Modeling Portfolio Efficiency using Stochastic Optimization with Incomplete Information and Partial Uncertainty

D. La Torre[1], F. Mendivil[2], M. Rocca[3]

November 22, 2021

[1] Corresponding author. SKEMA Business School, Université Côte d'Azur, Sophia Antipolis Campus, France
[2] Department of Mathematics and Statistics, Acadia University, Wolfville, Canada
[3] Department of Economics, Universitá degli Studi dell'Insubria, Varese, Italy

Abstract

Efficiency plays a crucial role in portfolio optimization. This notion is formulated by means of stochastic optimization techniques. Very often this problem is subject to partial uncertainty or incomplete information on the probability distribution and on the preferences expressed by means of the utility function. In this case both the objective function and the underlying probability measure are not known with precision. To address this kind of issues, we propose to model the notion of incomplete information by means of set-valued analysis and, therefore, we propose two different extensions of the classical model. In the first one we rely on the notion of set-valued function while the second one utilizes the notion of set-valued probability. For both of them we investigate stability properties. These results are also linked to the notion of robustness of the aforementioned problem. Finally we apply the obtained results to portfolio theory and stochastic dominance.

Keywords: Portfolio Optimization, Stochastic Dominance, Portfolio Efficiency, Partial Uncertainty, Incomplete Information, Set-Valued Analysis.

1 Introduction

Portfolio optimization and selection is a classical problem in Operations Research and Finance and it has been applied to different areas including financial markets, technological change, strategic investments, and many others.

Portfolio optimization provides to the investor a quantitative approach which will allow to select the best available investment plan. The first mathematical formulation and model in portfolio optimization goes back to the pioneering paper by Markowitz [30].

The efficient portfolio theory based on Markowitz's model is widely used in the construction of easy-to-read performance indicators. However, this model has been criticized for the strong assumptions imposed on the investor's utility function or the expected return distribution.

The notion of stochastic dominance, instead, considers the structure and behavior of the whole investment return distribution, and not only the mean and the variance (i.e. the first two moments) like in traditional Markowitz-based financial indexes. Recently, it has played a crucial role in financial performance analysis and risk evaluation.

Several papers in the literature have compared the mean-variance approach vs the stochastic dominance one with the aim of determining if there are any differences between the efficient sets of investments provided by both approaches [32, 33, 29, 26, 27].

In this paper we investigate the notion of portfolio efficiency with partial uncertainty and imprecise information. To cope with the presence of partial knowledge on the probability distribution and on the utility function we rely on techniques of set-valued analysis. Roughly speaking, to model the lack of information on the utility function we replace it with a set-valued mapping. The same philosophy is applied to the partial uncertainty on the underlying probability: the probability of a given event - which is in the classical meaning a positive number in $[0,1]$ - is replaced by a positive set-valued map with compact and convex values in $\mathbb{R}^d$.

Our approach is also related to existing techniques to deal with vague, imprecise, inconsistent and uncertain knowledge such as fuzzy set theory, evidence theory, and rough set theory (see, for instance, [48] and the references therein).

More in details, we propose two extensions. In the first one, we replace the single-valued utility function with a set-valued utility function. We suppose that the utility is not longer a real number but a subset of vectors which models the roughness and the lack of certainty on the objective function. The corresponding stochastic optimization model is a set-valued problem by construction. In the second extension, instead, we use the notion of set-valued probability to model the level of imprecision related to the underlying probability measure. The notion of set-valued probability is an extension of the notion of imprecise probability which has been widely investigated in the literature as it represents a quite natural way to extend the traditional notion of probability. While in the traditional approach associated with each event there is a number in $[0,1]$, in the theory of imprecise improbability this is replaced by a subset $[p_l, p_u]$, where p_l and p_u are the so-called lower and upper probabilities. The notion of imprecise probability has been used in several contexts and different applications have been proposed as well. Some classical examples in this theory are the Dempster-Shafer evidence theory ([10]), the coherent lower prevision theory ([45]), probability bound analysis ([13]), F-probability ([47]), the possibility theory ([12]). The definition of set-valued measures was first introduced for the needs of mathematical economics in [44] where it was used to study equilibria in exchange economies in which coalitions correspond to measurable sets and are the primary economic units (see also [9]). Moreover, the study of set-valued measures has been developed extensively because of its applications in other fields such as optimization and optimal control.

The notion of portfolio efficiency presented in this paper relies on set ordering. We study stability properties of the solution under perturbations on the utility function and on the underlying probability. We also provide generalized optimality conditions. The results in this paper also extend those proved in [6, 24, 23].

The paper proceeds as follows. Section 2 recalls the basic notion of portfolio optimization and stochastic dominance.

Section 3 is devoted to basic concepts about set-valued functions and set-valued probabilities. Section 4 is devoted to the extension of the notion of portfolio efficiency with incomplete information and partial uncertainty and, in particular, to the extension of Proposition 1 which states the optimality conditions portfolio efficiency (see [34]).

In Section 5 for a general stochastic optimization problem we present optimality conditions and stability results. In particular, in Subsection 5.1 we investigate optimality conditions and stability properties under perturbation of the utility (w.r.t. the uniform convergence) and of the underlying probability (w.r.t. the Monge-Kantorovich metric). In Subsection 5.2, we generalize the previous results to the case of set-valued utility. In Subsection 5.3 we consider the case in which the expectation is taken with respect to a set-valued probability and extend the previous results to this case as well. We apply these general results to the portfolio efficiency problem. In particular as a corollary of the optimality conditions for the general stochastic optimization problem, we prove the aforementioned extension of Proposition 1.

Section 6 presents some managerial insights and discusses some practical implementation of the method. Section 7 concludes the paper.

2 Basics on Portfolio Efficiency

As discussed in the previous section, the notion of stochastic dominance plays a crucial role to identify the optimal portfolio allocation. In the sequel of this section, we recall some classical definitions and results in stochastic dominance theory.

If X_j is a random variable which represents the possible assets, a portfolio is a linear combination of them, namely $Y_\lambda = \sum_i \lambda_i X_i$, where λ is the vector that describes the asset allocation and it belongs to the set Λ defined as:

$$\Lambda = \{\lambda \in \mathbb{R}_+^N : \sum_{i=1}^{N} \lambda_i = 1\}.$$

The following three definitions are classical [42, 43] and provide an ordering between random variables.

Definition 1. *Given two random variables X_1 and X_2 we say that X_1 dominates X_2 (in the first order) if*

$$P[X_1 \geq x] \geq P[X_2 \geq x] \tag{1}$$

for all x, and for some x, $P[X_1 \geq x] > P[X_2 \geq x]$.

Definition 2. *Given two random variables X_1 and X_2 we say that X_1 dominates X_2 (in the second order) if*

$$\mathbb{E}[u(X_1)] \geq \mathbb{E}[u(X_2)] \tag{2}$$

for all nondecreasing and concave utility functions u.

Definition 3. *Given two random variables X_1 and X_2, we say that X_1 dominates X_2 (in the third order) if and only if $\mathbb{E}[u(X_1)] \geq \mathbb{E}[u(X_2)]$ for all non-decreasing, concave utility functions u that are "'positively skewed"' (i.e. with positive third derivative).*

These are important concepts in portfolio theory and they have been explored in several papers including [34, 35, 36, 28, 15] in which the authors also provide statistical tests for stochastic dominance.

In the sequel of this paper we focus on the notion of second-order stochastic dominance. The following definition recalls stochastic portfolio efficiency.

Definition 4. *Given $\hat{\lambda} \in \Lambda$, we say that a portfolio $Y_{\hat{\lambda}}$ is stochastically efficient with respect to the probability ϕ and the utility $u \in \mathcal{U}$ if*

$$\hat{\lambda} = \text{argmax}_{\lambda \in \Lambda} \mathbb{E}_\phi(u(Y_\lambda)) = \text{argmax}_{\lambda \in \Lambda} \int_\Omega u(Y_\lambda(\omega)) d\phi(\omega) \tag{3}$$

$$= \text{argmax}_{\lambda \in \Lambda} \int_\Omega u\left(\sum_i \lambda_i X_i(\omega)\right) d\phi(\omega) \tag{4}$$

$$\tag{5}$$

where

$$\mathcal{U} = \{u : \mathbb{R} \to \mathbb{R}, u \textit{ is non-decreasing and concave}\} \tag{6}$$

More precisely, $Y_{\hat{\lambda}}$ is stochastically efficient with respect to the probability ϕ and the utility $u \in \mathcal{U}$ if

$$\mathbb{E}_\phi(u(Y_{\hat{\lambda}})) \geq \mathbb{E}_\phi(u(Y_\lambda)) \tag{7}$$

for any $\lambda \in \Lambda$.

In other words, once a probability ϕ and the utility $u \in \mathcal{U}$ are fixed, a portfolio $Y_{\hat{\lambda}}$ is stochastically efficient if $\hat{\lambda}$ is solution to the optimization problem:

$$\min_{\lambda \in \Lambda} \mathbb{E}_\phi(u(Y_\lambda)). \tag{8}$$

The following result was proved in the case of a discrete probability space in [34]. The extension to the case of a generic probability space is quite straightforward.

Proposition 1. *A given portfolio $Y_{\hat{\lambda}}$ is stochastically efficient with respect to the probability ϕ and the utility $u \in \mathcal{U}$ if and only if it obeys the following first-order optimality conditions:*

$$\mathbb{E}_\phi \left(u'(Y_{\hat{\lambda}}(\cdot)) \left[X_i(\cdot) - Y_{\hat{\lambda}}(\cdot) \right] \right) = \int_\Omega u'(Y_{\hat{\lambda}}(\omega)) \left[X_i(\omega) - Y_{\hat{\lambda}}(\omega) \right] d\phi(\omega) \leq 0 \tag{9}$$

for all $i = 1...N$.

3 Mathematical Preliminaries

In this section we recall some basic properties of set-valued functions and set-valued probabilities.

3.1 Set-Valued Functions

In this section we present some basic facts related to sets and set-valued functions. More details can be found in [2, 3]. In the sequel we denote by $\mathcal{K}$ the collection of all nonempty compact and convex subsets of $\mathbb{R}^d$. Addition of sets and scalar multiplication ($\lambda \in \mathbb{R}$) for $\mathcal{K}$ are defined by

$$A + B := \{a + b : a \in A, b \in B\} \text{ and } \lambda A = \{\lambda a : a \in A\}.$$

For $A \in \mathcal{K}$, we say that A *is nonnegative* ($A \geq 0$) if $0 \in A$. Given $A \in \mathcal{K}$ the *support function* $spt(\cdot, A) : \mathbb{R}^d \to \mathbb{R}$ is defined by

$$spt(p, A) = \sup\{p \cdot a : a \in A\}.$$

The support function completely defines A since

$$A = \bigcap_{\|p\|=1} \{x : x \cdot p \leq spt(p, A)\}. \tag{10}$$

Furthermore, $A \subseteq B$ if and only if $spt(p, A) \leq spt(p, B)$ for any $p \in S^1 = \{p : \|p\| = 1\}$. The function $spt(\cdot, \cdot)$ also satisfies the following properties: For all $\lambda \geq 0$ and $A, B \in \mathcal{K}$,

$$spt(p, \lambda A + B) = \lambda spt(p, A) + spt(p, B), \ spt(p, -B) = spt(-p, B)$$

but it is usually the case that $spt(p, -A) \neq -spt(p, A)$.

For any $A \in \mathcal{K}$, we can also define the *norm* of A using the support function as follows

$$\|A\| := \sup\{\|x\| : x \in A\} = \sup_{\|p\|=1} spt(p, A).$$

This definition satisfies all of the classical properties of a norm. There is a nice connection between the support function and the Hausdorff distance [7]: for $A, B \in \mathcal{K}$

$$d_H(A, B) = \sup_{\|p\|=1} |spt(p, A) - spt(p, B)|.$$

It is also the case that both addition and scalar multiplication on $\mathcal{K}$ are continuous in the Hausdorff distance.

A set $A \subset \mathbb{R}^d$ is *balanced* if $\lambda A \subseteq A$ for all $|\lambda| \leq 1$. For us a *unit ball* in $\mathbb{R}^d$ is any balanced set $\mathbb{B} \in \mathcal{K}$ with $0 \in int(\mathbb{B})$. Any such unit ball defines a norm on $\mathbb{R}^d$ via the Minkowski functional

$$\|x\| = \sup\{\lambda \geq 0 : \lambda x \in \mathbb{B}\}.$$

Given a unit ball $\mathbb{B}$, the *dual sphere* is defined as

$$\mathbb{S}^* = \{y : \sup\{y \cdot x : x \in \mathbb{B}\} = 1\} \subset \mathbb{R}^d$$

and is also a nonempty compact set. Notice that since $\mathbb{B}$ is compact, for each $y \in \mathbb{S}^*$, there is some $x \in \mathbb{B}$ with $y \cdot x = 1$.

A *set-valued function* or *multifunction* taking compact and convex values is a map from $\mathbb{R}^n$ to $\mathcal{K}$. For a given set-valued function $f : \mathbb{R}^n \to \mathcal{K}$ and measure μ, we can define the integral of f with respect to μ as an element of $\mathcal{K}$ via support functions using the property (see [3])

$$spt\left(q, \int_{\mathbb{R}^n} f(x) \, d\mu(x)\right) = \int_{\mathbb{R}^n} spt(q, f(x)) \, d\mu(x),$$

which defines the set as in (10). For more results on set-valued analysis see [3].

A set-valued map $f : \mathbb{R}^n \to \mathcal{K}$ is upper semicontinuous (u.s.c.) at $x_0 \in \mathbb{X}$ when for every $\varepsilon > 0$ there exists a neighborhood U of x_0 such that

$$f(x) \subseteq f(x_0) + \varepsilon\mathbb{B}$$

for every $x \in U$ (here $\mathbb{B}$ denotes the closed unit ball in $\mathbb{R}^d$). A set-valued map f is lower semicontinuous (l.s.c.) at $x_0 \in \mathbb{X}$ when for every $\varepsilon > 0$ there exists a neighborhood U of x_0 such that

$$f(x) \cap (f(x_0) + \varepsilon\mathbb{B}) \neq \emptyset$$

for every $x \in U$. A set-valued map f is Hausdorff continuous at $x_0 \in \mathbb{X}$ when it is both u.s.c. and l.s.c. at x_0. This is equivalent to saying that for every $\varepsilon > 0$ there exists a neighborhood U of x_0 such that

$$d_H(f(x), f(x_0)) < \varepsilon$$

for every $x \in U$. If the previous definitions holds for every $x_0 \in \mathbb{X}$, we will say that f is respectively upper semicontinuous, lower semicontinuous and continuous.

Given a compact subset Θ of $\mathbb{R}^n$ and a set-valued function $f : \Theta \subseteq \mathbb{R}^n \to \mathcal{K}$, consider the optimization problem

$$\max_{x\in\Theta} f(x). \tag{11}$$

We say that $x_0 \in \Theta$ is a *global maximizer for f over Θ* if for any $x \in \Theta$ we have $f(x) \subseteq f(x_0)$. Notice that we are using the natural ordering of sets given by inclusion. Let us now recall that a set-valued function $f : \mathbb{R} \to \mathcal{K}$ is *non decreasing* if

$$f(x) \subseteq f(y) \tag{12}$$

for $x, y \in \mathbb{R}$, $x \leq y$. A set-valued function $f : \mathbb{R}^n \to \mathcal{K}$ is *concave* if

$$tf(x) + (1 - t)f(y) \subseteq f(tx + (1 - t)y) \tag{13}$$

for $x, y \in \mathbb{R}^n$, $t \in [0, 1]$. Using the support function, we have that f is concave if and only if the function $spt(p, f(x))$ is concave for all $\|p\| = 1$.

3.2 Set-valued probabilities

In the sequel let us define by $\mathbb{X}$ a compact and convex subset of $\mathbb{R}^n$, (Ω, d) be a compact metric space, $f : \mathbb{X} \times \Omega \to \mathbb{R}$ be Borel measurable and continuous in x (we also sometimes will assume that it is jointly continuous or even Lipschitz). Ω models the set of all possible outcomes associated with a random parameter.

We now discuss the space of all possible probability measures supported on the set Ω. Let $\mathcal{B}$ be the Borel σ-algebra on Ω, then a *probability measure* on $(\Omega, \mathcal{B})$ with values in $[0, 1]$ is a function $\phi : \mathcal{B} \to [0, 1]$ such that $\phi(\emptyset) = 0$, $\phi(\Omega) = 1$, and

$$\phi\left(\bigcup_i A_i\right) = \sum_i \phi(A_i) \tag{14}$$

for any sequence of disjoint sets $A_i \in \mathcal{B}$. Let $\mathcal{M}$ be the space of all probability measures defined on Ω. The Monge-Kantorovich metric on $\mathcal{M}$ is defined as

$$d_M(\phi_1, \phi_2) = \sup_{f \in Lip_1(\Omega)} \left\{ \int_\Omega f(x) d\phi_1(x) - \int_\Omega f(x) d\phi_2(x) \right\} \tag{15}$$

where $Lip_1(\Omega)$ is the set of all Lipschitz functions defined on Ω with Lipschitz constant equal to 1 and ϕ_1 and ϕ_2 are in $\mathcal{M}$. It is well known that the metric space $(\mathcal{M}, d_M)$ is compact. In the special case where $\Omega \subset \mathbb{R}$, then it is known that

$$d_M(\phi_1, \phi_2) = \int_\Omega |F_{\phi_1}(x) - F_{\phi_2}(x)| \, dx, \tag{16}$$

where F_{ϕ_1} (respectively F_{ϕ_2}) is the CDF of ϕ_1 (respectively ϕ_2).

For any $\phi \in \mathcal{M}$, let us consider the stochastic optimization problem (see e.g. [38]):

$$\max_{x \in \mathbb{X}} \int_\Omega f(x, \omega) d\phi(\omega) := \max_{x \in \mathbb{X}} \mathbb{E}_\phi \left(f(x, \cdot) \right). \tag{17}$$

We now provide only basic definitions and those properties of set-valued probabilities that we will need; for more information and proofs see [1, 2, 3, 16, 17, 18, 37, 39]. The notion of set-valued probability extends the classical notion of probability. Within this context we suppose that the probability of a certain event is no longer a positive number but a set: this definition can be very useful to model situations in which there is total ignorance and absolutely no information about the system or subject under study. The classical axioms of probability theory are extended by means of the Minkowski sum and the inclusion order. This notion also extends the notion of imprecise probability [45] to the case of probabilities taking convex and compact-valued images. Within the imprecise probability formulation a single probability specification ϕ is replaced with an interval specification by means of lower and upper probabilities, namely ϕ_- and ϕ_+, and for a given event A we have that $\phi(A)$ is replaced by the positive interval $[\phi_-(A), \phi_+(A)] \subset [0, 1]$. The definition of imprecise probability recast in the framework of set-valued probability theory by assuming that the probability of a given event A is the interval $[-\phi_-(A), \phi_+(A)] \subset [-1, 1]$.

Given a set Ω and a σ-algebra $\mathcal{A}$ on Ω a *probability measure* on $(\Omega, \mathcal{A})$ with values in $[0, 1]$ is a function $\phi : \mathcal{A} \to [0, 1]$ such that $\phi(\emptyset) = 0$, $\phi(\Omega) = 1$, and

$$\phi \left(\bigcup_i A_i \right) = \sum_i \phi(A_i) \tag{18}$$

for any sequence of disjoint sets $A_i \in \mathcal{A}$. Similarly one can define the notion of *vector-valued probability measure* on $(\Omega, \mathcal{A})$. This is a function $\Phi : \mathcal{A} \to$

$[0,1]^d$, where $d \in \mathbb{N}$, such that $\Phi(\emptyset) = (0, ..., 0) \in \mathbb{R}^s$, $\Phi(\Omega) = (1, ..., 1) \in \mathbb{R}^d$, and

$$\Phi\left(\bigcup_i A_i\right) = \sum_i \Phi(A_i) \tag{19}$$

This last property is meant to be satisfied componentwise. Note that with this definition a vector-valued probability measure is simply a vector of probability measures.

Given a set Ω and a σ-algebra $\mathcal{A}$ on Ω a *set-valued measure* or *multimeasure* on $(\Omega, \mathcal{A})$ with values in $\mathcal{K}$ is a function $\Phi : \mathcal{A} \to \mathcal{K}$ such that $\Phi(\emptyset) = \{0\}$ and

$$\Phi\left(\bigcup_i A_i\right) = \sum_i \Phi(A_i) \tag{20}$$

for any sequence of disjoint sets $A_i \in \mathcal{A}$. The right side of (20) is the infinite Minkowski sum defined as

$$\sum_i K_i = \left\{\sum_i k_i : k_i \in K_i, \sum_i |k_i| < \infty\right\}.$$

One could also define the infinite sum by requiring the right side of (20) to converge in the Hausdorff distance.

We will say that a multimeasure Φ is *nonnegative*, and we write $\Phi(A) \geq 0$, if $0 \in \Phi(A)$ for all A. Nonnegative multimeasures are monotone: if $A \subseteq B$ then $\Phi(A) = \{0\} + \Phi(A) \subseteq \Phi(B \setminus A) + \Phi(A) = \Phi(B)$. This makes nonnegative multimeasures a nice generalization of (nonnegative) scalar measures.

The *total variation* of a multimeasure Φ is defined in the usual way as

$$|\Phi|(A) = \sup \sum_i \|\Phi(A_i)\|,$$

where the supremum is taken over all finite measurable partitions of $A \in \mathcal{A}$. The set-function $|\Phi|$ defined in this fashion is a (nonnegative and scalar) measure on Ω. If $|\Phi|(\Omega) < \infty$ then Φ is of *bounded variation.*

If Φ is a multimeasure and $p \in \mathbb{R}^d$ then the *scalarization* Φ^p defined by

$$\Phi^p(A) = spt(p, \Phi(A)) \tag{21}$$

is a signed measure on Ω in general and is a probability measure if Φ is a probability multimeasure.

One simple way to construct a multimeasure is by integrating a *set-valued density function* f with respect to a measure μ:

$$\Phi(A) = \int_A f(x)\ d\mu(x). \tag{22}$$

There are several approaches to defining this integral (see [3]). If the set-valued function f is nonnegative (that is, $0 \in f(x)$ for all x), then the resulting multimeasure will also be nonnegative. In addition, if $0 \leq f(x) \leq g(x)$ and Φ is a positive multimeasure, then

$$\int f(x)\ d\Phi(x) \subseteq \int g(x)\ d\Phi(x),$$

the convexity of the values of Φ is crucial.

Definition 5. *Let $\mathbb{B} \in \mathcal{K}$ be a unit ball. A $\mathbb{B}$* set-valued probability *or* probability multimeasure (or *pmm) on $(\Omega, \mathcal{A})$ is a nonnegative multimeasure Φ with $\Phi(\Omega) = \mathbb{B}$.*

A pmm Φ defines a parameterized family, Φ^p for $p \in \mathbb{S}^*$, of probability measures. However, in general Φ^p and Φ^q are related and the relationship can be quite complicated (the main constraint on this relationship is that $p \mapsto \Phi^p(A)$ is convex).

We can construct a pmm by integrating an appropriate density f against a finite measure μ, as in (22). For this to define a pmm we need f to satisfy some properties. The simplest conditions are to assume that $f(x) \in \mathcal{K}$ is balanced for each x, $\|f(x)\| \leq C$ for some C and all x, and

$$0 \in int \int_{\Omega} f(x)\ d\mu = int(\mathbb{B}).$$

Given a specific $\mathbb{B}$, it is difficult to find a density f which will give $\mathbb{B}$; it is better to use the integral of the density to define $\mathbb{B}$.

As usual, by a *random variable* on $(\Omega, \mathcal{A})$ we mean a Borel measurable function $X : \Omega \to \mathbb{R}$ and its *expectation* with respect to a pmm Φ is defined in the usual way as

$$\mathbb{E}_{\Phi}(X) = \int_{\Omega} X(\omega)\ d\Phi(\omega). \tag{23}$$

This integral can be constructed using support functions (that is, using the Φ^p) and each part of the decomposition $X = X^+ - X^-$ separately (since support functions work best with nonnegative scalars); see [18] for another approach. Since $0 \in \Phi(A)$ for each A, it is easy to see that $0 \in \mathbb{E}_{\Phi}(X)$ as well.

Classical results in probability theory as the strong law of Large Numbers, the Glivenko-Cantelli, and the Central Limit Theorems can be extended to pmm. More details on this topic can be found in [24].

We will denote by $\mathbf{M}$ the space of all probability multimeasures defined on Ω. The extended Monge-Kantorovich distance between $\Phi_1, \Phi_2 \in \mathbf{M}$, as defined in [22], is:

$$d_{\mathbf{M}}(\Phi_1, \Phi_2) = \sup_{\|p\|=1} d_M(\Phi_1^p, \Phi_2^p) \tag{24}$$

where d_M is the classical Monge-Kantorovich distance between two probabilities (as given in (15)).

4 Portfolio Efficiency with Incomplete Information and Partial Uncertainty

In the sequel we present two extensions of the notion of portfolio efficiency. The first case illustrates an extension of the notion of stochastic efficiency by means of the notion of set-valued utility while the second one considers the notion of set-valued probability.

4.1 Case 1: Incomplete information on the utility u

Let X_j be random variables which represent the possible assets and $Y_\lambda = \sum_i \lambda_i X_i$ is a portfolio with $\lambda \in \Lambda$ where we have

$$\Lambda = \left\{ \lambda \in \mathbb{R}^N_+ : \sum_{i=1}^N \lambda_i = 1 \right\}.$$

By using the notion of the expected value of a set-valued function, one can define the notion of stochastic portfolio efficiency that is based on the notion of stochastic dominance.

Definition 6. *Given $\hat{\lambda} \in \Lambda$, we say that $Y_{\hat{\lambda}}$ is stochastically efficient with respect to the probability ϕ and the set-valued utility $u \in \mathcal{U}$ if*

$$\hat{\lambda} = \operatorname{argmax}_{\lambda \in \Lambda} \mathbb{E}_\phi(u(Y_\lambda)) = \operatorname{argmax}_{\lambda \in \Lambda} \int_\Omega u(Y_\lambda(\omega)) d\phi(\omega) \tag{25}$$

where

$$\mathcal{U} = \{u : \mathbb{R} \to H_{cc}(\mathbb{R}^d), u \text{ is non-decreasing and concave}\} \tag{26}$$

and $H_{cc}(\mathbb{R}^d)$ is the collection of all nonempty compact intervals of $\mathbb{R}^d$. More precisely, $Y_{\hat{\lambda}}$ is stochastically efficient with respect to the probability ϕ and the set-valued utility $u \in \mathcal{U}$ if

$$\mathbb{E}_\phi(u(Y_\lambda)) \subseteq \mathbb{E}_\phi(u(Y_{\hat{\lambda}})) \tag{27}$$

for any $\lambda \in \Lambda$.

Let us also notice that if $u \in \mathcal{U}$ then $u^p(x) := spt(p, u(x))$ is a non-decreasing and concave function for any $p \in S^1$. This implies that if $Y_{\hat{\lambda}}$ is stochastically efficient with respect to the probability ϕ and the set-valued utility $u \in \mathcal{U}$ then $Y_{\hat{\lambda}}$ is stochastically efficient with respect to the probability ϕ and the utility function u^p for any $p \in S^1$. The converse of this implication is also true and fully characterizes the notion of stochastic efficiency with respect to a set-valued utility.

It is possible to extend the results presented in [34] for a classical utility function and in the case of a discrete probability space. The next section will be devoted to them.

We can now extend the notion of stochastic inefficiency presented in [25, 35]. First let us introduce the space

$$\mathcal{M}^* = \{\xi \in \mathcal{M} : \exists u \in \mathcal{U} \text{ such that}$$

$$\int_\Omega spt(p, u(\cdot))'(Y_{\hat{\lambda}}(\omega))(X_i(\omega) - Y_{\hat{\lambda}}(\omega))d\xi(\omega) \leq 0, i = 1, \ldots, N, \forall p \in S^1\}$$

and, for any probability $\phi \in \mathcal{M}$, let us define the function

$$SIM(\phi) = \min_{\xi \in \mathcal{M}^*} d_M(\xi, \phi) = d'_M(\phi, \mathcal{M}^*).$$

In other words $SIM(\phi)$ is the distance point-to-set between ϕ and the set $\mathcal{M}^*$. A given portfolio Y_λ is stochastically efficient relative to a given $\phi \in \mathcal{M}$ if and only if $SIM(\phi) = 0$. If $SIM(\phi) > 0$ we say the portfolio Y is stochastically inefficient.

4.2 Case 2: Partial uncertainty on the probability distribution

Similarly to the example presented in the previous section, let X_j be random variables which represent the possible assets and $Y_\lambda = \sum_i \lambda_i X_i$ is a portfolio with $\lambda \in \Lambda$ with the same definition of Λ as before. By using the notion of the expected value with respect to a set-valued probability Φ, one can define the notion of stochastic portfolio efficiency with respect to a set-valued probability.

Definition 7. *Given $\hat{\lambda} \in \Lambda$, we say that $Y_{\hat{\lambda}}$ is stochastically efficient with respect to the set-valued probability Φ if*

$$\hat{\lambda} = \text{argmax}_{\lambda \in \Lambda} \mathbb{E}_\Phi(u(Y_\lambda)) = \text{argmax}_{\lambda \in \Lambda} \int_\Omega u(Y_\lambda(\omega)) d\Phi(\omega) \tag{28}$$

for some utility function $u \in \mathcal{U}$ where

$$\mathcal{U} = \{u : \mathbb{R} \to \mathbb{R}, u \text{ is non-decreasing and concave}\}. \tag{29}$$

This corresponds to solve, for any $u \in \mathcal{U}$, the optimization problem:

$$\max_{\lambda \in \Lambda} \mathbb{E}_\Phi(u(Y_\lambda)). \tag{30}$$

This is similar to the problem presented in the previous section, but the expected value is now calculated with respect to a set-valued probability Φ.

Even in this case it is possible to generalize the results presented in [34]. And it is also possible to introduce the notion of stochastic inefficiency. Given the following space as follows:

$$\mathbf{M}^* = \{\Xi \in \mathbf{M} : \exists u \in \mathcal{U} \ \textit{such that}$$

$$\int_\Omega u'(Y_{\hat{\lambda}}(\omega))(X_i(\omega) - Y_{\hat{\lambda}}(\omega))d\Xi^p(\omega) \leq 0, i = 1...N, \forall p \in S^1\}.$$

and, for any set-valued probability $\Phi \in \mathbf{M}$, let us define the function

$$SIM(\Phi) = \min_{\Xi \in \mathbf{M}^*} d_{\mathbf{M}}(\Xi, \Phi) = d'_{\mathbf{M}}(\Phi, \mathbf{M}^*)$$

In other words $SIM(\Phi)$ is the distance point-to-set between ϕ and the set $\mathbf{M}^*$. A given portfolio Y_λ is stochastically efficient relative to a given $\Phi \in \mathbf{M}$ if and only if $SIM(\Phi) = 0$. If $SIM(\Phi) > 0$ we say the portfolio Y is stochastically inefficient.

5 Optimality and Stability Results

In this section we consider a general setting as follows. Find the optimal solution to the following problem:

$$\max_{x \in \mathbb{X}} \mathbb{E}_\phi \left(f(x, \cdot) \right) \tag{31}$$

where X is the feasible set (usually assumed to be compact). We assume that $(\Omega, \mathcal{F}, \phi)$ is a probability space, $f : \mathbb{X} \times \Omega \to R$ is the objective function depending on different possible scenarios $\omega \in \Omega$, and $\mathbb{E}_\phi$ is the expected value of $f(x, \cdot)$ with respect to a probability measure ϕ. Let us recall that Ω is the space of events, $\mathcal{F}$ is a σ-algebra of events defined on Ω, and ϕ is a probability measure defined over $\mathcal{F}$. Of course if we replace $f(x, \omega)$ with $u\left(\sum_i \lambda_i X_i(\omega)\right)$ and $\mathbb{X}$ with the compact set:

$$\Lambda = \left\{ \lambda \in \mathbb{R}^N_+ : \sum_{i=1}^N \lambda_i = 1 \right\}$$

then (31) reduces to the portfolio efficiency problem.

The motivation for considering such an extended formulation is quite simple: very often, in order to determine stability results of a certain mathematical problem, it is necessary to embed it into a more general family or into a more general space and use this setting to establish convergence results by outer approximations.

The results presented in the following subsections apply to this extended formulation.

5.1 The case of single-valued functions and single-valued probabilities

In this section we investigate the case in which the uncertainty is described by means of a classical probability measure. Our focus is then to investigate thye main properties of the problem

$$\max_{x \in \mathbb{X}} \mathbb{E}_\phi \left(f(x, \cdot) \right) \tag{32}$$

where $\phi \in M$ is a classical probability measure. These properties will be extended in the next sections to the case of set valued functions f and set-valued probabilities.
In order to analyze this problem we need to make some additional assumptions on f, in particular on its dependence on $\omega \in \Omega$.

Definition 8. *The function $f : \mathbb{X} \times \Omega \to \mathbb{R}$ is* lower semi-equicontinuous *if for all $x \in \mathbb{X}$ and $\epsilon > 0$ there is a $\delta > 0$ so that $f(y, \omega) > f(x, \omega) - \epsilon$ for all $y \in B_\delta(x)$ and all ω. We say that f is* upper semi-equicontinuous *if $-f$ is lower semi-equicontinuous.*

The following observation is simple but is important enough for us to record.

Proposition 2. *Let $f : \mathbb{X} \times \Omega \to R$ be such that $\{f(\cdot, \omega) : \omega \in \Omega\}$ is an upper (lower) semi-equicontinuous family and $\|f(x, \cdot)\|_{L^1(\Omega,\phi)} \leq K$ for all $x \in \mathbb{X}$. Then $G(x) = \mathbb{E}_\phi(f(x, \omega))$ is upper (lower) semi-continuous.*

Linearity of the expectation implies that if f is concave in x for all $\omega \in \Omega$ then $G(x) = \mathbb{E}_\phi(f(x, \cdot))$ is also a concave map. The following results explore the convergence of the above model under perturbation of the integrand f and the probability measure ϕ and their implications on the convergence of sequences of maximizers.

Proposition 3. *Let $f_n, f : \mathbb{X} \times \Omega \to \mathbb{R}$ with $f_n(x, \cdot) \to f(x, \cdot)$ in $L^1(\Omega, \phi)$ uniformly over $x \in \mathbb{X}$. Then $G_n(x) = \mathbb{E}_\phi(f_n(x, \cdot))$ uniformly converges to $G(x) = \mathbb{E}_\phi(f(x, \cdot))$ in the sup distance. In particular, if $f_n \to f$ uniformly over $\mathbb{X} \times \Omega$ then the conclusion holds.*

Proof. The proof is quite straightforward and it follows from the definitions and the fact that ϕ is a probability (and so finite) measure. ■

Proposition 4. *Let $f_n : \mathbb{X} \times \Omega \to \mathbb{R}$ be a sequence such that $f_n(x, \cdot) \to f(x, \cdot)$ in $L^1(\Omega, \phi)$ uniformly over $x \in \mathbb{X}$. If for each n, $x_n \in \mathbb{X}$ is a maximizer of $G_n(x) = \mathbb{E}_\phi(f_n(x, \cdot))$, then for every subsequence x_{n_k} converging to a point $\bar{x}$ it holds that $\bar{x}$ is a maximizer for $G(x)$.*

Proof. Since x_n is a maximizer of G, uniform convergence of G_n to G (by Proposition 3) implies that

$$G(\bar{x}) = \lim_{k\to+\infty} G_{n_k}(x_{n_k}) \geq \lim_{k\to+\infty} G_{x_k}(x) = G_{n_k}(x), \ \forall x \in \mathbb{X} \tag{33}$$

that is $\bar{x}$ is a maximizer of G. ∎

The proofs of the next results are straightforward adaptations of standard arguments and thus we omit them.

Proposition 5. *Suppose that $\phi_n \in \mathcal{M}$ converge to ϕ in the Monge-Kantorovich distance. Suppose further that f is uniformly Lipscthitz in ω (i.e., $|f(x,\omega_1) - f(x,\omega_2)| \leq Kd(\omega_1,\omega_2)$ for all $x \in \mathbb{X}$). Then $G_n \to G(x)$ uniformly over $\mathbb{X}$. Furthermore, if x_n is a maximizer for G_n over $\mathbb{X}$ and $x_{n_k} \to \bar{x}$ then $\bar{x}$ is a maximizer for G over $\mathbb{X}$.*

Proposition 6. *Suppose that $\phi_n \to \phi$ in the Monge-Kantorovich distance and $f_n \to f$ uniformly on $\mathbb{X} \times \Omega$ with $|f(x,\omega_1) - f(x,\omega_2)| \leq Kd(\omega_1,\omega_2)$ for all $x \in \mathbb{X}$. Then $G_n \to G$ uniformly on $\mathbb{X}$. Furthermore, if x_n is a maximizer for G_n over $\mathbb{X}$ and $x_{n_k} \to \bar{x}$ then $\bar{x}$ is a maximizer for G over $\mathbb{X}$.*

5.2 The case of Set-Valued Functions

Now we extend problem (32) to the case in which f is a set-valued function. We investigate the main properties of the problem

$$\max_{x\in\mathbb{X}} \mathbb{E}_\phi \left(f(x,\cdot)\right) \tag{34}$$

where $\phi \in M$ is a classical probability measure and $f : \mathbb{X} \times \Omega \to \mathcal{K}$ is a set-valued function taking compact and convex values. This problem can be seen as a way of taking into account the fact that exact values of function f are not known. In the following of this section let us define $G(x) := \mathbb{E}_\phi\left(f(x,\cdot)\right)$ and $G^p(x) := spt\left(p, \mathbb{E}_\phi\left(f(x,\cdot)\right)\right) = \mathbb{E}_\phi\left(spt(p, f(x,\cdot))\right)$.
First we state continuity and concavity properties for Problem (34).

Proposition 7. *Suppose that $f(x,\omega) \subseteq K$ for all $(x,\omega) \in \mathbb{X} \times \Omega$. Then $G(x) \subseteq K$ for all $x \in \mathbb{X}$ and consequently $G^p(x)$ is bounded by $\|K\|$ for any $x \in \mathbb{X}$ and any $p \in S^1$.*

Proof. The inclusion $f(x,\omega) \subseteq K$ implies that $spt(p, f(x,\omega)) \leq spt(p,K)$ for any $p \in S^1$ and thus

$$spt(p, G(x)) = \int_\Omega spt(p, f(x,\omega)) \ d\phi(\omega) \leq \int_\Omega spt(p,K) \ d\phi(\omega) = spt(p,K)$$

which implies $G(x) \subseteq K$. ∎

The extension of the definitions of upper and lower semi-equicontinuous (Definition 8) to a multifunction $f : \mathbb{X} \times \Omega \to \mathcal{K}$ is straightforward. We use them in the following result.

Proposition 8. *Suppose that the set-valued map f is upper (lower) semi-equicontinuous. Then the map $x \to spt(p, f(x, \omega))$ is upper (lower) semi-continuous for all $p \in S^1$. Furthermore, both $G^p(x)$ and $G(x)$ are u.s.c. (l.s.c.) as well. Finally, if f is both upper and lower semi-equicontinuous then G^p is continuous for all $p \in S^1$ and G is Hausdorff continuous.*

Proof. Since f is upper semi-equicontinuous at x_0, for every $\epsilon > 0$ there exists a neighborhood U of x_0 such that

$$f(x) \subseteq f(x_0) + \epsilon\mathbb{B}$$

for every $x \in U$ and for every $\omega \in \Omega$ (where $\mathbb{B} \subset \mathbb{R}^d$ is the unit ball). It follows

$$spt(p, f(x, \omega)) \leq spt(p, f(x_0, \omega)) + \epsilon$$

for every $x \in U$ and $\omega \in \Omega$, that is the supprt function is upper semi-equicontinuous at x_0. This implies that

$$G^p(x) \leq G^p(x_0) + \epsilon$$

and that

$$G(x) \subseteq G(x_0) + \epsilon\mathbb{B}$$

for every $x \in U$, which concludes the proof. ■

Proposition 9. *Suppose that f takes polyhedral values, that is there exist vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ and $f_i : \mathbb{X} \times \Omega \to \mathbb{R}_+$, $i = 1..m$, such that*

$$f(x, \omega) = \text{conv}\{\mathbf{v}_i f_i(x, \omega)\} \tag{35}$$

where $\text{conv}\, A$ denotes the convex hull of the set A. Suppose that $\bar{x} \in \mathbb{X}$ is such that

$$\bar{x} = \text{argmax}_{x \in \mathbb{X}} \mathbb{E}_\phi \left(f_i(x, \cdot)\right), \textit{ for all } i = 1, 2, \ldots, m. \tag{36}$$

Then $\bar{x}$ is the maximizer of G over $\mathbb{X}$ and

$$G(\bar{x}) = \text{conv}\{\mathbf{v}_i \mathbb{E}_\phi \left(f_i(\bar{x}, \cdot)\right)\} \tag{37}$$

Proof. The proof is straightforward and it follows by noticing that $G(x) \subset G(x_0)$ if and only if $G^{\mathbf{v}_i}(x) \leq G^{\mathbf{v}_i}(x_0)$ for any $x \in \mathbb{X}$ and any $i = 1, \ldots, m$. ■

Proposition 10. *Let $f(x, \omega)$ be a concave set-valued function w.r.t. x for all ω. Then $G^p(x) = \mathbb{E}_\phi \left(spt(p, f(x, \cdot))\right)$ is a concave function for every p with $\|p\| = 1$ and $G(x) = \mathbb{E}_\phi \left(f(x, \cdot)\right)$ is a concave set-valued map.*

Proof. Concavity of f in the first variable x gives

$$spt(p, f(tx + (1-t)y, \omega) \geq tspt(p, f(x, \omega)) + (1-t)spt(p, f(y, \omega))$$

for every $t \in [0,1]$ and $x, y \in \mathbb{X}$. Integrating we get

$$G^p(tx + (1-t)y) \geq tG^p(x) + (1-t)G^p(y)$$

i.e. concavity of G^p and consequently concavity of the set-valued map G. ■

Next result states optimality conditions for Problem (34).

Proposition 11. *Let f be a set-valued map and suppose that $x_0 \in \mathbb{X}$ is a maximizer of $G(x) = \mathbb{E}_\phi (f(x, \cdot))$ over $\mathbb{X}$. Suppose that for any p such that $\|p\| = 1$ we have that $x \to spt(p, f(x, \omega))$ is C^1 for all ω. Then*

$$\mathbb{E}_\phi(\nabla spt(p, f(x, \cdot))d) \leq 0 \tag{38}$$

for any direction $d \in K_{\mathbb{X}}(x_0) = \{d \in \mathbb{R}^n : \exists \bar{t} > 0, x_0 + td \in \mathbb{X}, \forall t \in [0, \bar{t}]\}$ and for any p such that $\|p\| = 1$. If furthermore f is concave, and condition (38) holds, then x_0 is a maximizer for problem (34).

Proof. As $x_0 \in \mathbb{X}$ is a maximizer with respect to the inclusion ordering, we get that

$$G(x) \subseteq G(x_0) \tag{39}$$

for any $x \in \mathbb{X}$. This implies that $x_0 \in \mathbb{X}$ is a maximizer for the map $G^p(x) = spt(p, G(x)) = \mathbb{E}_\phi (p, spt(f(x, \cdot)))$. The regularity hypothesis on $spt(p, f(x, \omega))$ allows to conclude that G^p is C^1 as well. A classical optimality condition implies then that $\nabla G^p(x)d \geq 0$ for any internal direction $d \in K_{\mathbb{X}}(x_0)$ which turns out to imply that

$$\mathbb{E}_\phi(\nabla spt(p, f(x_0, \cdot))d) \leq 0 \tag{40}$$

for any $p \in S^1$ and $d \in K_{\mathbb{X}}(x_0)$. The proof of the second part is immediate since $G^p(x)$ are concave.

■

The following corollary can be easily deduced from the above result by replacing the function $f(x, \omega)$ with $u\left(\sum_i \lambda_i X_i(\omega)\right)$. It also extends the results presented in [34] for a classical utility function and in the case of a discrete probability space.

Corollary 1. *A given portfolio $Y_{\hat{\lambda}}$ is stochastically efficient with respect to the probability ϕ and the set-valued utility $u \in \mathcal{U}$ if and only if it obeys the following first-order optimality conditions:*

$$\mathbb{E}_\phi \left(spt(p, u(\cdot))'(Y_{\hat{\lambda}}(\cdot)) \left[X_i(\cdot) - Y_{\hat{\lambda}}(\cdot) \right] \right) \leq 0 \tag{41}$$

for all $i = 1 \ldots N$ and for any $p \in S^1$ provided that $spt(p, u(\cdot))'(Y_{\hat{\lambda}}(\cdot))$ exists.

The next results are direct analogues of Propositions 3 - 6 from section 5.1. The technique of using the scalarizations makes the proofs of the set-valued results very similar to those for the scalar-valued cases.

Proposition 12. *Let $f_n, f : \mathbb{X} \times \Omega \to \mathcal{K}$ with $\mathbb{E}_\phi(d_H(f_n(x, \cdot), f(x, \cdot))) \to 0$ uniformly over $x \in \mathbb{X}$. Then the set-valued maps G_n converge uniformly over $\mathbb{X}$ to G in the Hausdorff distance. In particular, if $f_n \to f$ in the Hausdorff distance uniformly over $\mathbb{X} \times \Omega$ then the conclusion holds.*

Proof. We have, for $p \in S^1$ and all $x \in \mathbb{X}$ that

$$\begin{aligned} |G_n^p(x) - G^p(x)| &= \left| \int_\Omega spt(p, f_n(x, \omega)) - spt(p, f(x, \omega)) \; d\phi(\omega) \right| \\ &\leq \int_\Omega |spt(p, f_n(x, \omega)) - spt(p, f(x, \omega))| \; d\phi(\omega) \\ &\leq \mathbb{E}_\phi(d_H(f_n(x, \omega), f(x, \omega))). \end{aligned}$$

Hence

$$\sup_{x \in \mathbb{X}} d_H(G_n(x), G(x)) = \sup_{x \in \mathbb{X}} \sup_{p \in S^1} |G_n^p(x) - G^p(x)| =$$

$$\sup_{p \in S^1} \sup_{x \in \mathbb{X}} |G_n^p(x) - G^p(x)| \to 0$$

■

Proposition 13. *Let f_n be a sequence of multifunctions such that we have $\mathbb{E}_\phi(d_H(f_n(x, \cdot), f(x, \cdot))) \to 0$ uniformly over $x \in \mathbb{X}$. If $x_n \in \mathbb{X}$ is a maximizers of $G_n(x)$, for every subsequence x_{n_k} converging to a point $\bar{x}$ it holds that $\bar{x}$ is a maximizer for $G(x)$.*

Proof. Since x_n is a maximizer of G_n^p for every $p \in S^1$, uniform convergence of f_n to f gives,

$$G^p(\bar{x}) = \lim_{k \to +\infty} G_{n_k}^p(x_{n_k}) \geq \lim_{k \to +\infty} G_{x_k}^p(x) = G_{n_k}^p(x), \; \forall x \in \mathbb{X} \tag{42}$$

that is $\bar{x}$ is a maximizer of G^p for every $p \in S^1$ and hence a maximizer of G. ■

Proposition 14. *Suppose that $\phi_n \to \phi$ in the Monge-Kantorovich distance and $d_H(f(\omega_1, x), f(\omega_2, x)) \leq K d(\omega_1, \omega_2)$ for all $x \in \mathbb{X}$. Then $G_n(x)$ converges to $G(x)$ in the Hausdorff distance uniformly over $\mathbb{X}$. Furthermore, if x_n is a maximizer for G_n over $\mathbb{X}$ and $x_{n_k} \to \bar{x}$, then $\bar{x}$ is a maximizer for G over $\mathbb{X}$.*

Proof. Computing, we have for any $x \in \mathbb{X}$:

$$
\begin{aligned}
d_H(G_n(x), G(x)) &= \sup_{\|p\|=1} |spt(p, G_n(x)) - spt(p, G(x))| \\
&\leq \sup_{\|p\|=1} \int_\Omega |spt(p, f(x,\omega))| \, d(\phi_n(\omega) - \phi(\omega)) \quad (43) \\
&\leq K d_M(\phi^n, \phi) \to 0, \quad (44)
\end{aligned}
$$

and this implies the thesis. ∎

Proposition 15. *Suppose that $\phi_n \to \phi$ in the Monge-Kantorovich distance and $d_H(f_n(x,\omega), f(x,\omega)) \to 0$ uniformly over $\mathbb{X} \times \Omega$. Then $G_n \to G$ in the Hausdorff distance uniformly over $\mathbb{X}$. Furthermore, if x_n is a maximizer for G_n over $\mathbb{X}$ and $x_{n_k} \to \bar{x}$ then $\bar{x}$ is a maximizer for G over $\mathbb{X}$.*

Proof. Omitted as immediate. ∎

By replacing $f(x,\omega)$ with $u\left(\sum_i \lambda_i X_i(\omega)\right)$ from the previous results one can prove stability properties for the portfolio efficiency problem. These properties are relevant to investigate computational methods for portfolio efficiency with incomplete information on the utility function u.

5.3 The case of Set-Valued Probability

This section is devoted to the case in which the underlying level of uncertainty is modelled through a set-valued probability measure. In particular when Φ maps into $2^{\mathbb{R}}$, it describes uncertainty w.r.t. the "true" classical probability distribution involved in the stochastic optimization problem. In other words, we extend problem (32) focusing on the solution to the problem

$$\max_{x \in \mathbb{X}} \mathbb{E}_\Phi \left(f(x, \cdot) \right) \quad (45)$$

Let us notice that, for each $\Phi \in \mathbf{M}$ (here $\mathbf{M}$ denotes the space of all probability multimesures on Ω) the problem

$$\max_{x \in \mathbb{X}} \mathbb{E}_\Phi \left(f(x, \cdot) \right) \quad (46)$$

is a set-valued optimization program. In the following let us define $G(x) := \mathbb{E}_\Phi \left(f(x, \cdot) \right)$ and $G^p(x) := spt\left(p, \mathbb{E}_\Phi \left(f(x, \cdot) \right)\right) = \mathbb{E}_{\Phi^p} \left(f(x, \cdot) \right)$.

As in the previous section we begin studying continuity and convexity properties for Problem (46).

Proposition 16. *Suppose that f is uniformly bounded by a constant K for any $x \in \mathbb{X}$ and $\omega \in \Omega$. Then $\|G(x)\|$ is bounded by K for any $x \in \mathbb{X}$.*

Proof. For any $p \in S^1$, we have that

$$|spt(p, G(x))| = |G_p(x)| \leq \int_\Omega |f(x,\omega)| d\Phi^p \leq K\Phi^p(\Omega) = K$$

and then

$$\|G(x)\| = \sup_{\|p\|=1} |spt(p, G(x))| \leq K.$$

■

The following results state some regularity conditions for the maps G and G^p.

Proposition 17. *Suppose that f is upper (lower) semi-equicontinuous. Then both $G(x)$ and $G^p(x)$, for all p, are upper (lower) semicontinuous as well. If f is equicontinuous then G^p is continuous and G is Hausdorff continuous.*

Proof. By assumption for every $x_0 \in \mathbb{X}$ and $\varepsilon > 0$ there exists a neighbourhood U of x_0 such that

$$f(x,\omega) \leq f(x_0,\omega) + \varepsilon \tag{47}$$

for every $x \in U$, $\forall \omega \in \Omega$. Hence, for $x \in U$ and $p \in \mathbb{R}^d$ with $\|p\| = 1$ it holds

$$\mathbb{E}_{\Phi^p}(f(x,\cdot)) = \int_\Omega f(x,\omega) d\Phi^p \leq \int_\Omega (f(x_0,\omega) + \varepsilon) d\Phi^p =$$

$$= \int_\Omega f(x_0,\omega) d\Phi^p + \varepsilon = \mathbb{E}_{\Phi^P}(f(x,\cdot)) + \varepsilon,$$

which means that $G^p(x)$ is u.s.c. at x_0. This is equivalent to say that

$$\int_\Omega f(x,\omega) d\Phi \subseteq \int_\Omega f(x_0,\omega) d\Phi + \varepsilon\mathbb{B}.$$

The proof for the lower semicontinuous case is analogous. ■

More on this can be found in [3].

Proposition 18. *Suppose that $x_0 \in \mathbb{X}$ is a maximizer. Then, for any $p \in S^1$, x_0 is a global solution to the problem*

$$\max_{x \in \mathbb{X}} G^p(x).$$

Conversely if x_0 is a maximizer for $G^p(x)$ over $\mathbb{X}$ for all $\|p\| = 1$. Then x_0 is also a maximizer for $G(x)$ over $\mathbb{X}$.

Proof. The result follows by using the equivalence $G(x) \subseteq G(x_0)$ iff $G^p(x) \leq G^p(x_0)$ for all $\|p\| = 1$. ■

Remark 1. *Solving problem (45) means to hedge against uncertainty about a probability distribution, which can be seen since a solution x_0 of problem (45) is a minimizer for every stochastic optimization problem of the form (17) when the probability distribution ranges in $P = \{\Phi^p, p \in S^1\}$.*

This condition might be too complicated to be checked as it requires to verify the optimality of x_0 for all $p \in S^1$. This context can be simplified under the hypothesis that Φ is a set-valued probability taking polyhedral values in a specific way. Suppose that the exists $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ vectors in $\mathbb{R}^d$ such that

$$\Phi(A) = \text{conv}\{\phi_s(A)\mathbf{v}_s, s = 1...m\}, \tag{48}$$

where ϕ_s are classical probability measures. Then it is easy to check that in this case

$$G(x) = \text{conv}\left\{G_s(x)\mathbf{v}_s, s = 1...m\right\}$$

where

$$G_s(x) = \int_\Omega f(x,\omega)d\phi_s(\omega).$$

The following result states a sufficient condition under the hypothesis of a set-valued probability taking polyhedral values.

Proposition 19. *Suppose that $\Phi(A)$ is defined as in (48) and that $x_0 \in \mathbb{X}$ is simultaneously a solution to all the problems*

$$\max_{x \in \mathbb{X}} G_s(x)$$

for $s = 1, 2, \ldots, m$. Then $x_0 \in \mathbb{X}$ is a maximizer of $G(x)$ over $x \in \mathbb{X}$.

Proof. The proof follows from the observation that $G(x) \subseteq G(x_0)$ iff $G_i(x) \leq G_i(x_0)$ for all $i = 1, 2, \ldots, m$. ■

There is a geometrical condition we can use to check whether a given point x_0 is a maximizer for G over $\mathbb{X}$. For any $p \in S^1$, let $x^p_{\max}$ be the maximizer of G^p over $\mathbb{X}$ and construct the following set:

$$A_{\max} = \bigcap_{\|p\|=1} \{x : x \cdot p \leq G^p(x^p_{\max})\}. \tag{49}$$

The set $A_{\max}$ is the maximal set as it contains any $G(x)$ for any $x \in \mathbb{X}$. The following result is then immediate.

Proposition 20. *If there exists $x_0 \in \mathbb{X}$ such that $G(x_0) = A_{\max}$ then x_0 is a global maximizer of G over $\mathbb{X}$.*

Proposition 21. *Let $f(x, \omega)$ be a concave function w.r.t. x for all ω and Φ a probability multimeasure. Then $G^p(x)$ is concave for every p and $G(x)$ is a concave multifunction.*

Proof. Since f is concave with respect to x, it holds that

$$f(tx_1 + (1-t)x_2, \omega) \geq tf(x_1, \omega) + (1-t)f(x_2, \omega)$$

for every $x_1, x_2 \in X$ and $t \in [0,1]$. Hence for p with $\|p\| = 1$ it holds

$$\begin{aligned}\int_\Omega f(tx_1 + (1-t)x_2, \omega) d\Phi^p &\geq \int_\Omega (tf(x_1, \omega) + (1-t)f(x_2, \omega))\, d\Phi^p \\ &= t\int_\Omega f(x_1, \omega) d\Phi^p + (1-t)\int_\Omega f(x_2, \omega) d\Phi^p\end{aligned}$$

which gives concavity of G^p. This entails the concavity of the set-valued map G. ■

Next proposition gives optimality conditions for Problem (46).

Proposition 22. *Let Φ be a probability multimeasure. Suppose that $f(x,\omega)$ is C^1 as a function of x for all ω and $x_0 \in \mathbb{X}$ be a maximizer. Then*

$$0 \in \mathbb{E}_\Phi(\nabla f(x_0, \cdot)d) \tag{50}$$

for any internal direction $d \in K_\mathbb{X}(x_0)$. Furthermore if f is concave, and condition (50) holds, then x_0 is a maximizer for problem (45).

Proof. As $x_0 \in \mathbb{X}$ is a maximizer with respect to the inclusion ordering, we get that

$$G(x) \subseteq G(x_0)$$

for any $x \in \mathbb{X}$. This implies that $x_0 \in \mathbb{X}$ is a maximizer for the map $G^p(x)$ for all p. The regularity hypothesis on f allows us to conclude that G^p is C^1 as well. A classical optimality condition implies then that $\nabla G^p(x)d \geq 0$ for any internal direction $d \in K_\mathbb{X}(x_0)$ which in turn implies that

$$\mathbb{E}_{\Phi^p}(\nabla f(x, \cdot)d) \leq 0$$

for any $p \in S^1$. This result allows to conclude that

$$0 \in \mathbb{E}_\Phi(\nabla f(x, \cdot)d)$$

that is the thesis. The proof of the second part is immediate since $G^p(x)$ are concave. ■

The following corollary can be easily deduced from the above result by replacing $f(x,\omega)$ with $u\left(\sum_i \lambda_i X_i(\omega)\right)$. It also extends the results presented in [34] for a classical utility function and in the case of a discrete probability space.

Corollary 2. *A given portfolio $Y_{\hat{\lambda}}$ is stochastically efficient with respect to the set-valued probability Φ and the utility $u \in \mathcal{U}$ if and only if it obeys the following first-order optimality conditions:*

$$\mathbb{E}_{\Phi^p}\left(u'(Y_{\hat{\lambda}}(\cdot))\left[X_i(\cdot) - Y_{\hat{\lambda}}(\cdot))\right]\right) \leq 0 \tag{51}$$

where $\Phi^p = spt(p, \Phi)$ is a classical probability measure and this inequality holds for all $i = 1...N$ and for any $p \in S^1$ (provided that $u'(Y_{\hat{\lambda}}(\cdot))$ exists).

We close this subsection with stability results for problem (46).

Proposition 23. *Let Φ be a probability multimeasure and $f_n \to f$ uniformly over $\mathbb{X} \times \Omega$. Then $G_n(x) \to G(x)$ uniformly on $\mathbb{X}$.*

Proof. The proof is quite straightforward. Let us first observe that $G_n^p(x)$ converges to $G^p(x)$ uniformly, as

$$\begin{aligned}
|G_n^p(x) - G^p(x)| &= \left| spt\left(p, \int_\Omega f_n(x,\omega) - f(x,\omega)\ d\Phi(\omega)\right)\right| \\
&= \left|\int_\Omega f_n(x,\omega) - f(x,\omega) d\Phi^p(\omega)\right| \\
&\leq \int_\Omega |f_n(x,\omega) - f(x,\omega)| d\Phi^p(\omega) \\
&\leq |f_n(x,\omega) - f(x,\omega)| \to 0
\end{aligned}$$

for any $p \in S^1$. Now the thesis follows by recalling that

$$\begin{aligned}
\sup_{x\in\mathbb{X}} d_H(G_n(x), G(x)) &= \sup_{x\in\mathbb{X}} \sup_{p\in S^1} |G_n^p(x) - G^p(x)| \\
&\leq \sup_{p\in S^1} \sup_{x\in\mathbb{X}} |G_n^p(x) - G^p(x)| \to 0.
\end{aligned}$$

■

Proposition 24. *let Φ be a probability multimeasure an $f_n \to f$ uniformly on $\mathbb{X} \times \Omega$. If $x_n \in \mathbb{X}$ is a sequence of maximizers of the map $G_n(x)$, for every subsequence x_{n_k} converging to a point $\bar{x}$ it holds that $\bar{x}$ is a maximizer for $G(x)$.*

Proof. Since x_n is a maximizer of G^p for every $p \in S^1$, uniform convergence of f_n to f gives, by Proposition 23

$$G^p(\bar{x}) = \lim_{k\to+\infty} G_{n_k}^p(x_{n_k}) \geq \lim_{k\to+\infty} G_{x_k}^p(x) = G_{n_k}^p(x),\ \forall x \in \mathbb{X}$$

that is $\bar{x}$ is a maximizer of G^p for every $p \in S^1$ and hence a maximizer of G. ■

Proposition 25. *Suppose that $\Phi_n \to \Phi$ in the Monge-Kantorovich distance for probability multimeasures. Suppose further that $f_n \to f$ uniformly on $\mathbb{X} \times \Omega$ and $|f(x,\omega_1) - f(x,\omega_2)| \leq Kd(\omega_1,\omega_2)$ for all $x \in \mathbb{X}$. Then $G_n \to G$ in the Hausdorff distance uniformly over $\mathbb{X}$. Furthermore, if x_n is a maximizer for G_n over $\mathbb{X}$ and $x_{n_k} \to \bar{x}$, then $\bar{x}$ is a maximizer for G over $\mathbb{X}$.*

Proof. Computing, we have for any $x \in \mathbb{X}$:

$$\begin{aligned} d_H(G_n(x), G(x)) &= \sup_{\|p\|=1} |spt(p, G_n(x)) - spt(p, G(x))| \\ &\leq \sup_{\|p\|=1} \int_\Omega |f(x,\omega)| \, d(\Phi_n^p(\omega) - \Phi^p(\omega)) \\ &\leq d_{\mathbf{M}}(\Phi^n, \Phi) \to 0 \end{aligned}$$

and thus $G_n \to G$ uniformly. Once this is established the second part follows as in the proof of Proposition 24. ■

Stability properties for the portfolio efficiency problem easily follow from the above results by replacing $f(x,\omega)$ with $u\left(\sum_i \lambda_i X_i(\omega)\right)$. These properties might be important to prove convergence results for computational methods with partial uncertainty on the probability distribution.

6 Managerial Insights and Implementation

Dealing with uncertainty is one of the most challenging issues in the decision making process. Decision makers have to include different forms of exogenous noise and errors which can be generated from data, from estimation techniques, from exogenous shocks, and so on.

Assuming different forms of the underlying distribution as well as relying on sampling techniques can partially overcome these difficulties. In our approach, instead, we propose somehow to include noise in the model itself. Our model has been inspired by other existing approaches to deal with noise, inconsistent, imprecise, and uncertain knowledge such as fuzzy set theory or rough set theory.

In fuzzy decision making, for instance, the information in complex and uncertain environments is modelled by means of fuzzy sets and the notion of membership function plays a key role.

In rough set theory, the information is approximated by two precise boundary lines which describe the imprecise concepts. A rough set is then approximated by two regular sets.

In our approach we assume that information is described by compact and convex sets which include any possible form of noise.

From the decision maker point-of-view there are several benefits of adopting this approach to portfolio management. Firstly, one positive aspect is related to the structural stability of the method. The optimal solution can be seen as an averaged solution since the optimization model considers several and different scenarios simultaneously. Secondly, statistical perturbations and data noise are somehow neutralized and estimations are less dependent on the sample size and the statistical approach. Thirdly, this approach can easily be controlled via upper and lower optimization problems.

From a computational perspective this approach is not more complicated to be implemented or more computationally challenging than other existing techniques. In the simple case of interval-valued utility or interval-valued probability, it relies on the solution of two optimization problems including the upper and the lower limits, respectively. More precisely, if u is a given positive interval-valued utility $[u_-, u_+]$ ($u_- < 0 < u_+$) and ϕ is a probability measure, then portfolio efficiency requires to solve the optimization problem:

$$\min_{\lambda \in \Lambda} \mathbb{E}_\phi(u(Y_\lambda)) \tag{52}$$

This can be split into the following two problems:

$$\min_{\lambda \in \Lambda} -\mathbb{E}_\phi(u^-(Y_\lambda)) \tag{53}$$

$$\min_{\lambda \in \Lambda} \mathbb{E}_\phi(u_+(Y_\lambda)) \tag{54}$$

In fact, it is evident that if $\hat{\lambda}$ is a common solution to both problems Eq. (53) and (54) then $\hat{\lambda}$ solves problem Eq. (52).

The same kind of passages can be repeated in the case of interval-valued probabilities in which the model formulation can be simplified by means of lower and upper probabilities.

7 Conclusion

In this paper we have modelled two different extensions of the notion of portfolio efficiency with incomplete information and partial uncertainty. Both these formulations make use of the notion of set-valued function and set-valued probability measure to model the lack of certainty on the objective function and on the underlying probability distribution and both of them are formulated as set-valued optimization problems by construction. We have proved stability results, optimality conditions, and provided different scalarization techniques.

Future research avenues include:

- an extension of these approaches to the case in which the inclusion ordering is replaced by a more general ordering induced by a convex closed cone with nonempty interior

- the extension of the set-valued approach to the case of higher-order stochastic dominance

- the study of empirical tests for set-valued stochastic dominance.

References

[1] Z. Artstein, Set-valued measures, *Transactions of the American Mathematical Society*, **165** (1972), 103-125.

[2] Z. Artstein, On the calculus of closed set-valued functions, *Indiana Univeristy Mathematical Journal*, **24** (1974), 433-441.

[3] J.P. Aubin, H. Frankowska, *Set-valued analysis*, Birkhäuser, Boston, 1990.

[4] F. Ben Abdelaziz, H. Masri, Stochastic programming with fuzzy linear partial information on probability distribution, *European Journal of Operational Research*, **162**, (2005), no. 3, 619-629.

[5] F. Ben Abdelaziz, H. Masri, A compromise solution for the multiobjective stochastic linear programming under partial uncertainty, *European Journal of Operational Research*, **202**, (2010), 55-59.

[6] F. Ben Abdelaziz, D. La Torre, Robust generalized Merton-type financial portfolio models with generalized utility, *Annals of Operations Research*, (2021), in press. https://doi.org/10.1007/s10479-021-04051-x

[7] G. Beer, *Topologies on closed and closed convex sets*, Kluwer, Netherlands, 1993.

[8] L. A. Ben-Tal, L. El Ghaoui, A. Nemirovski, *Robust optimization*, Princeton University Press, USA, 2009.

[9] G. Debreu, D. Schmeidler (1967) The Radon-Nikodym derivative of a correspondence. In: Lecam, L., Neyman, J., Scott, E.L. (eds.) Proceedings of Fifth Berkeley Symposium Math.Stat. and Probability. I, Part 1, pp. 351–372. Berkely: University of California Press.

[10] Dempster, A. P (1966) New approaches for reasoning towards posterior distributions based on sample data. *Ann. Math. Statist.* 37 355–374.

[11] Dempster A.P. (2008) Upper and Lower Probabilities Induced by a Multivalued Mapping. In: Yager R.R., Liu L. (eds) Classic Works of the Dempster-Shafer Theory of Belief Functions. *Studies in Fuzziness and Soft Computing*, vol 219. Springer, Berlin, Heidelberg.

[12] D. Dubois, H. Prade (1988) *Possibility theory an approach to computerized processing of uncertainty*, Plenum Press, New York.

[13] Ferson, S., V. Kreinovich, L. Ginzburg, D.S. Myers, and K. Sentz. (2003). Constructing Probability Boxes and Dempster–Shafer Structures Archived 22 July 2011 at the Wayback Machine. SAND2002-4015. Sandia National Laboratories, Albuquerque, NM.

[14] Y. Ermoliev, A. Gaivoronski, Stochastic optimization problems with incomplete information on distribution functions, *SIAM Journal on Control and Optimization*, **23**, (1985), no. 5, 697–716.

[15] Y. Fang, T. Post, Higher-Degree Stochastic Dominance Optimality and Efficiency, *European Journal of Operational Research*, **261**, (2017), 984-993.

[16] C. Hess, Set-valued integration and set-valued probability theory: an overview, in *Handbook of Measure Theory, vols 1, II*, North-Holland, Amsterdam (2002).

[17] F. Hiai, Radon-Nikodým theorems for set-valued measures, *J. Multivariate Anal.* **8** (1978), no. 1, 96-118.

[18] D. Kandilakis, On the extension of multimeasures and integration with respect to a multimeasure, *Proceedings of the AMS*, **116**, (1992), no.1, 85-92.

[19] H.Kunze, D. La Torre, F.Mendivil, E.R.Vrscay, *Fractal-based methods in analysis*, Springer, New York, 2012.

[20] D. Kuroiwa, Existence theorems of set optimization with set-valued maps, *Journal of Information and Optimization sciences*, **24**, (2003), no. 1, 73-84.

[21] D. La Torre, F. Mendivil, Iterated function systems on multifunctions and inverse problems *Journal of Mathematical Analysis and Applications*, **340** (2007), no. 2, 1469–1479.

[22] D. La Torre, F. Mendivil, The Monge–Kantorovich metric on multimeasures and self-similar multimeasures, *Set-Valued and Variational Analysis* **23** (2015), no. 2, 319-331.

[23] D. La Torre, F. Mendivil, Portfolio optimization under partial uncertainty and incomplete information: a probability multimeasure-based approach, *Annals of Operations Research*, **267** (2018), 267–279.

[24] D. La Torre, F. Mendivil, Stochastic linear optimization under partial uncertainty and incomplete information using the notion of probability

multimeasure, *Journal of the Operational Research Society*, **69** (2018), no. 10, 1549–1556.

[25] La Torre, D., Mendivil, F. Stochastic efficiency and inefficiency in portfolio optimization with incomplete information: a set-valued probability approach, *Annals of Operations Research*, (2020), in press. https://doi.org/10.1007/s10479-020-03886-0

[26] H. Levy, Stochastic dominance, efficiency criteria, and efficient portfolios: The multi-period case, *American Economic Review*, **63**(5) (1973), 986–994.

[27] H. Levy, *Stochastic dominance: Investment decision making under uncertainty*, Springer, New York, 2006.

[28] O. Linton, T. Post, Y. Whang, Testing for the stochastic dominance efficiency of a given portfolio, *Econometrics Journal*, **17**, 2, (2014), S59–S74.

[29] M. Malavasi, S. Ortobelli Lozza, S. Trück, Second order of stochastic dominance efficiency vs mean variance efficiency, *European Journal of Operational Research*, **290** (2021), 1192-1206.

[30] H. Markowitz, Portfolio Selection, *The Journal of Finance* **7** (1952), 77-91.

[31] F. Mendivil, Computing the Monge-Kantorovich distance, *Comp. Appl. Math.*, **36**(3) (2017), 1989-1402.

[32] T.O. Meyer, X.M. Li, L.C. Rose, Comparing mean variance tests with stochastic dominance tests assessing international of portfolio diversification benefits, *Financial Services Review*, **14** (2005), 149–168.

[33] R.B. Porter, J.E. Gaumnitz, Stochastic dominance vs. mean-variance portfolio analysis: An empirical evaluation, *American Economic Review*, **62** (3) (1972), 438–446.

[34] T. Post, Empirical Tests for Stochastic Dominance Efficiency, *Journal of Finance*, **58** (2003), 1905–1932.

[35] T. Post, V. Poti, Portfolio Analysis using Stochastic Dominance, Relative Entropy and Empirical Likelihood, *Management Science*, **63** (2017), 1, 153-165.

[36] T. Post, M. Kopa, Portfolio Choice based on Third-order Stochastic Dominance, *Management Science*, **63**, 3381-3392.

[37] R.T. Rockafellar, R. J-B. Wets, *Variational analysis*, Springer, New York, 1998.

[38] A. Shapiro, D. Dentcheva, and A. Ruszczynski. *Lectures on stochastic programming: modeling and theory*, SIAM, 2014.

[39] M. Stojaković, Set valued probability and its connection with set valued measure, *Statist. Probab. Lett.* **82** (2012), 1043-1048.

[40] B. Urli, R. Nadeau, Stochastic MOLP with incomplete information: An interactive approach with recourse, *Journal Operational Research Society*, **41**, (1990), no. 12, 1143–1152.

[41] B. Urli, R. Nadeau, PROMISE/scenarios: An interactive method for multiobjective stochastic linear programming under partial uncertainty, *European Journal Operational Research*, **155**, (2004), 361–372.

[42] R.G. Vickson, Stochastic Dominance Tests for Decreasing Absolute Risk Aversion. I. Discrete Random Variables, *Management Science*, 21 (12), (1975), 1438–1446.

[43] R.G. Vickson, Stochastic Dominance Tests for Decreasing Absolute Risk Aversion. II. General random Variables, *Management Science*. 23 (5), 1977, 478–489.

[44] K. Vind (1964), Edgeworth-allocations in an exchange economy with many traders, *Int. Econ. Rev.* 5(2), 165–177.

[45] P. Walley (1991), *Statistical reasoning with imprecise probabilities*, Chapman and Hall, London, New York.

[46] Y. Wang (2008), Imprecise probabilities with a generalized interval form. In R.L. Muhanna and R.L. Mullen, eds., Proc. 3rd Int. Workshop on Reliability Engineering Computing (REC'08), Savannah, Georgia, 45-59.

[47] K. Weichselberger (2000), The theory of interval-probability as a unifying concept for uncertainty, International Journal of Approximate Reasoning, 24 (2–3), pp. 149-170.

[48] Q. Zhang, Q. Xie, G. Wang (2016), A survey on rough set theory and its applications, CAAI Transactions on Intelligence Technology, 1 (4), 323-333.