbb{E}[X_1] = \mathbb{E}[X_2]$. □

The very important implication of Lemma 3 is that if (ρ, S) ranks the distributions, then we can unambiguously identify a distribution once we know its mean and variance. Combining this result with Theorem 2 and 3, we are able to characterize the distributions that imply MV preference functionals. Since the main application of such results concerns optimal portfolio choice, we shall put ourselves in that framework. Note that in spite of the hypotheses over the market structure, our analysis remains valid for analogous linear problems.

For, assume there is a frictionless market with n risky securities whose net return over the period is given by the random variable X. We set $\Delta := \mathbb{E}[X]$ and $\Sigma := \mathbb{V}[X]$, where Σ is a non singular matrix. We shall consider a riskless asset with net return $r \in \mathbb{R}$. The admissible set of portfolios will be

$$\mathcal{A} := \{p \equiv \alpha' X + \alpha_0 r \mid \alpha \in \mathbb{R}^n, \alpha_0 \in \mathbb{R}\} \subset \mathcal{X}, \tag{11}$$

with the convention $\alpha_0 r \equiv 0$ if the riskless investment is not permitted. Note that a portfolio is identified by the pair (α, α_0). We set no short-sale constraints, so the choice of an individual whose initial endowment is $W^0 \in \mathbb{R}$ must only satisfy the budget constraint $C(p) := \alpha' \mathbf{1} + \alpha_0 r \leq W^0$, where $\mathbf{1}$ is the $n \times 1$ vector of ones. Let $\rho_{|\mathcal{A}}$ be the restriction of ρ to the set of attainable portfolios. First, consider the case where the riskfree asset is available. We are able to give a characterization of the distributions that imply MV preferences in the case of law invariant functional representations. This is the main result of our paper.

Theorem 4. *Assume $\rho_{|\mathcal{A}}$ is law invariant in S and $(\rho_{|\mathcal{A}}, S)$ ranks the distributions. Then*

$$\rho_{|\mathcal{A}} \text{ is MV in } S \text{ iff } X \text{ is elliptically distributed.}$$

Proof. $(\Rightarrow)$ If $X \in E_n(\Delta, \Sigma; \Psi)$ then every admissible portfolio satisfies $p \in E_1(\alpha'\Delta + \alpha_0 r, \alpha'\Sigma\alpha; \Psi)$. By Theorem 2 it follows that the distribution of p is identified by its mean and variance. Note that here the assumption of $(\rho_{|\mathcal{A}})$ ranking the distributions is superfluous. $(\Leftarrow)$ By Lemma 3, *(ii)*, whenever two portfolios have the same mean and variance, it follows that they have the same distribution too. Therefore, the distribution of every p is characterized by its mean and variance. By Theorem 2 we conclude. Here the assumption of $\rho_{|\mathcal{A}}$ being law invariant is not necessary. □

The previous theorem is a powerful extension of the characterization shown by Chamberlain [7]. Just as an example, it is possible to slightly adjust our arguments to encompass continuous versions of the Subjective Expected Utility by Savage, Rank Dependent Utility by Quiggin and Choquet Expected Utility, as considered in Diecidue and Wakker [12]. If the riskless asset is not available and $\Delta \neq 0$ we can apply Theorem 3 and see that a result similar to Theorem 4 holds but the conclusion will be: $\rho_{|\mathcal{A}}$ is MV in S iff there is a non-singular matrix T such that $TX = (m, Z)$, where, conditional on m, Z is spherically distributed about the origin. However, if $\Delta = 0$ then Theorem 3 leads again to elliptically distributed stocks' returns.

Unfortunately, the interesting case of CPT is not encompassed in our Theorem 4. This is because the CPT functional evaluate gains/losses instead of admissible portfolios in $\mathcal{A}$. Intuitively, the reference wealth shifts the set of the distributions of the attainable terminal lotteries so that considering $\rho^{CPT}_{|\mathcal{A}}$ is no more sensible. This case deserves a separate analysis due to its theoretical and practical relevance. Incidentally, we shall answer a conjecture advanced by Pirvu and Schulze [33]. As in that article, let the reference wealth of the decision maker be

$$W^{ref} := a'X + b'\alpha + c, \tag{12}$$

for some $a, b \in \mathbb{R}^n$ and $c \in \mathbb{R}$. The coefficients a, b, c possibly depend on some of the parameters regarding the initial wealth, the market structure or even the portfolio. In other words, they are allowed to depend on α, α_0, W^0, Δ, Σ and r.

In this way the terminal wealth of the agent will be $W^1 := W^0 + p$. Following CPT, the gains/losses that the agent shall evaluate are of the form $Y = W^1 - W^{ref}$. Using (12), we find

$$Y = (\alpha - a)'X + \delta, \qquad \delta := \begin{cases} W^0 + \alpha_0 r - c - b'\alpha & \text{if the riskfree asset is available,} \\ W^0 - c - b'\alpha & \text{otherwise.} \end{cases} \tag{13}$$

We remark that if riskfree investment is allowed, the budget constraint holds with equality since ρ^{CPT} is monotone decreasing.[6] We now give a sufficient condition for ρ^{CPT} to be MV in $\mathcal{V}_0$. As we already observed, ρ^{CPT} will be MV in $\mathcal{V}_1$ too. Moreover, the result holds independently on the availability of the riskfree asset.

Proposition 7. *Assume W^{ref} as in* (12)*. If X is elliptically distributed then ρ^{CPT} is MV in $\mathcal{V}_0$.*

Proof. If $a = \alpha$ there is nothing to prove since Y is constant. So let $a \neq \alpha$. Then $Y = \gamma' X + \delta$ for some $\gamma \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$, where δ is possibly null. By Lemma 2, *(i)*, and Theorem 2 it follows that the distribution of Y is characterized by its mean and variance. Therefore, ρ^{CPT} is a function that depends uniquely on $\mathbb{E}[Y]$ and $\mathbb{V}[X]$. $\square$

We have a number of observations concerning the previous proposition.

1. A deeper analysis shows that if $a \neq \alpha$, $\Delta \neq 0$ and $\delta = 0$ then it is sufficient to assume the existence of a non-singular matrix T such that $TX = (m, Z)$ with Z spherically distributed about the origin conditional on m; see Theorem 3. Notably, this assumption is weaker than $X \in E_n(\Delta, \Sigma; \Psi)$. The condition $\delta = 0$ requires some coordination between the reference wealth W^{ref}, the initial wealth W^0 and the availability of the riskfree asset.

2. Some particular cases of the previous observation turn out to be interesting. Assume $a = b = 0$ so that the reference wealth is constant and independent on the portfolio. If there is a riskless asset and $c = W^0 + \alpha_0 r$ then the attainable gains/losses are exclusively of the form $Y = \alpha' X$. Hence the decision maker denies herself the evaluation of a riskfree investment but not the opportunity to safely invest. Conversely, if there is not a riskless asset and $c \neq W^0$ then $Y = \alpha' X + \delta$, $\delta \neq 0$. Consequently, the investor preferences reflect some riskless investment evaluation without an effective riskfree investment. In the latter case we must also require X to be elliptical distributed. Resuming, the benchmark wealth possibly distorts the way the distribution of the terminal prospects are perceived as we already observed.

3. If we are aware of the initial wealth W^0 and the parameters a, b, c that describe the reference wealth, then the distribution of Y can be characterized by the mean and the variance of the portfolio. Specifically, we have $Y \in E_1(\mathbb{E}[p] + W^0 - b'\alpha - c - a'\Delta, \mathbb{V}[p] + a'\Sigma a - 2\alpha'\Sigma a; \Psi)$.

4. W^{ref} as in (12) makes sense since it involves some quantities the investor is aware of. However, we could suppose W^{ref} to be different from (12).[7] In such case, it is clear that ρ^{CPT} can fail to be MV in $\mathcal{V}_0$. Notably, if W^{ref} is independent of X and X is elliptically distributed then we retain the MV form for ρ^{CPT}. Indeed, we have $p \in E_1(\alpha'\Delta, \alpha'\Sigma\alpha; \Psi)$ so that the distribution of p depends only on its mean and variance. Furthermore, we compute $F_Y(x) = \int_{\mathbb{R}} F_p(x - \alpha_0 r - W^0 - z) dF_{W^{ref}}(z)$, where F_p and $F_{W^{ref}}$ are the distribution of p and W^{ref} respectively. Thus, F_Y is a function of $\mathbb{E}[p]$ and $\mathbb{V}[p]$, which in turn are given by $\mathbb{E}[p] = \mathbb{E}[Y] - W^0 + \mathbb{E}[W^{ref}]$ and $\mathbb{V}[p] = \mathbb{V}[Y] - \mathbb{V}[W^{ref}]$.

The complete converse of Proposition 7 does not hold. Intuitively, the reason lies in the additional degrees of freedom brought by the reference wealth in (12). Consequently, we distinguish three cases in the following proposition.

---

[6] Del Vigna [11] shows that in a portfolio optimization problem for a CPT agent, we can find as optimal solution the pair $(\alpha, \alpha_0) = (0, W^0)$, that is null risky investment by the CPT agent. However, investment in every portfolio $(\alpha, \alpha_0) \in \mathbb{R}^{n+1}$ is still permitted.

[7] The literature offers a variety of specifications for the reference wealth. In most cases, W^{ref} is assumed to be constant; see Bernard and Ghossoub [3] and He and Zhou [18] among others. Alternatively, it can be stochastic as in Jin and Zhou [20] or even endogenous as in Kőszegi and Rabin [24]. We remark that the choice of a constant benchmark is classical in equilibrium models with CPT investors; see De Giorgi, Hens, Levy [8] and Del Vigna [11] among others.

Proposition 8. *Assume W^{ref} as in* (12) *and ρ^{CPT} is MV in $\mathcal{V}_1$.*

(i) If $a = \alpha$ then the distribution of X is unrestricted.

(ii) If $a \neq \alpha$, $\Delta \neq 0$ and $\delta = 0$ then there is a non-singular matrix T such that $TX = (m, Z)$, where, conditional on m, Z is spherically distributed about the origin.

(iii) If ($a \neq \alpha$ and $\delta \neq 0$) or ($a \neq \alpha$, $\Delta = 0$ and $\delta = 0$) then X is elliptically distributed.

Proof. By Proposition 6, $(\rho^{CPT}, \mathcal{V}_1)$ ranks the distributions. Applying Lemma 3, *(ii)*, we find that the mean and the variance of Y characterize its distribution. *(i)* If $a = \alpha$ then Y is constant so that ρ^{CPT} is MV no matter how X is distributed. *(ii)* With these hypotheses, we have $Y = (\alpha - a)'X$. Thus we apply Theorem 3. *(iii)* In this case $Y = (\alpha - a)'X + \delta$, $\delta \neq 0$. Hence apply Theorem 2. □

We now turn to the question raised by Pirvu and Schulze [33]. The problem is: Does the prospect value of a portfolio p depend only on its mean and variance if and only if asset returns X are elliptically distributed?. For the sufficient part, the authors provide a positive answer by showing directly the dependence of ρ^{CPT} on the mean and variance of some transformation of asset returns. Specifically, they assume the presence of a riskfree asset. Moreover, they suppose elliptical distributed excess returns over the riskfree rate and the benchmark return in the reference wealth. In our notation, their assumption becomes $X - r\mathbf{1} - b =: y \in E_n(\hat{\Delta}, \hat{\Sigma}; \hat{\Psi})$ for some parameters $\hat{\Delta}$, $\hat{\Sigma}$ and $\hat{\Psi}$. Furthermore, they suppose y to possess a density. Here are our observations. First, looking at the proof of our Proposition 7 we note that the presence of a density is not necessary. Second, by Lemma 2, *(iii)*, we see that if a density exists then we could take X to be location-scale distributed. Third, it seems more natural to assume elliptically distributed returns for the basic assets X instead of y. In fact, from a mathematical point of view things are equivalent thanks to Lemma 2, *(i)*. Finally, we observe that the presence of a riskless asset is not necessary. Our Theorem 4 thus confirms their answer and we highlight that their hypotheses can be substantially weakened; see also the previous observations 1 and 4.

For the necessary part, Pirvu and Schulze [33] advance a conjecture but they are not able to prove it. We remark that in the aforementioned article, portfolios are generated by linear combinations of $y := X - r\mathbf{1} - b$ and not by combinations of X. Together with W^{ref}, this introduces a non negligible bias between the distribution of X and that of the attainable gains/losses Y. Now, consider Theorem 4. By its proof, we see that the mean and the variance of the gains/losses Y characterize the distribution of Y but not that of the original assets X. Effectively, if we exclude the pathological cases *(i)* and *(ii)* of Proposition 8 then the conjecture is correct. Notably, only case *(iii)* allows for the strongest conclusion of elliptical asset returns. We thus positively answer the conjecture with the following corollary.

Corollary 4. *Assume W^{ref} as in* (12) *with ($a \neq \alpha$ and $\delta \neq 0$) or ($a \neq \alpha$, $\Delta = 0$ and $\delta = 0$). ρ^{CPT} is MV in $\mathcal{V}_1$ iff X is elliptically distributed.*

We conclude this section observing that Lemma 3 can be applied to the DT too, since we know that $(\rho^Y, \mathcal{T}_1)$ ranks the distributions. However, we have to be careful since ρ^Y only evaluates non negative terminal wealths. To fix ideas, assume there is a riskfree asset so that $W^1 = \alpha'(\mathbf{1} + X) + \alpha_0(1 + r)$. If asset returns X are elliptical, then the distribution of W^1 is characterized by its mean and variance for every (α, α_0). Consequently, ρ^Y is MV in $\mathcal{T}_1$. But in order to have a sensible economic model, we should put further restrictions on X and/or (α, α_0) leading to $W^1 \geq 0$ almost surely. The other cases can be studied similarly.

5 Conclusions

In this paper, we studied in a rigorous way law invariant preference functionals and their connections with stochastic dominance and Mean-Variance analysis. First, we introduced the intuitive notion of a preference functional ranking the distributions of random variables. This condition is substantially the converse of the monotonicity of the functional with respect to distributions. Putting together the two conditions, we saw that every functional can be made independent of the initial wealth of the decision maker. After that, we extended First and Second order Stochastic Dominance to the Dual Theory of Choice. Interestingly, we found the same rankings over the distributions as those for Expected Utility with the only restriction to non negative lotteries. The most interesting case remains that of Cumulative Prospect Theory. Even considering S-shaped value functions exhibiting loss aversion and reversed S-shaped probability distortions, we were able to prove that First Stochastic Dominance is indeed equivalent to prospect value dominance, a result which is completely new to the literature.

The previous findings opened the door for extending Mean-Variance analysis to more general preference paradigms than Expected Utility. In particular, we saw how a law invariant functional that ranks the distributions naturally leads to a functional that depends only on the mean and variance of the lotteries. Mean-Variance preferences are also strictly connected to elliptical distributions of the underlying prospects. Our main result is the characterization of the distributions that imply Mean-Variance preferences, namely the elliptical ones as expected. However, the Cumulative Prospect Theory paradigm deserves a special treatment. It was already known that elliptical asset's returns lead to Mean-Variance prospect value. As a matter of fact, we proved under mild assumptions that whenever the prospect value depends only on the mean and variance, the distribution of the stocks' returns is elliptical. Furthermore, the presence of a riskless asset does not influence the result. This fact closes Mean-Variance analysis in the Cumulative Prospect Theory to elliptical distributions, as it happens for Expected Utility.

Here is a list of some topics left uncovered by this paper. To us, they deserve a special attention and we left them for future research.

First, Mean-Variance analysis has profound links with the mutual fund separation property. For the Expected Utility paradigm, Cass and Stiglitz [6] characterized the utility functions that imply separability for every distribution of the stocks' returns. Conversely, Ross [35] characterized the distributions that imply two-fund separability for every concave increasing utility functions. Chamberlain [7] proved that spherical distributions (with a riskless asset) and conditional spherical distributions (without a riskless asset) of the risky assets' return leads to separability without restrictions over the utility functions. We would like to establish a full parallel with the Cumulative Prospect Theory. In an earlier version of their working paper, De Giorgi, Hens and Levy [8] showed that two-fund separability holds if asset returns are multivariate normal and the riskfree asset is available for some special cases of the Cumulative Prospect Theory functional. Pirvu and Schulze [33] extended the result to elliptical distribution with mild assumptions over the value functions and the probability distortions. The technicalities in Cumulative Prospect Theory lies in the possible ill-posedness of the portfolio optimization problem. We ask whether it is possible to characterize the S-shaped utility functions and/or the weighting functions that imply mutual fund separability. Similarly, we would like to find the distributions that lead to separation without imposing conditions over the Cumulative Prospect Theory preference functional.

Second, Mean-Variance analysis is the basis for the Security Market Line Theorem and the Capital Asset Pricing model. Early works on the subject by Sharpe, Lintner and Mossin during the 60's assumed Mean-Variance preferences to be increasing in the mean and decreasing in the variance. For Expected Utility maximizers, this is accomplished through a quadratic utility function or via assumptions on the distributions of the assets. Once such limitations are imposed, one can identify a frontier of efficient portfolios. However, Chamberlain [7] showed that even for risk averse investors, Expected Utility can be non increasing in the mean. Similarly, we highlighted that for Cumulative Prospect Theory a similar situation is possible too. De Giorgi, Hens and Rieger [10] showed that Capital Asset Pricing equilibria need not to exist for some specifications of Cumulative Prospect Theory preferences. De Giorgi, Hens and Levy [9] gave a unified framework for a Capital Asset Pricing model where investors following different preference paradigms coexist. In particular, they showed the robustness of the Security Market Line Theorem under the assumption of normally distributed asset returns. Subsequently, Del Vigna [11] provided sufficient conditions for the existence of Capital Asset Pricing equilibria with positive prices with heterogeneous Expected Utility and Cumulative Prospect Theory maximizers. As a by-product, endogenous market segmentation can arise. Now, it is clear that the analysis over the existence of such equilibria can go through assuming elliptical returns without heavy restrictions over the preference paradigms. However, a crucial hypothesis for Capital Asset Pricing models is common knowledge about stocks' payoffs. In other words, every investor should have the same beliefs about the distributions of the original assets. Our analysis showed that for Cumulative Prospect Theory maximizers, different reference wealths possibly lead to different perceptions of such distributions. We wonder about robustness of equilibria with respect to heterogeneity in such perceptions.

Third, our findings pointed out that the Mean-Variance criterion should be commonly employed when evaluating risky prospects. As long as the distribution of the lotteries remains elliptical, we may suppose that decision makers have the same efficient set and consequently follow the same pattern of choice. This is totally independent of the preference paradigm the individuals adhere, as long as the representing functional satisfies basic assumptions such as law invariance and monotonicity in the distribution. Laboratory contexts should highlight such behavior. Additional empirical and experimental investigation is thus needed.

Finally, our very first definition involved law invariance and monotone decreasing preference functionals with respect to the distribution of the lotteries. This definition is strictly connected to the wide ambit of risk measures. As is knows, a coherent law invariant risk measure implies such monotonicity. However, the assumptions for coherent risk measures are quite restrictive, in particular sub-additivity, homogeneity and cash-invariance. Apparently they are not fulfilled by most common preference functionals. We would like to find conditions over the preferences or over the representing functionals that ensure monotonicity. More generally, we would like to investigate on the links between risk measures, generalized reward-risk measures and non-expected utilities.

6 Appendix

Proof of Proposition 3. Firstly, we give an intuition of the proof. As already observed in text, we proved the result for $\mathcal{T}_0$ using a weighting function T which did not explicitly exploited the right continuity of the distributions. This is because we only needed rough estimates for ρ^Y and we could choose $T(x)$ to be identically null in a neighborhood of $x = 0$. However, when using the class $\mathcal{T}_1$ we have to be careful, since T strictly positive near 0 can lead to unbounded ρ^Y depending on the choice of (T, F). In order to accommodate the presence of any distribution F, we must be able to find a suitable distortion T providing boundedness for ρ^Y. The argument is as follows.

Assume $F_1(\bar{x}) > F_2(\bar{x})$ for some $\bar{x} \geq 0$. Choose a sequence $\{T_n\}_n \subset \mathcal{T}_1$ such that for sufficiently big n it holds

$$T_n(x) = \begin{cases} \tilde{T}_n & x \in [0, 1 - F_1(\bar{x})], \\ \frac{n-2}{n} \frac{x-(1-F_1(\bar{x}))}{F_1(\bar{x})-F_2(\bar{x})} + \frac{1}{n} & x \in (1 - F_1(\bar{x}), 1 - F_2(\bar{x})), \\ 1 - \frac{1}{n} \left[\frac{1-x}{F_2(\bar{x})}\right]^{\frac{(n-2)F_2(\bar{x})}{F_1(\bar{x})-F_2(\bar{x})}} & x \in [1 - F_2(\bar{x}), 1], \end{cases}$$

where $\tilde{T}_n$ is such that $T_n \in \mathcal{T}_1$, i.e. $\tilde{T}_n(0) = 0$, $\tilde{T}_n(1 - F_1(\bar{x})) = \frac{1}{n}$, $\tilde{T}_n' > 0$, and

$$\lim_{n\to\infty} \int_{\bar{x}}^{+\infty} \tilde{T}_n(1 - F_1(x))dx = 0.$$

Some clarifications are in order. First, we select a S-shaped T_n, being a straight line in the middle interval and non linear elsewhere. As n grows to infinity, T_n becomes steeper in the middle interval and flatter elsewhere. Observe that if $F_1(\bar{x}) = 1$ then we only need the last two pieces of T_n. Second, we explicitly defined these latter pieces in order to give precise estimates, but any other analogous specification would be correct. Third, T_n converges uniformly to the distortion that we used for $\mathcal{T}_0$ but we can apply a limit argument as in Proposition 1. Finally, and more important, we can always find such a sequence $\{\tilde{T}_n\}_n$ and it can be chosen to be monotone non increasing. For example, if F_1 is continuous we can set $\tilde{T}_n(x) \simeq \frac{1}{[F_1^{-1}(1-x)]^n}$ in a neighborhood of $x = 0$. If F_1 is not continuous, then we can replace F_1^{-1} with its convex envelope. If F_1 is $O(x^{-m})$, $m \in \mathbb{N}$, for x tending to infinity, then we can even explicitly set

$$\tilde{T}_n(x) = \frac{1}{n}\left[\frac{x}{1-F_1(\bar{x})}\right]^{\frac{(n-2)(1-F_1(\bar{x}))}{F_1(\bar{x})-F_2(\bar{x})}}.$$

Similarly to the proof for the class $\mathcal{T}_0$, we now proceed to estimate

$$\rho^Y(T_n, F_1) \le \bar{x} + \int_{\bar{x}}^{+\infty} \tilde{T}_n(1-F_1(x))dx,$$

and

$$\rho^Y(T_n, F_2) \ge (1-\tfrac{1}{n})\bar{x} + \int_{\bar{x}}^{F_2^{-1}(F_1(\bar{x}))} \left[\frac{n-2}{n}\frac{F_1(\bar{x})-F_2(x)}{F_1(\bar{x})-F_2(\bar{x})} + \frac{1}{n}\right] dx.$$

For the reader's convenience, we recall the existence of $\beta > 0$ such that F_1 and F_2 are continuous over $[\bar{x}, \bar{x}+\beta]$, $F_1(\bar{x}) > F_2(x)$ for every $x \in [\bar{x}, \bar{x}+\beta]$ and $F_2^{-1}(F_1(\bar{x})) \ge \bar{x}+\beta$. Setting η as in (5) and noticing that η does not depend on n, we find

$$\rho^Y(T_n, F_1) - \rho^Y(T_n, F_2) \le \frac{1}{n}\bar{x} + \int_{\bar{x}}^{+\infty} \tilde{T}_n(1-F_1(x))dx - \left(\frac{n-2}{n}\right)\eta\beta.$$

Passing to limit for $n \to \infty$, we conclude since $\eta > 0$. □

Proof of Proposition 6. ($\Leftarrow$) Follows from Proposition 5. ($\Rightarrow$) The proof is once again by contradiction, similar to that of Proposition 3. We shall exhibit a sequence $\{v_n\}_n \subset \mathcal{V}_1$, $v_n = (u_n^+, u_n^-, T_n^+, T_n^-)$ providing good estimates for the gain and the loss part of ρ^{CPT} respectively. Then we pass to the limit, showing the convergence to a strictly negative value and we conclude.

Now, assume there is $\bar{x} \ge 0$ with $F_1(\bar{x}) > F_2(\bar{x})$. Thus we can find $\beta > 0$ such that F_1, F_2 are continuous and satisfy $F_1(\bar{x}) > F_2(x)$ for $x \in [\bar{x}, \bar{x}+\beta]$. Choose the sequence $\{v_n\}_n$ as follows. The value functions are given by

$$u_n^+(x) = \begin{cases} Ax & x \in [0, \bar{x}+\beta], \\ A(\bar{x}+\beta) + \frac{1-\exp\{-2nA[x-(\bar{x}+\beta)]\}}{2n} & x \in (\bar{x}+\beta, +\infty), \end{cases} \tag{14}$$

$$u_n^-(x) = \begin{cases} Bx & x \in [-(\bar{x}+\beta), 0], \\ -B(\bar{x}+\beta) - \frac{1-\exp\{nB[x+(\bar{x}+\beta)]\}}{n} & x \in (-\infty, -(\bar{x}+\beta)), \end{cases} \tag{15}$$

where $0 < A < B < 2A$. Observe that this condition is sufficient to guarantee loss aversion as specified in (10). Moreover, u_n^+, u_n^- are bounded, so that we can apply integration by parts. We now compute estimates independently on the particular choice of T_n^+, T_n^-. For the losses part, we have

$$\begin{aligned}
\rho_-^{CPT}(v_n, F_1) &= \int_{-\infty}^{-(\bar{x}+\beta)} u_n^-(x)d[T_n^-(F_1(x))] + \int_{-(\bar{x}+\beta)}^{0} u_n^-(x)d[T_n^-(F_1(x))] \\
&\le -B(\bar{x}+\beta)\int_{-\infty}^{-(\bar{x}+\beta)} d[T_n^-(F_1(x))] + B\int_{-(\bar{x}+\beta)}^{0} xd[T_n^-(F_1(x))] \\
&= -B\int_{-(\bar{x}+\beta)}^{0} T_n^-(F_1(x))dx.
\end{aligned}$$

Similar computations give

$$\begin{aligned}
\rho_-^{CPT}(v_n, F_2) &\ge -[B(\bar{x}+\beta) + \tfrac{1}{n}]\int_{-\infty}^{-(\bar{x}+\beta)} d[T_n^-(F_2(x))] + B\int_{-(\bar{x}+\beta)}^{0} xd[T_n^-(F_2(x))] \\
&= -\tfrac{1}{n}T_n^-(F_2(-(\bar{x}+\beta))) - B\int_{-(\bar{x}+\beta)}^{0} T_n^-(F_2(x))dx,
\end{aligned}$$

leading to

$$\begin{aligned}
I_n^- &:= \rho_-^{CPT}(v_n, F_1) - \rho_-^{CPT}(v_n, F_2) \\
&\qquad\qquad\qquad \leq B\int_{-(\bar{x}+\beta)}^{0} [T_n^-(F_2(x)) - T_n^-(F_1(x))]dx + \tfrac{1}{n}T_n^-(F_2(-(\bar{x}+\beta))). \qquad (16)
\end{aligned}$$

Now observe that $F_2(0) < 1$, otherwise we can not have $\bar{x} \geq 0$. If $F_2(0) = 0$, then we are done since equation (16) reduces to

$$I_n^- \leq -B\int_{-(\bar{x}+\beta)}^{0} T_n^-(F_1(x))dx \leq 0, \quad \text{for every } T_n^- \in \mathcal{T}_{RS}.$$

Else $F_2(0) > 0$ and we choose

$$T_n^-(x) = \begin{cases} \frac{1}{n}\left[\frac{x}{F_2(0)}\right]^{\gamma} & x \in [0, F_2(0)], \\ 1 - (1-\frac{1}{n})\left[\frac{1-x}{1-F_2(0)}\right]^{\frac{\gamma(1-F_2(0))}{(n-1)F_2(0)}} & x \in (F_2(0), 1], \end{cases} \tag{17}$$

with $\gamma \in (0,1)$. Recalling that F_2 is non decreasing, we find

$$\begin{aligned}
I_n^- &\leq B\int_{-(\bar{x}+\beta)}^{0} T_n^-(F_2(x))dx + \tfrac{1}{n}T_n^-(F_2(-(\bar{x}+\beta))) \\
&\leq B\tfrac{1}{n}(\bar{x}+\beta) + \tfrac{1}{n}T_n^-(F_2(-(\bar{x}+\beta))),
\end{aligned}$$

which converges to 0 as n grows to infinity. Note that passing to the limit is harmless since we are integrating a bounded function over a compact set (the same holds for I_n^+ below). In each case we see that the loss part of the prospect value can be made as small as we desire. Now, for the gains part we proceed as before obtaining

$$\begin{aligned}
I_n^+ &:= \rho_+^{CPT}(v_n, F_1) - \rho_+^{CPT}(v_n, F_2) \\
&\leq A\int_{0}^{\bar{x}+\beta} [T_n^+(1-F_1(x)) - T_n^+(1-F_2(x))]dx + \tfrac{1}{n}T_n^+(1-F_1(\bar{x}+\beta)).
\end{aligned}$$

We distinguish four cases depending on the value of $F_2(\bar{x}+\beta)$ and $F_1(\bar{x})$.

(i) $F_2(\bar{x}+\beta) = 0$ and $F_1(\bar{x}) = 1$. Then $F_2(x) = 0$ for $x \in [0, \bar{x}+\beta]$ and $F_1(x) = 1$ for $x \in [\bar{x}, \bar{x}+\beta]$. This leads to

$$I_n^+ \leq -A\int_{\bar{x}}^{\bar{x}+\beta} dx = -A\beta < 0, \quad \text{for every } T_n^+ \in \mathcal{T}_{RS}.$$

(ii) $F_2(\bar{x}+\beta) > 0$ and $F_1(\bar{x}) = 1$. For $\gamma \in (0,1)$, set

$$T_n^+(x) = \begin{cases} (1-\frac{1}{n})\left[\frac{x}{1-F_2(\bar{x}+\beta)}\right]^{\frac{\gamma(1-F_2(\bar{x}+\beta))}{(n-1)F_2(\bar{x}+\beta)}} & x \in [0, 1-F_2(\bar{x}+\beta)], \\ 1 - \frac{1}{n}\left[\frac{1-x}{F_2(\bar{x}+\beta)}\right]^{\gamma} & x \in (1-F_2(\bar{x}+\beta), 1]. \end{cases}$$

Then the integrand appearing in I_n^+ is dominated by $\frac{1}{n}$ for $x \in [0, \bar{x}]$ and it is lower than $(\frac{1}{n} - 1)$ for $x \in [\bar{x}, \bar{x}+\beta]$. Hence we find

$$\lim_n I_n^+ \leq \lim_n A[\tfrac{1}{n}\bar{x} + (\tfrac{1}{n} - 1)\beta] = -A\beta < 0.$$

(iii) $F_2(\bar{x}+\beta) = 0$ and $F_1(\bar{x}) < 1$. For $\gamma \in (0,1)$, choose

$$T_n^+(x) = \begin{cases} \frac{1}{n}\left[\frac{x}{1-F_1(\bar{x})}\right]^{\gamma} & x \in [0, 1-F_1(\bar{x})], \\ 1 - (1-\frac{1}{n})\left[\frac{1-x}{F_1(\bar{x})}\right]^{\frac{\gamma F_1(\bar{x})}{(n-1)(1-F_1(\bar{x}))}} & x \in (1-F_1(\bar{x}), 1]. \end{cases}$$

Now we estimate

$$\begin{aligned}I_n^+ &\le A\int_0^{\bar{x}}[T_n^+(1-F_1(x))-1]dx + A\int_{\bar{x}}^{\bar{x}+\beta}(\tfrac{1}{n}-1)dx + \tfrac{1}{n}T_n^+(1-F_1(\bar{x}+\beta))\\ &\le A(\tfrac{1}{n}-1)\beta + \tfrac{1}{n}T_n^+(1-F_1(\bar{x}+\beta)),\end{aligned}$$

which implies $\lim_n I_n^+ \le -A\beta < 0$.

(iv) $F_2(\bar{x}+\beta) > 0$ and $F_1(\bar{x}) < 1$. For $\gamma \in (\frac{1}{4}, 1)$, set

$$T_n^+(x) = \begin{cases}(1-\frac{1}{n})\frac{x}{1-F_2(\bar{x}+\beta)} & x\in[0,1-F_2(\bar{x}+\beta)],\\ 1-\frac{1}{2n}-\frac{1}{2n}\left[\frac{1-x-F_2(\bar{x}+\beta)/2}{F_2(\bar{x}+\beta)/2}\right]^{\frac{(n-1)F_2(\bar{x}+\beta)}{1-F_2(\bar{x}+\beta)}} & x\in(1-F_2(\bar{x}+\beta), 1-\frac{F_2(\bar{x}+\beta)}{2}],\\ 1-\frac{1}{2n}+\frac{1}{6n}\left[\frac{x-1-F_2(\bar{x}+\beta)/2}{F_2(\bar{x}+\beta)/3}\right]^{4\gamma} & x\in(1-\frac{F_2(\bar{x}+\beta)}{2}, 1-\frac{F_2(\bar{x}+\beta)}{6}],\\ 1-\frac{1}{3n}\left[\frac{1-x}{F_2(\bar{x}+\beta)/6}\right]^{\gamma} & x\in(1-\frac{F_2(\bar{x}+\beta)}{6},1].\end{cases} \tag{18}$$

See Figure 3. Note that T_n^+ are linear over $[0, 1-F_2(\bar{x}+\beta)]$ and they do not depend on $F_1(\bar{x})$. Moreover, γ must be greater than $\frac{1}{4}$ to ensure convexity in a neighborhood of $x=1$. Again, the integrand in I_n^+ is lower than $\frac{1}{n}$ for $x\in[0,\bar{x}]$. Recalling that F_1, F_2 are non decreasing, we estimate

$$\begin{aligned}I_n^+ &\le A\tfrac{1}{n}\bar{x} + A\beta[T_n^+(1-F_1(\bar{x})) - T_n^+(1-F_2(\bar{x}+\beta))] + \tfrac{1}{n}T_n^+(1-F_1(\bar{x}+\beta))\\ &\le A\tfrac{1}{n}\bar{x} + A\beta[(1-\tfrac{1}{n})\tfrac{1-F_1(\bar{x})}{1-F_2(\bar{x}+\beta)} - (1-\tfrac{1}{n})] + \tfrac{1}{n},\end{aligned}$$

which implies $\lim_n I_n^+ \le -A\beta\frac{F_1(\bar{x})-F_2(\bar{x}+\beta)}{1-F_2(\bar{x}+\beta)} < 0$.

Hence, in each case we see that $\rho_+^{CPT}(v_n, F_1) - \rho_+^{CPT}(v_n, F_2)$ tends to a strictly negative quantity. We observe that the unique connection between the estimates for the gain part and the loss part is given by the value of $F_2(0)$. However, this value only affects the analysis for the loss part in (17). Combining this with the previous estimates concludes. Finally, if $F_1(-\bar{x}) > F_2(-\bar{x})$ for some $\bar{x} > 0$ then we find $\beta > 0$ such that $-\bar{x}+\beta < 0$, F_1, F_2 are continuous and $F_1(-\bar{x}) > F_2(x)$ for $x\in[-\bar{x}, -\bar{x}+\beta]$. We adjust u_n^+ (u_n^-) in equation (14) ((15)) choosing a linear shape for $x\in[0,\bar{x}]$ ($[-\bar{x},0]$) and a concave (convex) shape elsewhere. The subsequent analysis proceeds as before, adjusting minor details and choosing appropriate T_n^+, T_n^-. □

We now add some remarks to the previous proof. First, we provided explicit expressions for the value functions and the probability distortions. It is evident that qualitatively similar functions do the job as well. For u_n^+, u_n^-, it is sufficient to assume $\lim_{x\to+\infty} u_n^+(x) = L_n^+$ with $\lim_n L_n^+ = A(\bar{x}+\beta)$ and $\lim_{x\to-\infty} u_n^-(x) = -L_n^-$ with $\lim_n L_n^- = -B(\bar{x}+\beta)$. Analogous considerations hold for T_n^+, T_n^-, where we also provided families of functions depending on a real parameter γ. Second, observe that the weighting functions are in some sense necessary. If they are not present then we just have a S-shaped utility function and we are led to SD relations as in [30]. Similarly, it is necessary to have strictly convex-concave value functions at least over some regions. In fact, if we choose linear u_n^+, u_n^-, then integrations by parts lead us to a model analogous to DT. In turn, we are obliged to select S-shaped probability distortions as in the proof of Proposition 3, abandoning the class $\mathcal{T}_{RS}$.

Example 2. The example is twofold. First, we provide a case where $\rho^{CPT}(v,X) = h(v,\mathbb{E}[X],\mathbb{V}[X])$ is not increasing in the mean. We build on Chamberlain [7] assuming m independent of Z, with $\mathbb{P}(m=3) = \mathbb{P}(m=-1) = \frac{1}{2}$; $\mathbb{P}(Z=1) = \mathbb{P}(Z=-1) = \theta$, $\mathbb{P}(Z=0) = 1-2\theta$ for $\theta\in(0,\frac{1}{2})$. Let $X_1 = m$ and $X_2 = -m$, so that $\mathbb{E}[X_1] = 1$, $\mathbb{E}[X_2] = -1$ and $\mathbb{V}[X_1] = \mathbb{V}[X_2] = 4$. Recalling the notation $v = (u^+, u^-, T^+, T^-)$, we require $u^+(x) = x$ for $x\le 3$ and $u^-(x) = \frac{3}{2}x$ for $x\ge -3$. Next, we require

$$T_n^+(x) = \tfrac{1}{n}(2x)^{1/n}, \quad T_n^-(x) = \tfrac{1}{2}(2x)^{1/n}, \quad \text{for } x\in[0,\tfrac{1}{2}), \quad n\in\mathbb{N}.$$

Define $X^* := \sqrt{\frac{2}{\theta}}Z$. Then $\mathbb{E}[X^*] = 0$ and $\mathbb{V}[X^*] = 4$ for every θ. Straightforward computations show that $\rho^{CPT}(v_n, X_1) = \frac{3}{n} - \frac{3}{4}$, $\rho^{CPT}(v_n, X_2) = \frac{1}{n} - \frac{9}{4}$, and $\rho^{CPT}(v_n, X^*) = \sqrt{\frac{2}{\theta}}T_n^+(\theta) - \frac{3}{2}\sqrt{\frac{2}{\theta}}T_n^-(\theta)$. If θ

is sufficiently small and n is chosen appropriately, then we find $h(v_n, 0, 4) < h(v_n, -1, 4) < h(v_n, 1, 4)$ as claimed. For a direct check, set $\theta = \frac{1}{10}$ and $n = 2$.

Second, we show a case where ρ^{CPT} is not decreasing in the variance. We build on De Giorgi, Hens and Levy [8] and Del Vigna [11]. For, select u^+ to be strictly increasing and convex and set $u^-(x) := -\lambda u^+(-x)$ for $x \leq 0$, with $\lambda > 1$ to induce loss aversion. Next we choose a special CPT functional, the so-called rank dependent one. Specifically, we assume $T^+(1-x) = 1 - T^-(x)$ for $x \in [0,1]$, with T^- continuous, piecewise differentiable and strictly increasing. Now, for any symmetrically distributed random variable Y we have

$$\begin{aligned}\rho^{CPT}(v, Y) &= \int_0^{+\infty} u^+(x) d[T^-(F_Y(x))] - \lambda \int_{-\infty}^0 u^+(-x) d[T^-(F_Y(x))] \\ &= \int_0^{+\infty} u^+(x)[(T^-)'(F_Y(x)) - \lambda (T^-)'(1 - F_Y(x))] dF_Y(x).\end{aligned}$$

In particular, if $Y \equiv 0$ then $\rho^{CPT} = 0$. On the other hand, let Y be a standard normal random variable. In order to conclude the example, we only need to find $\lambda > 1$ and T^- such that for every $x \in (\frac{1}{2}, 1]$ it holds $(T^-)'(x) - \lambda (T^-)'(1-x) > 0$. Indeed, u^+ is positive as well as dF_Y. We choose the probability distortion to be

$$T^-(x) = \begin{cases} 1.1x & x \in [0, 0.1), \\ 0.1x + 0.1 & x \in [0.1, 0.5), \\ 1.7x - 0.7 & x \in [0.5, 1]. \end{cases}$$

An explicit computation shows that $\lambda < \frac{3}{2}$ does the job.

References

[1] N. Barberis and M. Huang (2008): Stocks as lotteries: The implications of probability weighting for security prices. *The American Economic Review*, 98(5), 2066-2100.

[2] M. Baucells and F.H. Heukamp (2006): Stochastic Dominance and Cumulative Prospect Theory. *Management Science*, 52(9), 1409-1423.

[3] C. Bernard and M. Ghossoub (2010): Static portfolio choice under cumulative prospect theory. *Mathematics and financial economics*, 2(4), 277-306.

[4] L. Campi and M. Del Vigna (2012): Weak Insider Trading and Behavioral Finance. *SIAM Journal on Financial Mathematics*, 3(1), 242-279.

[5] G. Carlier and R.A. Dana (2006): Law invariant concave utility functions and optimization problems with monotonicity and comonotonicity constraints. *Statistics & Decisions*, 24, 127-152.

[6] D. Cass and J. Stiglitz (1970): The Structure of Investor Preferences and Asset Returns, and Separability in Portfolio Allocation. *Journal of Economic Theory*, 2, 122-160.

[7] G. Chamberlain (1983): A Characterization of the Distributions That Imply Mean-Variance Utility Functions. *Journal of Economic Theory*, 29, 185-201.

[8] E.G. De Giorgi, T. Hens and H. Levy (2011): CAPM equilibria with prospect theory preferences. *Working Paper*, available at http://ssrn.com/abstract=420184.

[9] E.G. De Giorgi, T. Hens and H. Levy (2012): Two Paradigms and Nobel Prizes in Economics: a Contradiction or Coexistence? *European Financial Management*, 18(2), 163-182.

[10] E.G. De Giorgi, T. Hens and M.O. Rieger (2010): Financial market equilibria with cumulative prospect theory. *Journal of Mathematical Economics*, 46, 633-651.

[11] M. Del Vigna (2011): Financial market equilibria with heterogeneous agents: CAPM and market segmentation. *Working Paper*, available at http://www.dmd.unifi.it/RePEc/flo/WorkingPapers/2011/DiMaDWP2011-08.pdf.

[12] E. Diecidue and P.P. Wakker (2001): On the intuition of Rank-Dependent Utility. *The Journal of Risk and Uncertainty*, 23(3), 281-298.

[13] K. Fan and G.G. Lorentz (1954): An Integral Inequality. *The American Mathematical Monthly*, 61(9), 626-631.

[14] I. Fortin and J. Hlouskova (2011): Optimal asset allocation under linear loss aversion. *Journal of Banking & Finance*, 35, 2974-2990.

[15] A. Galichon and M. Henry (2012): Dual theory of choice with multivariate risks. *Journal of Economic Theory*, 147(4), 1501-1516.

[16] J. Hadar and W.R. Russell (1969): Rules for Ordering Uncertain Prospects. *The American Economic Review*, 59(1), 25-34.

[17] G. Hanoch and H. Levy (1969): The Efficiency Analysis of Choices Involving Risk. *The Review of Economic Studies*, 36(3), 335-346.

[18] X.D. He and X.Y. Zhou (2011): Portfolio Choice Via Quantiles. *Mathematical Finance*, 21(2), 203-231.

[19] J.E. Ingersoll (1987): Theory of Financial Decision Making. Rowman and Littlefield.

[20] H. Jin and X.Y. Zhou (2008): Behavioral portfolio selection in continuous time. *Mathematical Finance*, 18, 385-426.

[21] D. Kahneman and A. Tversky (1979): Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263-292.

[22] D. Kelker (1970): Distribution Theory of Spherical Distributions and a Location-Scale Parameter Generalization. *Sankhyā : The Indian Journal of Statistics, Series A*, 32(4), 419-430.

[23] V. Köbberling and P.P. Wakker (2005): An index of loss aversion. *Journal of Economic Theory*, 122, 119-131.

[24] B. Köszegi and M. Rabin (2006): A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 71(4), 1133-1165.

[25] S. Kusuoka (2001): On Law Invariant Coherent Risk Measures. *Advances in Mathematical Economics*, 3, 83-95.

[26] E.L. Lehmann (1955): Ordered Families of Distributions. *The Annals of Mathematical Statistics*, 26(3), 399-419.

[27] M. Leshno and H. Levy (2002): Preferred by "All" and Preferred by "Most" Decision Makers: Almost Stochastic Dominance. *Management Science*, 48(8), 1074-1085.

[28] H. Levy (1992): Stochastic Dominance and Expected Utility: Survey and Analysis. *Management Science*, 38(4), 555-593.

[29] H. Levy and R. Duchin (2004): Asset Return Distributions and the Investment Horizon. *Journal of Portfolio Management*, 30(3), 47-62.

[30] H. Levy and M. Levy (2002): Prospect Theory: Much Ado about Nothing? *Management Science*, 48(10), 1334-1349.

[31] H. Levy and Z. Wiener (1998): Stochastic Dominance and Prospect Dominance with Subjective Weighting Functions. *Journal of Risk and Uncertainty*, 16(2), 147-163.

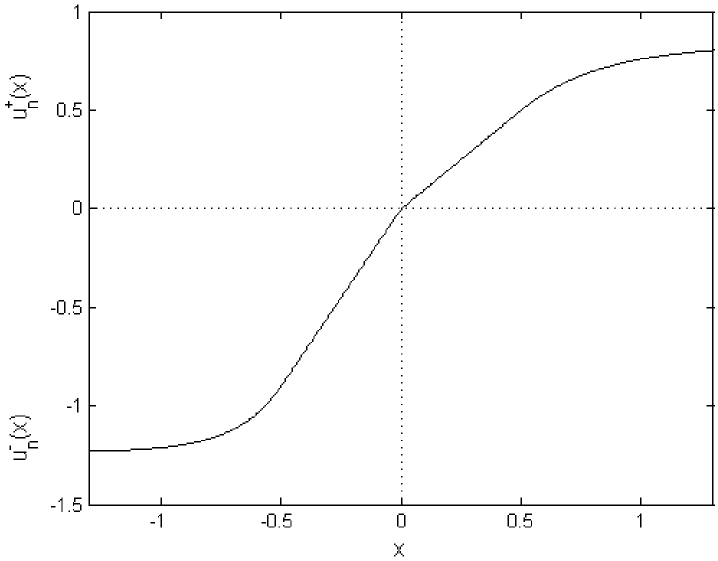

Figure 1: the value functions u_n^+, u_n^- in equations (14) and (15) for $n = 3$, $A = 1$, $B = 1.8$, $\bar{x}+\beta = 0.5$. This is a S-shaped utility function as suggested by CPT.

[32] J. Owen and R. Rabinovitch (1983): On the Class of Elliptical Distributions and their Applications to the Theory of Portfolio Choice. *The Journal of Finance*, 38(3), 745-752.

[33] T.A. Pirvu and K. Schulze (2012): Multi-stock portfolio optimization under prospect theory. To appear in *Mathematics and Financial Economics*.

[34] J.P. Quirk and R. Saposnik (1962): Admissibility and Measurable Utility Functions. *The Review of Economic Studies*, 29(2), 140-146.

[35] S.A. Ross (1978): Mutual fund separation in financial theory – the separating distributions. *Journal of Economic Theory*, 17, 254-286.

[36] U. Schmidt and H. Zank (2005): What is loss aversion? *The Journal of Risk and Uncertainty*, 30:2, 157-167.

[37] A. Tversky and D. Kahneman (1992): Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297-323.

[38] W.K. Wong and R.H. Chan (2008): Prospect and Markowitz stochastic dominance. *Annals of Finance*, 4, 105-129.

[39] M. E. Yaari (1987): The Dual Theory of Choice under Risk. *Econometrica*, 55(1), 95-115.

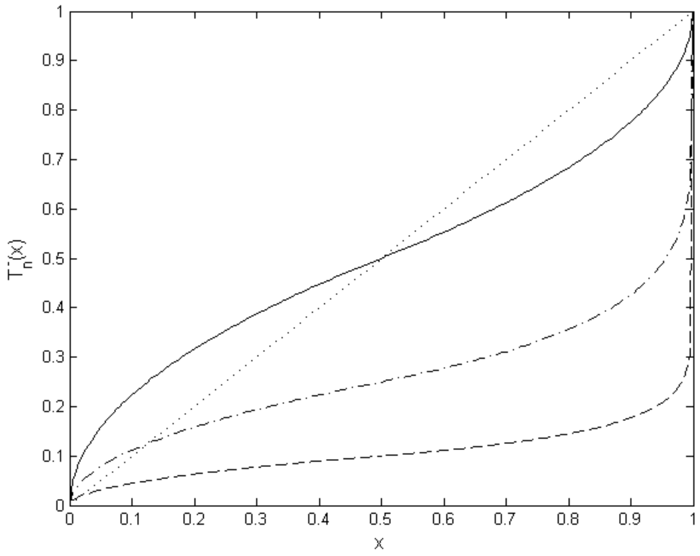

Figure 2: the reversed S-shaped weighting functions T_n^- in equation (17) for $\gamma = 0.5$, $F_2(0) = 0.5$, $n = 2$ (solid), $n = 4$ (dash-dot) and $n = 10$ (dashed).

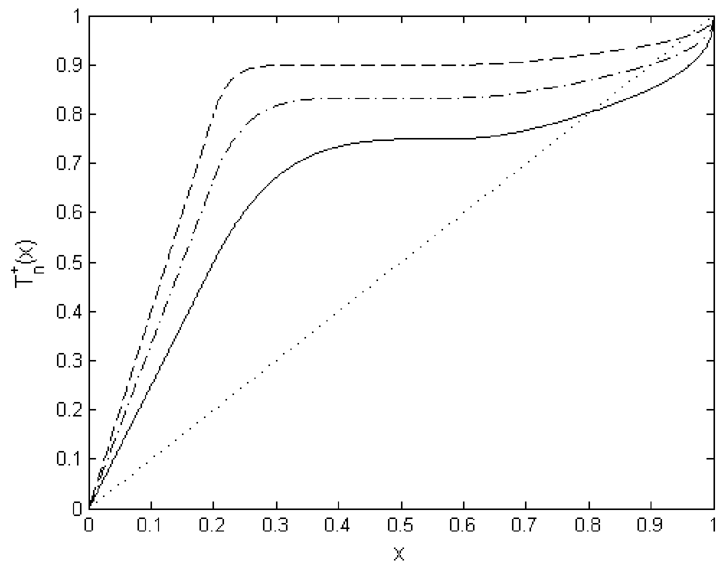

Figure 3: the reversed S-shaped weighting functions T_n^+ in equation (18) for $\gamma = 0.4$, $F_2(\bar{x} + \beta) = 0.8$, $n = 2$ (solid), $n = 3$ (dash-dot) and $n = 5$ (dashed).