A lattice approach to the Beta distribution induced by stochastic dominance: Theory and applications

Yann Braouezéc and John Cagnolb,c,1

bIESEG School of Management, Univ. Lille, CNRS, UMR 9221 - LEM - Lille Economie Management, F-59000 Lille, France; cUniversité Paris-Saclay, CentraleSupélec, MICS, Gif-sur-Yvette, France; 1Fédération of Mathematics CNRS FR3487, Paris, France

ABSTRACT
We provide a comprehensive analysis of the two-parameter Beta distributions seen from the perspective of second-order stochastic dominance. By changing its parameters through a bijective mapping, we work with a bounded subset instead of an unbounded plane. We show that a mean-preserving spread is equivalent to an increase of the variance, which means that higher moments are irrelevant to compare the riskiness of Beta distributions. We then derive the lattice structure induced by second-order stochastic dominance, which is feasible thanks to the topological closure of . Finally, we consider a standard (expected-utility based) portfolio optimization problem in which its inputs are the parameters of the Beta distribution. We explicitly characterize the subset of for which the optimal solution consists of investing 100% of the wealth in the risky asset and we provide an exhaustive numerical analysis of this optimal solution through (color-coded) graph.

1. Introduction

From the seminal and fundamental article of Rothschild and Stiglitz (see Rothschild & Stiglitz, 1970), we know that the statement the random variable Y is riskier than a random variable X can be defined in three equivalent ways: (1) every risk-averse (expected-utility) decision prefers X to Y, (2) Y is equal to X plus a noise term (i.e., with zero mean) uncorrelated to X and finally (3) Y has more weight in the tails than X. In the particular case in which the mean of X and Y are equal, Y is a mean-preserving spread of X and the distribution function of X is said to second-order stochastically dominate the distribution function of Y.

While the variance plays an important role in finance (it is the square of the so-called volatility of the log-returns of a stock or an index), the fundamental contribution of Rothschild and Stiglitz (1970) is precisely to show that the variance as a measure of riskiness may not be equivalent to the three aforementioned definitions. However, as observed in his review paper on the subject (see Levy, 1992), in some cases, the variance can be safely used as a measure of riskiness (i.e., it is thus consistent with the above definitions) and the best well-known example is the case of Gaussian (or normal) random variables. This actually holds true because the Gaussian distribution function belongs to a location-scale family.

The normal distribution plays a central role in probability theory and related fields because of its striking properties such as stability, location-scale family, absence of fat tails, existence of all moments, etc. It is, however, particular because it is symmetric around its mode/mean (i.e., zero skewness) and its excess kurtosis, a disputable quantity frequently used in finance, is also equal to zero. In finance, many popular models, for example, the capital asset pricing model (CAPM) or the Black-Scholes model, actually rely on the Gaussian distribution to model the rate of return of a stock (possibly an index) while it is well-known that the distribution of the observed log-returns are not symmetric (i.e., it is skewed) and may exhibit a positive excess kurtosis (see the well-known review Cont (2001)). From a decision theory point of view, this absence of asymmetry of the normal distribution fails to capture a possible preference for a skewed distribution, generally measured by the third standardized central moment called skewness (see Arnold and Groeneveld (1995) for different measures of skewness, and see also Groeneveld and Meeden (1984)).

The two-parameter Beta distribution thus appears as an interesting candidate to consider due to the various shapes it can take (e.g., Moitra, 1990) since it can be single-peaked (i.e.,-shaped, positively or negatively skewed, but it can also be U-shaped, J-shaped (increasing), or even decreasing.
1.1. Aim of the paper

It is to provide a comprehensive analysis of the two-parameter Beta distribution in which we explicitly adopt a lattice approach induced by second-order stochastic dominance. It should be pointed out that there is already a large body of literature on the Beta distribution (see for instance, Moitra (1990), Gupta and Nadarajah (2004), a handbook solely devoted to the Beta distribution) but no lattice analysis has been yet offered. In their well-known paper, Ebrahimi et al. (1999) consider an interesting related problem. They analyse the ordering of various special distributions (e.g., Beta, Gamma, inverse Gamma, Pareto Weibull, etc.) according to the variance and entropy, and they note that the Beta distribution is the most complicated case. By definition, the set of parameters of the Beta distribution, denoted frequently \( \alpha \) and \( \beta \), is unbounded since both \( \alpha \) and \( \beta \) are positive (see e.g., Pitman, 2009).

As a result, Ebrahimi et al. (1999) represents their finding (i.e., the ordering) in a subset of parameters such as \([0, 4] \times [0, 4]\) and not in the overall \(xy\)-plane (i.e., \( \mathbb{R}^2_+ \)). With stochastic dominance in view, the usual definition of the Beta distribution thus has three major drawbacks.

1. The set of parameters, that is, the \( xy \)-plane (i.e., \( \mathbb{R}^2_+ \)) is unbounded.
2. The two parameters \( x > 0 \) and \( \beta > 0 \) have no natural economic interpretation.
3. The limiting distributions (i.e., Dirac masses and convex combination of Dirac masses) are excluded from the analysis in \( xy \)-plane.

Throughout the paper, instead of working with \( \mathbb{R}^2_+ \), as the natural set of parameters, we shall consider a bounded subset of \( \mathbb{R}^2_+ \) and the new parameters will be the mean \( m \) of the Beta random variable and the variance \( v \). As a result, with a mean-preserving spread in view, it becomes quite easy to leave the mean constant. Thanks to the equinumerosity property of \( \mathbb{R} \) with any open subset of \( \mathbb{R} \), one can design a bijective mapping (possibly differentiable) between \( \mathbb{R}^2_+ \) and an open bounded subset of \( \mathbb{R}^2_+ \). While this bounded subset can possibly be a square or a disk, we find it convenient from a topological point of view to choose a subset delimited by the \( x \)-axis, whose upper boundary is a parabola of the form \( y = m - m^2 \). For this reason, we call this new set of parameters the \( \mathbb{MV} \)-dome (denoted \( \mathcal{D} \)). For a given mean \( m \in (0, 1) \), as long as the variance \( v < m - m^2 \), this couple of parameters \((m, v)\) lies in the \( \mathbb{MV} \)-dome and corresponds, by design, to a unique point in the \( xy \)-plane. As is well-known in the \( xy \)-plane (see e.g., Pitman, 2009), when both \( \alpha \) and \( \beta \) are higher (lower) than one, the Beta density is respectively \( \cap \)-shaped or single-peaked (\( U \)-shaped). An interesting aspect of the \( \mathbb{MV} \)-dome is that the limiting \( \cap \)-shaped Beta distributions, as well as the limiting \( U \)-shaped, are located on the boundary of the \( \mathbb{MV} \)-dome. The particular feature of these limiting Beta distributions is that they do not admit a density since they are Dirac masses (or convex combination of Dirac masses).

1.2. Contribution of the paper

It is a major contribution of this paper to show that, by considering the topological closure of the \( \mathbb{MV} \)-dome (its boundary), we are able to derive the non-trivial lattice structure of the two-parameter Beta distribution induced by second-order stochastic dominance. The topological closure of the \( \mathbb{MV} \)-dome thus does not appear as pure mathematical construction designed to simply extend the set of Beta distributions; it is at the heart of the construction of the lattice structure. This, in turn, allows us to construct the Hasse diagram, that is, the path from the minimum element in general denoted \( \perp \) to the maximum element in general denoted \( \top \). In particular, we show that the set of Beta distribution functions with a constant mean are ordered with respect to their variance. More precisely, if \( F \) and \( G \) are the distribution functions of two Beta random variables \( X_F \) and \( X_G \), respectively with the same mean but different variance, that \( F \) second-order stochastically dominates \( G \), something denoted as usual as \( G \leq_{ssd} F \), is equivalent to \( V_F \leq V_G \) where \( V_F \) and \( V_G \) are the variance of \( X_F \) and \( X_G \), respectively. Interestingly, in the \( \mathbb{MV} \)-dome, the set of Beta distribution functions with a constant mean is just a vertical segment.

While the ordering of Beta distribution functions along a vertical segment of the \( \mathbb{MV} \)-dome is an interesting result, it is actually not enough to derive the particular lattice structure. We also need to derive the order of the Beta distribution functions for which the parameters are located on the boundary of the \( \mathbb{MV} \)-dome. For this reason, we show that the Beta distribution functions with zero variance (the parameters are located on the \( x \)-axis of the \( \mathbb{MV} \)-dome) but also those with maximal variance (the parameters are located on the parabola of the \( \mathbb{MV} \)-dome), which are ordered with respect to second-order stochastic dominance. Taken now together, these orders allow us to derive the particular lattice structure (of the Beta distribution whose parameters are in \( \mathbb{MV} \)-dome) and the Hasse diagram.

At this stage, it is fairly natural to inquire whether or not the distribution functions, the parameters of which are located in the \( \mathbb{MV} \)-dome,
can be totally ordered with respect to stochastic dominance. For instance, if one considers two distribution functions with a different mean, but with the same variance \( \nu \), are these two distribution functions ordered according to stochastic dominance? For the Gaussian case, the answer is positive and the distributions are even comparable according to first-order stochastic dominance. For the Beta distribution case, the answer may be negative in that the distributions may be non-comparable according to second-order stochastic dominance. As we shall see, the very reason for this somewhat surprising result is related to the “change of regime” of the Beta distribution functions when the parameters lie on the boundary of the dome; these distributions do not admit a density anymore and reduce to Bernoulli random variables for which the distribution functions is flat.

1.3. Economic foundations and related literature in operational research

In their groundbreaking paper, Diamond and Stiglitz (1974) introduce the notion of a mean preserving spread (an example of second order stochastic dominance) with respect to a family of distribution functions that depends upon a single parameter and offer many economic applications. More generally, the notion of stochastic dominance is widely analyzed in Economics (e.g., Levy (1992) for an early survey and Levy (2015) and Sriboonchita et al. (2009) for an elementary and more advanced textbook respectively) but also in the Operational Research community, e.g., Bunn (1979), Denuit et al. (2013), Kopa et al. (2018), Bruni et al. (2017), Egozcue and Wong (2010), Post and Kopa (2013), Batur and Choobineh (2012), Fang and Post (2017), Liesiö et al. (2020, and Liesiö et al. (2020)) to quote old and recent papers. However, to the best of our knowledge, Ali (1975) seems to be the only paper in which the Beta distribution is explicitly considered with dominance stochastic in view and applications to portfolio choices, as we do in the present paper, although no lattice structure is offered. Interestingly, since portfolio choice problems are at the interface of many disciplines such as probability theory, optimization and economics, it is not surprising that these problems constitute an active area in operational research (e.g., Bernard et al., 2019; Bilbao et al., 2006; Cesarone & Colucci, 2017; Tassak et al., 2017) to quote few operational research papers.

1.4. Organization of the paper

We introduce in Section 2 the notations used throughout the paper and we prove a simple, yet essential result. In Section 3, the essence of the paper, by changing the set of parameters through a bijective mapping, we derive the lattice structure of the two-parameter Beta distribution with respect to second-order stochastic dominance. Finally, in Section 4, we apply these results to a portfolio optimization problem in which a decision-maker must allocate their wealth between a default risk-free asset and a risk one (a Beta random variable). Thanks to the known lattice structure of the Beta distribution function, we are able to fully characterize the region of the Dome for which the decision-maker decides to invest 100% of the wealth in the risky asset. Moreover, since the set of parameters is bounded, we are able to provide an exhaustive numerical analysis.

2. Notations, definitions, and preliminary results

2.1. Basic definitions and result

Throughout the paper, unless stated otherwise, we consider the case of positive random variables, that is, those for which the support (of the underlying probability measure) is the compact subset of \( \mathbb{R}_+ \), such as \([0, 1]\). Let \( X \) be such a random variable and let \( F_X := F \) be its distribution function. The expectation \( \mathbb{E}(X) \) lies in \([0, 1]\), thus it is finite, and can be written as (see e.g., Pitman, 2009, p. 332)

\[
\mathbb{E}(X) = \int_0^1 (1 - F(x)) \, dx = \int_0^1 S(x) \, dx \quad (1)
\]

where \( S = 1 - F \) is called the survival function and note that Equation (1) holds more generally for any positive random variable with finite expectation.

Consider now two positive random variables \( X_1 \) and \( X_2 \) with distribution function \( F_1 \) and \( F_2 \), respectively (both supported by \([0, 1]\)) with \( F_1 \neq F_2 \) and such that \( \mathbb{E}(X_2) = \mathbb{E}(X_1) \). From Equation (1), we thus obtain that

\[
\mathbb{E}(X_2) = \mathbb{E}(X_1) \iff \int_0^1 [F_1(x) - F_2(x)] \, dx = 0. \quad (2)
\]

Following the terminology introduced in the influential paper Diamond and Stiglitz (1974), the distribution function \( F_1 \) is said the be riskier than the distribution function \( F_2 \) if it has the same mean but more weights in both tails. Formally (see Cohen (1995), or the textbook Sriboonchita et al. (2009)) \( F_1 \) is said to be a mean preserving spread of \( F_2 \) if and only if both statements hold:

\[
\mathbb{E}(X_1) = \mathbb{E}(X_2) \quad (3a),
\]

\[
\forall x \in [0, 1], \int_0^x F_2(t) \, dt \leq \int_0^x F_1(t) \, dt. \quad (3b)
\]
When Equation (3b) holds, as usual in Economics, it is said that $F_2$ dominates $F_1$ according to second order stochastic dominance (SSD), which we note as $F_1 \leq_{SSD} F_2$.

Note that SSD implies $\mathbb{E}(X_1) \leq \mathbb{E}(X_2)$ a mean preserving spread thus is the particular case of second-order stochastic dominance in which the means are identical.

Let $\mathbb{E}(u(X))$ be the expected utility associated to the random variable $X$ for some Von-Neumann Morgenstern utility function $u$ and let $U_{\mu}$ be the set of increasing concave functions. The following equivalence is a well-known result in the Economics of risk (see for instance Levy (1992, p. 557) or Sripoonchita et al. (2009, p. 81)).

$$F_1 \leq_{SSD} F_2 \iff \forall u \in U_{\mu}, \mathbb{E}(u(X_1)) \leq \mathbb{E}(u(X_2)).$$

In other words, as long as the distribution function $F_2$ second-order stochastically dominates $F_1$, all risk-averse expected utility decision-makers (weakly) prefer $F_2$ to $F_1$.

Let us now recall first-order stochastic dominance (FSD). It is said that $F_2$ first-order stochastically dominates $F_1$ (FSD) if, for each $x \in [0,1]$, $F_2(x) \leq F_1(x)$, with $F_2 \neq F_1$, something that we note $F_1 \leq_{FSD} F_2$. If $F_1 \leq_{FSD} F_2$, then Equation (3b) is satisfied but the converse is obviously not true. As is well-known, first order stochastic dominance (FSD) is a stronger than second-order stochastic dominance (SSD). Note also that if $F_1 \leq_{FSD} F_2$, then, $\mathbb{E}(X_1) < \mathbb{E}(X_2)$. By definition, when the mean of $X_1$ and $X_2$ are equal and when $F_1$ and $F_2$ have at least one crossing point, they cannot be FSD-ranked. The analysis of a mean preserving spread precisely consists in analysing situations in which distribution functions are not FSD-ranked but might be SSD ranked. When the distribution functions cross only once, following once again Diamond and Stiglitz (1974), the mean preserving spread is said to be simple (see also Chateauneuf et al., 2004).

With the Beta distribution in view, let $C^1([0,1])$ be the set of continuous non-decreasing distribution functions $F$ such that $F(0) = 0$ and $F(1) = 1$.

**Definition 1.** For $\mu \in (0,1)$, let $\mathcal{F}_\mu$ be the set of functions with the two following properties.

1. For each $F \in C^1([0,1])$, $\int_0^1 (1 - F(x)) dx = \mu$
2. For each $F, G \in C^1([0,1])$ with $F \neq G$, there exists a unique non trivial crossing point $x_c \in (0,1)$ such that $F(x_c) = G(x_c)$.

The set of distribution functions $\mathcal{F}_\mu$ is such that, by definition, the mean of each element $F$ is equal to $\mu$ and two different distribution functions have a unique non-trivial crossing point, that is, a crossing point which is not equal to a bound of the support, (i.e., zero or one). From the first property, two functions $F$ and $G$ with the same mean $\mu$ must cross at least once. From the second property, this (non-trivial) crossing point is unique.

**Proposition 1.** The set of distribution functions $\mathcal{F}_\mu$ is completely ordered with respect to second order stochastic dominance.

**Proof.** See the appendix.

Consider now, the three following elements of $\mathcal{F}_\mu$ denoted $F_1, F_2, F_3$. By definition, they have the same mean equal to $\mu$. Let $V_i$ denote the variance of the underlying random variable $X_i$ with distribution function $F_i$ that belongs to $\mathcal{F}_\mu$. From (Hadar & Russell, 1971, Theorem 3) and Tesfatsion (1976), the following result is true.

**Corollary 1.** If $F_1, F_2, F_3$ belong to $\mathcal{F}_\mu$ and are such that $F_3 \leq_{SSD} F_2 \leq_{SSD} F_1$, then, $V_3 \geq V_2 \geq V_1$.

Proposition 1 is interesting in that it says that when distribution functions cross only once in $(0,1)$, then, the distribution functions are ordered with respect to second order stochastic dominance. Assuming however that the distribution functions have a unique non-trivial crossing point (i.e., in $(0,1)$) is a fairly strong assumption. As we shall now see, when the distribution function belongs to a location-scale family, then, the non trivial crossing point is unique.

**2.2. Location scale family and second order stochastic dominance**

Let $X$ be a random variable with finite variance distributed according to a density that depends upon two parameters $\mu \in \mathbb{R}$ and $\sigma > 0$. Let $f(x, \mu, \sigma)$ and $F(x, \mu, \sigma)$ be the density and the distribution function of $X$. The two parameters $\mu$ and $\sigma$ are said to be location-scale parameters for the distribution function of $X$ if for all $x$, it satisfies

$$F_X(x, \mu, \sigma) = G \left( \frac{x - \mu}{\sigma} \right),$$

$$f_X(x, \mu, \sigma) = \frac{1}{\sigma} g \left( \frac{x - \mu}{\sigma} \right)$$

(4)

(5)

for some distribution function and density $G$ and $g$, respectively sometimes called the reduced or standard distribution function and density (see Balakrishnan and Nevzorov (2004), see also Ebrahimi et al. (1999)). It is important to point out that the (reduced) density $g$ and the distribution function $G$ depend upon $x$ and the two parameters $\mu$ and $\sigma$ in a specific way, that is $\sum_{m=0}^{\infty} \frac{(-\mu/\sigma)^m}{m!}$. When the distribution function of $X$ belongs to a location-scale family, it suffices to write for each $x$ that
$f_X(x, 0, 1) = g(x)$, hence the name of reduced density for $g$. Assume that $X \sim f_X(x, \mu, \sigma)$, that is, the density of $X$ is $f_X(x, \mu, \sigma)$. When this density belongs to a location-scale family, the reduced random variable $\frac{X-\mu}{\sigma} \sim f_X(x, 0, 1)$ or, equivalently, if $X \sim f_X(x, 0, 1)$, then the random variable $\mu + \sigma X \sim f_X(x, \mu, \sigma)$. The best well-known example of such a location-scale family is the Gaussian (or normal) density, see Ebrahimi et al. (1999, table 1) or Rinne (2011) for a longer list of location-scale distributions.

To see the implication of location-scale family in terms of stochastic dominance, assume that $X \sim g_X(x, 0, 1)$ and let $Y_1 = \sigma_1 X + \mu_1$ and $Y_2 = \sigma_2 X + \mu_2$. By definition, $F_{Y_1}$ and $F_{Y_2}$ belong to the same location-scale family and their distribution function are respectively equal to $F_{Y_1}(y) = G\left(\frac{y-\mu_1}{\sigma_1}\right)$ and $F_{Y_2}(y) = G\left(\frac{y-\mu_2}{\sigma_2}\right)$ where $G$ is here the distribution function of $X$. Since

$$G\left(\frac{y-\mu_1}{\sigma_1}\right) = G\left(\frac{y-\mu_2}{\sigma_2}\right) \iff y_c = \frac{\mu_1\sigma_2 - \mu_2\sigma_1}{\sigma_1 - \sigma_2} \quad (6)$$

it thus follows that $y_c$ is the unique crossing point of $F_{Y_1}$ and $F_{Y_2}$, which means that one distribution must second-order stochastically dominates the other one. The following result is then not difficult to prove.

If $\mu_1 = \mu_2$, then, the SSD order is equivalent to the variance order, that is,

$$Y_2 \leq_{SSD} Y_1 \iff \sigma_1 < \sigma_2. \quad (7)$$

When the distribution functions belong to a location scale family, the variance order is equivalent to a mean preserving spread (second order stochastic dominance with constant mean). Unfortunately, when the distribution functions do not belong to a location scale family, the variance order needs not be equivalent to the second order stochastic dominance. From this point of view, the Beta distribution is challenging because it does not belong, in general, to a location scale family. To prove the equivalence between the SSD order and the variance order, one may rely on Proposition 1 but it must be shown before that the distribution function of two Beta random variables with the same mean have a unique (non trivial) crossing point. Such a result will be proved in the sequel.

3. A lattice approach to the set of Beta distributions (5)

3.1. The Beta distribution: Definitions, properties, and applications

3.1.1. Definition and properties

Let X be a random variable distributed according to a two-parameter Beta distribution. It is said to be a Beta random variable if the density function is given by

$$f_X(x, \mu, \sigma) = \beta(x / \sigma)^{x-1}(1-x)^{\beta-1} / B(\alpha, \beta)$$

where $\alpha$ and $\beta$ are positive numbers and $B(\alpha, \beta)$ is a normalization constant:

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx. \quad (8)$$

The distribution function of the Beta distribution frequently $I(x, \alpha, \beta)$ has in general no closed formula and is classified as a special function in mathematics (see e.g., Temme, 2011). There are, however, various types of representations (i.e., integral representation, series expansions) but also asymptotic expansions (for $\alpha$ or $\beta$ large) and recurrence relations. A comprehensive review with statistical application in view is also provided in (Gupta & Nadarajah, 2004, Chapters 1 and 2). Throughout the paper, one may use the terms Beta density (or distribution) or Beta probability measure interchangeably. The Beta distribution flexible in that it can be $\cap$-shaped (single-peaked or Arched), increasing, decreasing or $U$-shaped as a function of the parameters (see Gupta and Nadarajah (2004, Chapter 2), see also Pitman (2009, p. 329) for an elementary textbook). For a paper on the skewness on the Beta distribution, see for instance Moitra (1990). Regarding now the moments, it is be shown that the expectation and the variance of $X$ are respectively equal to (see Gupta & Nadarajah, 2004, pp. 35-36).

$$E(X) = \frac{\alpha}{\alpha + \beta}, \quad V(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}. \quad (9)$$

In general, the two-parameter Beta distribution is not a location-scale family. To see this, let $X = \mu + \sigma Z$ where $Z$ is a two-parameter Beta distribution with parameter $\alpha$ and $\beta$. It is easy to show that the density of $X$ is equal to $f_X(x, \mu, \sigma) = \frac{1}{\sigma} f_Z\left(\frac{x-\mu}{\sigma}\right)$ so that

$$f_X(x, \mu, \sigma) = \frac{(x-\mu)^{x-1}(\sigma+\mu-x)^{\beta-1}}{\sigma^{\alpha+\beta-1}B(\alpha, \beta)}. \quad (10)$$

It is easy to see that Equation (12) fails to satisfy Equation (5) since the parameters $\alpha$ and $\beta$ appear in the density and thus do influence both the shape of the density and the distribution function. In such a situation, Equation (6) is not anymore valid. It turns out that for some values of the parameters $\alpha$ and $\beta$, the Beta distribution may be a location-scale family. A simple example of location-scale family is when $\alpha = \beta = 1$, that is, when the Beta distribution reduces to an uniform distribution. Less trivial is
the so-called Arcsine distribution. We refer to Rinne (2011) for other examples of particular cases of the Beta distribution that are location-scale family. From the discussion, since the two-parameter Beta distribution is in general not a location-scale family, if one wants to make use of Proposition 1, it remains to prove that when the mean is constant, the Beta distribution functions have a unique (non trivial) crossing point. To prove such a statement, we shall perform a change of variable.

3.1.2. Why is the Beta distribution interesting?
Because it is highly flexible, that is, as opposed to many probability distributions, it can take various shapes. Depending upon the parameters, it can be \( \cap \)-shaped when \( \alpha > 1 \) and \( \beta > 1 \) or \( U \)-shaped when \( \alpha < 1 \) and \( \beta < 1 \). But it can also be decreasing when \( \alpha < 1 \) and \( \beta \geq 1 \) or increasing when \( \alpha \geq 1 \) and \( \beta > 1 \).

The Gaussian distribution, which is symmetric around its mean (but also its mode or median), is often used in finance to model the rate of return of a stock while its well-known that for individual stocks (but not necessarily for portfolios, see Sun and Yan (2003) for more on this), the observed distribution is positively skewed and such a stylized fact has been called the “persistence of skewness” (Albuquerque, 2012; Beedles & Simkowitz, 1980; Chang et al., 2013; DeFusco et al., 1996; Lau et al., 1989; Singleton & Wingender, 1986). Unfortunately, since the Gaussian distribution has no skewness, it cannot reproduce the above stylized fact. Such a positive skewness can however be reproduced with a Beta distribution due to its flexibility (Moitra, 1990). In finance, the Beta distribution is also used to model the recovery rate (i.e., a percentage) of a credit product such as a bond or a Credit Default Swap. It is usual to consider this recovery rate as given, e.g., 20% or 40% while it is indeed stochastic. In Gambetti et al. (2019), relying on the so-called beta regression models, well-suited for its heteroskedastic and skewed distributions, the authors analyze the determinants of recovery rate distributions. In project management (PERT), an active area of operational research, the Beta distribution is once again used for its flexibility, i.e., it can be bell-shaped and yet positively skewed (see Johnson (1998), see also Farnum and Stanton (1987) and Golenko-Ginzburg (1988)).

3.2. The MV-dome and the topological closure of \( B \)
The two-parameter Beta distribution comes with a number of caveats in the context of investigating mean preserving-spreads.

(1) The parameters \( \alpha \) and \( \beta \) do not have any direct economic meaning.
(2) The iso-mean and iso-variance curves are fairly complex in the unbounded \((\alpha, \beta)\)-quadrant.
(3) The two-parameter Beta distribution is, in general, not a location-scale family.
(4) The support of any random variable \( Y = aX + b \) where \( X \) is a two-parameter Beta distribution is equal to \([b, a + b] \) and is never equal to \([0, 1] \) unless \( a = 1 \) and \( b = 0 \).

In what follows, in order to be able to compare Beta distributions supported by \([0, 1] \) with the same mean and variance \( V \) and no longer by \( \alpha \) and \( \beta \). Instead of representing the parameters of Beta distributions by \((\alpha, \beta) \in \mathbb{R}^2_{++} \) we will represent them by \((M, V) \in \mathcal{D} \). It is important to point out at this stage that the MV-dome is not only a transformation function.

As a result this function is a bijection. We shall note this function \( \phi^{-1} \) so that

\[
\begin{align*}
\alpha &= \frac{M(M-M^2-V)}{V}, \quad \text{(14a)} \\
\beta &= \frac{(1-M)(M-M^2-V)}{V}. \quad \text{(14b)}
\end{align*}
\]
Let $X$ be a random variable whose repartition function is $F_{M,V}$. The following convergence in distribution of $X$ toward Dirac masses holds true.

(1) For all $M \in (0,1)$, $\lim_{V \to 0} X = \delta_M$.

(2) $\lim_{(M,V) \to (0,0)} X = \delta_0$ and $\lim_{(M,V) \to (1,0)} X = \delta_1$.

(3) For all $M \in (0,1)$, $\lim_{V \to D(M)} X = (1-M)\delta_0 + M\delta_1$.

**Proof.** See Appendix A.

When $V$ approaches the boundary of the dome, i.e., the parabola defined by $V = D(M)$, the third point of the above result shows that the beta random variable tends to a convex combination of the Dirac Delta functions (also called Dirac masses) $\delta_0$ and $\delta_1$ with weights $1-M$ and $M$. Put it differently, the above lemma says that on the boundary of the dome, the Beta random variable is no longer a continuous random variable but is rather a discrete random variable. This result is actually not so surprising. When $(M,V)$ lies in a $U$-shaped region and when $V$ tends to $D(M)$, the limiting distribution assigns no weight to the open interval $(0,1)$. It thus becomes a *Bernoulli random variable* that takes two values, zero and one, with probability $1-M$ and $M$, respectively so that its expected value (mean) is equal to $M$. Since Dirac Delta functions should not be excluded of the analysis, Lemma 1 suggests extending by continuity $\phi^{-1}$ on the domain defined as

$$\bar{\mathcal{D}} = \{(M,V) \in [0,1] \times [0,1] \mid V \leq M-M^2\}$$

that is, the dome now includes its boundary, the parabola defined as $V = D(M)$ but also the bottom segment $[0,1] \times \{0\}$. The distribution function of the random variable $X$ is

\[
\phi(M,V) = (\alpha, \beta) = \left(\frac{M(M-M^2-V)}{V}, \frac{(1-M)(M-M^2-V)}{V}\right)
\]

mapping $\mathcal{D}$ to $\mathbb{R}^2_+$. Note the positivity of $M-M^2-V$ is a necessary condition for the positivity of $\alpha$. Since both $\phi$ and $\phi^{-1}$ are continuous, it thus defines a *homeomorphism* (bijectivity, continuity and continuity of the inverse mapping). An interesting property of homeomorphisms is that they "propagate" the topological properties from one space to the other.

Using this bijection, the Beta distributions will be parameterized by $(M,V) \in \mathcal{D}$ instead of $(\alpha, \beta) \in \mathbb{R}^2_+$. From now on, we shall note $F_{M,V}$ the distribution function associated to the parameters $(M, V)$ and $f_{M,V}$ the corresponding density. Thus, $f_{M,V}$ is given by (8) where $(M,V) = \phi^{-1}(\alpha, \beta)$.

We shall denote by $\mathcal{B}$ the set of distribution functions $F_{M,V}$ with $(M,V) \in \mathcal{D}$. With mean preserving spread in mind, this representation is particularly meaningful since it is now possible to immediately identify a Beta distribution with higher mean (going right) and with higher variance (going up). The four main categories of Beta distributions Arched (A), Increasing (I), Decreasing (D) and U-shaped (U) are shown on Figure 1. Lemma 4 in the appendix provide the details about the functions plotted on Figure 1. It is important to note that, by definition, the boundary of the dome is not considered at this stage.

**Lemma 1.** Let $X$ be a random variable whose repartition function is $F_{M,V}$. The following

\[
\frac{1}{4}
\]

\[
\begin{array}{c}
\text{U} \\
0 \\
\text{D} \\
0 \\
\text{A} \\
0 \\
\text{I} \\
0 \\
M \\
1
\end{array}
\]

Figure 1. The $MV$-dome $\mathcal{D}$. 

(1) $\lim_{V \to 0} X = \delta_M$.

(2) $\lim_{(M,V) \to (0,0)} X = \delta_0$ and $\lim_{(M,V) \to (1,0)} X = \delta_1$.

(3) For all $M \in (0,1)$, $\lim_{V \to D(M)} X = (1-M)\delta_0 + M\delta_1$. 

**Proof.** See Appendix A.
\[
F_{M,0}(x) = \begin{cases} 
0 & \text{if } x < M \\
1 & \text{if } x \geq M
\end{cases}
\] (15)

Note that the Dirac mass \( \delta_M \) and the distribution function \( F_{M,0} \) are equivalent in the sense that a random variable distributed according to a distribution function \( F_{M,0} \) is a Dirac mass on the constant \( M \).

In the same vein, the distribution function of the random variable \( X \) is
\[
F_{M,D}(x) = \begin{cases} 
1-M & \text{if } 0 < x < 1 \\
1 & \text{if } x = 1
\end{cases}
\] (16)

Once again, the weighted sum of Dirac masses \((1-M)\delta_0 + M\delta_1\) and the distribution function \( F_{M,D(M)} \) are equivalent. The topological closure of \( B \) thus is
\[
\bar{B} = B \cup \{F_{M,0}, M \in (0,1)\} \cup \{F_{M,D(M)}, M \in (0,1)\}
\]

As we shall see, the topological closure will be useful in deriving the lattice structure of the set \( \bar{B} \).

### 3.3. Properties of the distribution functions \( F_{M,V} \) and second-order stochastic dominance

The goal of this section is in part to show for any two-parameter Beta distribution with the same mean \( M \), the distribution functions cross only once.

**Proposition 2.** Let \( M \) be in \((0, 1)\). Let \( V_1, V_2 \) be in \((0, D(M))\) with \( V_1 \neq V_2 \).

1. The equation
\[
F_{M,V_1}(x) = F_{M,V_2}(x)
\] (17)
has one solution and one solution only in the open interval \((0, 1)\).

2. Assume that \( V_1 < V_2 \) and let \( x_c(M, V_1, V_2) \) be the unique solution of Equation (17). Then
\[
\forall x \in (0, x_c), F_{M,V_1}(x) < F_{M,V_2}(x),
\]
\[
\forall x \in (x_c, 1), F_{M,V_2}(x) < F_{M,V_1}(x).
\]

**Proof.** See Appendix A.

Figure 2 illustrates Proposition 2 by showing the 11 cumulative distribution functions \( F_{M,V} \) for \( M = 0.3 \) where \( V \) varies. They are represented with a shade of red going from lighter to darker as \( V \) goes from 0% to 100% of the maximum possible variance for this mean, that is \( V(M) \).

Figure 2. Cumulative distribution function \( F_{M,V} \) for \( M = 0.3 \) and several \( V \).

Let \( \bar{B}_M \) be the subset of \( \bar{B} \) for a given \( M \in (0,1) \) (thus \( V \) varies in \([0, D(M)]\)). We shall prove that all distributions in \( \bar{B}_M \), corresponding to a vertical segment of the \( MV \)-dome, can be totally ordered according to second order stochastic dominance. For any \( V \in [0, D(M)] \), define
\[
Y_{M,V} : x \mapsto \int_0^x F_{M,V}(t)dt.
\]

Lemmas 5 and 6 in Appendix A show that \( Y_{M,V} \) is increasing and convex. This implies that, given \( V_1 \) and \( V_2 \) in \((0, D(M))\) with \( V_1 < V_2 \), we have
stochastic dominance is equivalent to a variance according to second-order stochastic dominance that the skewness and kurtosis might be relevant. 

Pare two Beta distribution functions, which means that the maximum element by

and denote the minimum element by

The maximum element of \( B_M \) is the distribution function with mean \( M \) and zero variance, i.e., \( F_{M,0} \) while the minimum element is the distribution function with the maximum variance given the mean \( M \), equal to \( V = D(M) \) (this means that \((M, V)\) is located on the upper parabola, i.e., \( F_{M,D(M)} \)). Intermediate elements are distribution functions \( F_{M,V} \), with \( V \in (0, D(M)) \). To summarize, in \( B_M \), the Beta distribution functions are completely ordered with respect to second order stochastic dominance, that is, for any \( 0 < V_1 < V_2 < D(M) \), we have

To conclude, it is interesting to note that Corollary 2 has an important implication in terms of "moments comparisons". As long as \( F_{M,V_y} \) and \( F_{M,V_y} \), with \( M \in (0,1) \), second-order stochastic dominance is equivalent to a variance order, that is, \( F_{M,V_y} \), equivalent \( V_1 \leq V_2 \). When the mean \( M \) of the Beta random variable is constant, a risk-averse expected utility decision-maker only takes the variance into account to compare two Beta distribution functions, which means that the skewness and/or the (excess) kurtosis are irrelevant. Put it differently, this is only when the mean of two Beta distribution functions are different that the skewness and kurtosis might be relevant.

Remark 1. Let \( X \) and \( Y \) two random variables with distribution function \( F_X \) and \( F_Y \), respectively. From Hadar and Russell (1971, Theorem 5), we know that if \( F_Y \leq_{SSD} F_X \), then, any affine transformation of \( X \) and \( Y \) leaves invariant second order stochastic dominance, that is, \( F_{X+\mu} \leq_{SSD} F_{X+\mu} \). Let \( X_1 \) and \( X_2 \) be two Beta random variables. Using our notations through which the distribution function \( F \) is indexed by the mean \( M \) and the variance \( V \), assume that \( F_{M,V} \leq_{SSD} F_{M,V} \). From Hadar and Russell (1971, Theorem 5), it thus follows that if \( Y_1 = \mu X_1 + \mu \) and \( Y_2 = \mu X_2 + \mu \), then, \( F_{a_M+\mu, a_V} \leq_{SSD} F_{a_M+\mu, a_V} \).

3.4. Lattice structure of \( \bar{B} \) and Hasse diagram

Now that each vertical section \( B_M \subset \bar{B} \) has been completely ordered with respect to \( \leq_{SSD} \), let us investigate if \( \bar{B} \) can be ordered in some way with respect to \( \leq_{SSD} \). The boundary of the dome \( \partial B \) is the union of the upper parabola corresponding to a convex combination of \( \delta_0 \) and \( \delta_1 \) and the \( x \)-axis corresponding to \( \delta_0 \) for \( M \in [0,1] \). As we shall now see, both can be ordered from left to right. It is not surprising, but this fact should be pointed out as it becomes useful later.

Proposition 3. If \( M_1 \) and \( M_2 \) are such that \( 0 \leq M_1 < M_2 \leq 1 \), then

(i) \( F_{M_1,0} \leq_{SSD} F_{M_2,0} \),

(ii) \( F_{M_1,D(M_2)} \leq_{SSD} F_{M_2,D(M_2)} \).

Proposition 3 is an important and interesting intermediate result and its proof turns out to be very simple. For this reason, the proof appears here and not in the appendix. For (i), it suffices to note that when \( x < M_1 \) or when \( x \geq M_2 \), \( F_{M_1,0}(x) = F_{M_2,0}(x) \), while when \( x \in [M_1,M_2] \), \( F_{M_1,0}(x) = 0 \) and \( F_{M_2,0}(x) = 1 \). So, \( F_{M_1,0}(x) < F_{M_2,0}(x) \). As a result, \( F_{M_i,0} \leq_{SSD} F_{M_2,0} \). To prove (ii), it suffices to note for all \( x \in (0,1), F_{M_1,D(M_2)}(x) = 1 - M_1 \) and \( F_{M_2,D(M_1)}(x) = 1 - M_2 \). Since \( M_1 < M_2 \leq 1 \), it thus follows that for all \( x \in (0,1), F_{M_1,D(M_1)}(x) > F_{M_2,D(M_2)}(x) \) and this concludes the proof.

We have seen that, for any \( M \in [0,1], F_{M,0} \) the Dirac mass \( \delta_M \) (since we identify the distribution functions and the probability measures). It thus follows that the minimum element can also be denoted by \( \delta_0 \) and the maximum element by \( \delta_1 \). Since it is common to denote the minimum element by \( \bot \) and the maximum element by \( \top \), the above corollary can thus, with a slight abuse of notation, be written as follows.

Corollary 3. \( \bot = \delta_0 \) (i.e., \( F_{0,0} \)) is the minimum element while \( \top = \delta_1 \) (i.e., \( F_{1,0} \)) is the maximum element.

Let \( (X, \leq) \) be a partially ordered set (called poset for short), that is, a set on which there is a binary relation \( \leq \) which is reflexive, antisymmetric and transitive. Let \( x \) and \( y \) two elements of \( X \) and denote
the join (the least upper bound) of x and y as \( x \lor y \) and the meet (greatest lower bound) as \( x \land y \). The poset \( X \) is said to be a lattice (see Topkis (1998) or Milgrom and Shannon (1994)) if for every two pair of elements \( x \) and \( y \) of \( X \), the join and the meet do exist in \( X \). As an elementary example (see e.g., Topkis, 1998, p. 13), the set of real numbers \( \mathbb{R} \) is an example of lattice.

The Hasse diagram is a visual representation of the partial order, where the arrow \( a \rightarrow b \) indicates \( a \leq_{\text{sub}} b \). Using the above results and the transitivity property for comparable elements of the poset (the MV-dome), we obtain paths from the minimum element \( \delta_0 \) to the maximum element \( \delta_1 \). The simplest example of paths can be called trivial ones since they only compare Dirac masses (indeed distribution functions). It is obvious that, for any \( 0 < M < M' < 1 \), one has the following (trivial) path

\[
\text{Trivial path } : \quad \delta_0 \rightarrow \delta_M \rightarrow \delta_M' \rightarrow \delta_1.
\]

Such a path is also trivial from a utility theory point of view since it merely reflects the fact that the underlying utility function increases with (sure) wealth. In the same vein, for \( M \) and \( M' \) such that \( 0 < M < M' < 1 \), the following path

\[
\delta_0 \rightarrow (1-M)\delta_0 + M\delta_1 \rightarrow (1-M')\delta_0 + M'\delta_1 \rightarrow \delta_1
\]

is also trivial. We thus call the non-trivial path the one that makes use of distribution functions \( F_{M,V} \) that lies in \( B \) and not on its boundary. In Figure 3, we provide a representation of a Hasse diagram of the following non-trivial path.

\[
\text{Non-trivial path } : \quad \delta_0 \rightarrow (1-M)\delta_0 + M\delta_1 \rightarrow F_{M,V} \rightarrow \delta_M \rightarrow \delta_1.
\]

3.5. Mean, variance, skewness, and second-order stochastic dominance

3.5.1. Is skewness relevant for mean-preserving spread?

Assume that \( M < 0.5 \) so that the Beta distribution is right-skewed. For a given mean \( M < 0.5 \), the skewness denoted \( Sk \), defined as the third central standardized moment, increases when the variance increases. As a result, the comparison of two Beta distribution \( X_1 \) and \( X_2 \) with the same mean but variance \( V_1 \) and \( V_2 \) such that \( V_1 < V_2 \) and skewness \( Sk(X_1) := Sk_1 \) and \( Sk(X_2) := Sk_2 \) such that \( Sk_1 < Sk_2 \) generates, in principle, a trade-off between the variance and the skewness. A risk-averse (EU) decision-maker might prefer \( X_1 \) while another one might prefer \( X_2 \) because the investor is more prudent but less risk-averse. This is indeed not the case. From the previous result, we know that along a vertical section of the MV-dome in which the mean \( M \) is constant, as long as one decreases the variance from \( V_2 \) to \( V_1 \), the distribution function \( F_{M,V_1} \) second-order stochastically dominates \( F_{M,V_2} \) independently of the resulting skewness (and kurtosis). Two risk-averse (EU) decision-makers, whether they like right-skewed distribution (i.e., \( U'' \) is increasing) or not (i.e., \( U'' \) is decreasing) should prefer \( X_1 \) to \( X_2 \).
independently of the skewness (positive or not) and the kurtosis (positive or not).

3.5.2. What happens when the mean increases while leaving the variance constant?

Before considering the Beta distributions, consider the case of distribution functions that belong to a location-scale family such as the standard Gaussian distribution function denoted \( \Phi \) and let \( m \) and \( v \) denote the mean and the variance. As is well-known (see e.g., Srboonchita et al. (2009) or Müller et al. (2017)), the following property is true: if \( \Phi_{m,v} \leq \text{ssd} \Phi_{m_0,v_0} \), then, for each \( M > m_0 \), \( \Phi_{m,v} \leq \text{ssd} \Phi_{M,v} \), that is, the (second-order) stochastic dominance property is preserved when one increases the mean while leaving the variance constant.

Consider now the Beta distribution. We know that this can be ordered along a vertical segment of the MV-dome for which the mean is constant but the variance varies. It is now natural to inquire whether or not a similar property can be obtained. As the following result shows, the answer is no.

Proposition 4. Let \( F_{M,V} \) and \( F_{M',V} \) be two Beta distribution functions, with \( M \neq M' \). There exists a triplet \( M, M' \) and \( V \) such that \( F_{M,V} \) and \( F_{M',V} \) are not comparable according to second-order stochastic dominance.

Proof. See Appendix A.

Let \( F_{M,V} \) be a distribution function such that \( m \in (0,1) \) and \( v \in (0,D(m)) \) are given. Consider now the distribution function \( F_{M_0,V_0} \) where \( M_0 > m \) and \( V_0 < D(M_0) \). The following result is a consequence of the above proposition.

Corollary 4. If \( F_{M,V} \leq \text{ssd} F_{M_0,V_0} \), then, there exists \( M > M_0 \) such that \( F_{M,V} \) is not SSD-comparable to \( F_{M_0,V} \), where \( V_0 < D(M) \).

The corollary states that if \( F_{M_0,V_0} \) SSD-dominates \( F_{M,V} \), then, by increasing the mean, but leaving the variance constant equal to \( V_0 \), one can find a mean \( M > M_0 \) that is high enough (satisfying the constraint \( D(M) \geq V_0 \)) such that \( F_{M,V} \) is no longer SSD-comparable to \( F_{M_0,V} \). This is a fairly surprising property since it is never true for distribution functions that belong to the same location-scale family.

To see why Corollary 4 holds true, assume that \( M \) is such that \( V_0 = D(M) \), which means that the distribution function is \( F_{M,D(M)} \) for which we know that \( F_{M,D(M)}(x) = 1 - M \) for \( x \in [0,1] \). See Equation (16). Since the distribution \( F_{M,V} \) admits a density, the support of \( F_{M,V} \) is \([0,1]\) and \( F_{M,V} \) is a continuous and strictly increasing function from 0 to 1. This therefore means that there exists a single point \( x_c \in (0,1) \) for which \( F_{M,D(M)}(x_c) = F_{M,V}(x_c) \). When \( x < x_c \), \( F_{M,V}(x) < F_{M,D(M)}(x) \) so that \( \int_0^x F_{M,V}(z)dz < \int_0^x F_{M,D(M)}(z)dz \). Since \( M > m \), using the fact that the mean is the integral of the survival function, it thus follows that \( \int_0^1 F_{M,V}(z)dz > F_{M,D(M)}(z)dz \). Taken together, these two inequalities violate Equation (3b) so that \( F_{M,D(M)} \) and \( F_{M,V} \) are not SSD-comparable. At a more fundamental level, the reason why the above corollary is true for the Beta distribution is due to the following property. When the variance \( V \) is left constant and when one approaches the boundary of the MV-dome as \( M \) increases, the Beta distribution converges toward a discrete random variable.

4. Applications: Expected utility, portfolio choices, and exhaustive numerical analysis

We now consider a simple portfolio problem in which the risk-averse (expected utility) decision-maker (or investor) can invest a fraction \( \gamma \in [0,1] \) of their wealth in a risky asset \( X \), which is random variable following a Beta distribution with mean \( m \) (in percentage) and with a volatility equal to \( \sqrt{v} \), that is, a variance equal to \( v \). For simplicity, we consider the usual Beta distribution with a support equal to \([0,1]\) but it would also be possible, up to a linear transformation, to consider positive and negative of returns. As already said, the Beta distribution is interesting for a given stock because its returns can both be bell-shaped and right skewed (i.e., with a positive skewness) something consistent with the stylized fact called the persistence of skewness for stocks (e.g., Chang et al., 2013; Langlois, 2020; Singleton & Wingender, 1986; Sun & Yan, 2003). It should be pointed out that while we shall make use of the Beta distribution to model the stock returns, a number of alternative candidates are also available (see e.g., Simkowitz and Beedles (1980), see also Adcock et al. (2015) for a review). For instance, Kon (1984) considers a mixture of normal distributions to explain the observed positive skewness.

The remaining fraction, \( 1 - \gamma \) is invested in a risk-free asset for which the rate of return is equal to \( r > 0 \) with probability one. The aim is to analyze the optimal fraction invested in the risky asset. More particularly, we are interested in deriving conditions under which this optimal fraction is equal to 100%.

Let \( W_0 \) be the initial wealth of the investor and assume that \( \mathbb{E}(X) := m - r \), which means that the equity premium \( m - r \) is positive. For a given \( \gamma \in [0,1] \), the final (random) wealth \( W(\gamma) \) after one period (e.g., one year) is equal to

\[
W(\gamma) = W_0 \times \left[ 1 + (1-\gamma)r + \gamma X \right] = W_0 \times \left( 1 + r + \gamma(X-r) \right). \tag{19}
\]
It should be pointed out that the risk-free rate \( r \) is Dirac mass, a distribution function of \( B \), i.e., \( F_{r,0} \).

### 4.1. CARA utility and portfolio choices

Assume now that the decision-maker is endowed with a CARA utility function (i.e., \( U(w) = -e^{-\lambda w} \)) where \( w \) is the wealth) that depends upon a unique risk-aversion parameter \( \lambda \). Let \( \mathbb{E}U(W_f(\gamma)) \) denote the expected utility of the final wealth. The optimization problem thus is

\[
\max_{\gamma \in [0,1]} \mathbb{E}U(W_f(\gamma)) := -\int_0^1 e^{-\lambda w \times (1+r+\gamma(x-r))} dF_{m,v}(x) \tag{20}
\]

where \( \lambda > 0, r \in (0,1) \) and \( F_{m,v} \in B \). One can now analyze the optimal fraction of the initial wealth invested in the risky asset \( X \) as a function of the various parameters. For a given risk-aversion parameter \( \lambda > 0 \) and a risk-free rate \( r \in (0,1) \), let us note \( \gamma^*_{\lambda,r}(m, v) \) the fraction of the initial wealth invested in the risky asset.

**Remark 2.** We now want to solve the optimization problem defined in Equation (20), whose the solution depends upon the two parameters of the Beta distribution. Assume we focus on single peaked Beta distribution, those for which \( \alpha \geq 1 \) and \( \beta \geq 1 \). Since \( [1, \infty) \times [1, \infty) \) is an unbounded set, this means that an exhaustive analysis of the optimal solution \( \gamma^* \) as a function of the parameters is out of reach. Thanks to our change of variable, we are indeed able to provide such an exhaustive analysis since the set of parameters for which the Beta in single-peaked is bounded.

To perform our analysis, we first consider the “worst-case” scenario for the variance. From the previous section, such a “worst-case” scenario appears on the upper boundary of the MV-dome, i.e., the parabola. Since \( m \) is fixed, note that \( F_{m,v} \in B_m \). Since \( B_m \) is closed, there is a distribution function within \( B_m \) that will constitute this “worst-case” scenario and it is obviously \( F_{m,D(m)} \). Since \( F_{m,D(m)} \) reduces to the distribution of a Bernoulli random variable with parameter \( m \), we can explicitly obtain the expected utility in this worst-case scenario. We prove the following lemma in the appendix.

**Lemma 2.** Let \( r \in (0,1) \) and \( m \in (r,1) \). The optimal fraction of the initial wealth \( \gamma^*_{\lambda,r}(m, D(m)) \) invested in the risky asset is equal to

\[
\gamma^*_{\lambda,r}(m, D(m)) := \gamma^*_{\lambda,r}(m) = \min \left\{ \frac{1}{\lambda} \ln \left( \frac{m}{r} \frac{1-r}{1-m} \right) ; 1 \right\} \tag{21}
\]

From Equation (21), it is easy to see that if \( m \leq r \), that is, there is no risk-premium (or equity premium), then, \( \gamma^*_{\lambda,r}(m, D(m)) = 0 \). In Figure 4, we represent the set of parameters denoted \( S_r \) for which the decision-maker invests all their wealth in the default risk-free asset.

However, when \( m > r \), the quantity \( \frac{m(1-r)}{r(1-m)} \) is greater than one so that the optimal fraction invested in the risky asset is always positive, that is \( \gamma^*_{\lambda,r}(m, D(m)) > 0 \). As expected from Equation (21), when \( \lambda \) increases, everything else equal, \( \gamma^*_{\lambda,r}(m, D(m)) \) decreases and tends to zero when \( \lambda \) tends to infinity. On the other hand, from Equation (21), it is not difficult to see that when \( m \) (or \( r \)) increases, \( \gamma^*_{\lambda,r}(m, D(m)) \) increases (decreases).

Let \( \gamma^*_{\lambda,r}(m,v) \) be the optimal fraction of the initial wealth invested in the risky asset for a given mean \( m \) and a given variance \( v \leq D(m) \). In what follows,
From an economic point of view, Lemma 3 means that no matter how small is the equity premium \( r \) that is, the rational investor will always invest a positive fraction of their initial wealth in the risky asset. The CARA utility function will invest a positive fraction of their initial wealth in the risky asset no matter the parameters \((c,\gamma_2,\gamma_3)\) and the risk-free rate \( r \). By definition of the mean threshold, if \( m = \bar{m}_{\lambda,r} \), then, \( \gamma_{\lambda,r}^+(m,D(m)) = 1 \). Since \( F_{m,v} \) second-order stochastically dominates \( F_{m,D(m)} \), it thus follows that \( \gamma_{\lambda,r}^+(m,v) = 1 \). As long as it is optimal to invest 100% of their initial wealth in the risky asset when \( (m,D(m)) \) with \( m > r \), it is obviously still the case when the variance is lower than the its maximal value \( D(m) \). Define \( R_{\lambda,r} \) as follows:

\[
R_{\lambda,r} = \{ (m,v) \in \mathbb{D} : m \in [\bar{m}_{\lambda,r}, 1] \text{ and } v \in [0, D(m)] \}.
\]  

(22)

which is the subset of parameters \((m,v)\) of the MV-dome for which the decision-maker invests 100% of their initial wealth in the risky asset no matter the variance \( v \) (or the volatility \( \sqrt{v} \)) and the expected return on the risky asset \( m \). The next proposition summarizes the discussion.

**Proposition 5.** \( \forall (m,v) \in R_{\lambda,r}, \gamma_{\lambda,r}^+(m,v) = 1 \).

We illustrate Proposition 5 in Figure 4 and provide a graphical representation of the subset \( R_{\lambda,r} \). It is important to point out that since \( \bar{m}_{\lambda,r} \) depends upon the choice of the utility function (that is, the one-parameter CARA utility function), the subset \( R_{\lambda,r} \) also depends upon this choice. However, the
choice of an increasing and concave utility function will change the critical threshold $\bar{m}_{k,r}$ (for a one-parameter utility function) but not the shape of the subset $R_{k,r}$, see Figure 4. It is important to note at this stage that $(m,v) \in R_{k,r}$ is a sufficient condition for $\gamma^*_r(m,v) = 1$. It is thus possible that while $(m,v) \notin R_{k,r}$, $\gamma^*_r(m,v) = 1$. In Figure 5, each point under the parabola represents a solution to the optimization problem. One can clearly see that the region in which $\gamma^*_r(m,v) = 1$ is much larger than $R_{k,r}$ precisely due to the fact that we consider a particular utility function, the CARA.

5. Conclusion

We have shown in this paper how one can transform the set of parameters of the Beta distribution in a meaningful way to compare Beta distribution functions with respect to second-order stochastic dominance. This led us to derive the particular lattice of the Beta distribution and show its striking difference with the Gaussian case when the variance constant. Finally, we have also shown how our approach can be used to perform an exhaustive numerical analysis of an optimization problem that takes as input the parameters of the Beta distribution.

Generalizing our methodology to other special two-parameter probability distributions such as the two-parameter Gamma distribution (or the two-parameter Weibull distribution) would clearly be interesting. Since the mean of the Gamma (or Weibull) random variable as a function of the natural parameters is not bounded, finding a bijective mapping which yields a bounded subset for the set of new parameters, mean and variance, is therefore more difficult and left for further research.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes

1. Note that the notion of second-order stochastic dominance does not require an equal mean.

2. We discovered that this result appears as a corollary in Müller et al. (2017), though with a different approach.

3. Note that instead of $C^1([0,1])$, one could consider $C([0,1])$. In such a case, a random variable with such a distribution is no longer positive, but the assumption of identical mean can be made. One must further assume, however, that the random variables have finite variance.

4. It can be counterintuitive at first glance that an unbounded set and a bounded set can be equinumerous, that is, there exists a bijective mapping between them. This equinumerosity property is indeed very usual. To see this, consider the distribution function of the Gaussian random variable $\psi$. Since $\psi$ is strictly increasing (and continuous), it defines a bijection between $(0,1)$ and $\mathbb{R}$, which means that $(0,1)$ and $\mathbb{R}$ are equinumerous.

5. They both consist of rational fractions with no pole in their domain and thus are continuous. Put it differently, if $f$ and $g$ are two continuous functions, $\frac{f}{g}$ is continuous as long as $g$ is not zero.

6. That is, we note the minimum element as $\beta_0$ instead of $F_{R,0}$ and the same for the maximum element.

7. When $(M, V)$ lies inside the Dome, $F_{M,V}$ admits a density and thus is absolutely continuous with respect to the Lebesgue measure. When $M$ is such that $V = D(M)$, $F_{M,D(M)}$ does not admit a density and $F_{M,D(M)}$ is no longer absolutely continuous with respect to the Lebesgue measure.

8. With further computations we could replace $O(x^2)$ by $O(x^3)$. However, this is not necessary.

9. Further investigation would lead to determine that (iii) occurs. However, it is not necessary for our proof.

ORCID

Yann Braouezec http://orcid.org/0000-0002-2586-2534
John Cagnol http://orcid.org/0000-0002-5447-3335

References

Adcock, C., Eling, M., & Loperfido, N. (2015). Skewed distributions in finance and actuarial science: A review. The European Journal of Finance, 21(13-14), 1253–1281. https://doi.org/10.1080/1351847X.2012.720269

Albuquerque, R. (2012). Skewness in stock returns: Reconciling the evidence on firm versus aggregate returns. Review of Financial Studies, 25(5), 1630–1673. https://doi.org/10.1093/rfs/hhr144

Ali, M. M. (1975). Stochastic dominance and portfolio analysis. Journal of Financial Economics, 2(2), 205–229. https://doi.org/10.1016/0304-405X(75)90005-7

Arnold, B. C., & Groeneveld, R. A. (1995). Measuring skewness with respect to the mode. The American Statistician, 49(1), 34–38.

Balakrishnan, N., & Nevzorov, V. B. (2004). A primer on statistical distributions. John Wiley & Sons.

Batur, D., & Choobineh, F. (2012). Stochastic dominance based comparison for system selection. European Journal of Operational Research, 220(3), 661–672. https://doi.org/10.1016/j.ejor.2012.02.018

Beedles, W. L., & Simkowitz, M. A. (1980). Morphology of asset asymmetry. Journal of Business Research, 8(4), 457–468. https://doi.org/10.1016/0148-2963(80)90018-1

Bernard, C., De Staelen, R. H., & Vanduffel, S. (2019). Optimal portfolio choice with benchmarks. Journal of the Operational Research Society, 70(10), 1600–1621. https://doi.org/10.1080/01605682.2018.1470066

Bilbao, A., Arenas, M., Jiménez, M., Perez Gladish, B., & Rodríguez, M. (2006). An extension of sharpe’s single-index model: Portfolio selection with expert betas. Journal of the Operational Research Society, 57(12), 1442–1451. https://doi.org/10.1057/palgrave.jors.2602133
Brioni, R., Cesaroni, F., Scozzari, A., & Tardella, F. (2017). On exact and approximate stochastic dominance strategies for portfolio selection. *European Journal of Operational Research*, 259(1), 322–329. https://doi.org/10.1016/j.ejor.2016.10.006

Bunn, D. W. (1979). Stochastic dominance. *Journal of the Operational Research Society*, 30(9), 845–848.

Cesaroni, F., & Colucci, S. (2017). Minimum risk versus capital and risk diversification strategies for portfolio construction. *Journal of the Operational Research Society*, 69, 1–18.

Chang, B. Y., Christoffersen, P., & Jacobs, K. (2013). Market skewness risk and the cross section of stock returns. *Journal of Financial Economics*, 107(1), 46–68. https://doi.org/10.1016/j.jfineco.2012.07.002

Chateauneuf, A., Cohen, M., & Meilijson, I. (2004). Four notions of mean-preserving increase in risk, risk attitudes and applications to the rank-dependent expected utility model. *Journal of Mathematical Economics*, 40(5), 547–571. https://doi.org/10.1016/S0304-4068(03)00044-2

Cohen, M. D. (1995). Risk-aversion concepts in expected-and non-expected-utility models. *The Geneva Papers on Risk and Insurance: Issues and Practice*, 20(1), 73–91. https://doi.org/10.1017/S0227409300003783

Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223–236. https://doi.org/10.1080/713665670

DeFusco, R. A., Karels, G. V., & Muralidhar, K. (1996). Skewness persistence in us common stock returns: Results from bootstrapping tests. *Journal of Business Finance & Accounting*, 23(8), 1183–1195. https://doi.org/10.1111/j.1468-5956.1996.tb01164.x

Denis, M., Eeckhoudt, L., & Jokung, O. (2013). Non-differentiable transformations preserving stochastic dominance. *The Journal of Operational Research Society*, 64(9), 1441–1446. https://doi.org/10.1057/jors.2012.140

Diamond, P. A., & Stiglitz, J. E. (1974). Increases in risk and in risk aversion. *Journal of Economic Theory*, 8(3), 337–360. https://doi.org/10.1016/0022-0531(74)90090-8

Ebrahimi, N., Maasoumi, E., & Soofi, E. S. (1999). Ordering univariate distributions by entropy and variance. *Journal of Econometrics*, 90(2), 317–336. https://doi.org/10.1016/S0304-0221(98)00046-3

Egozcue, M., & Wong, W.-K. (2010). Gains from diversification based on stochastic dominance under incomplete probability information. *European Journal of Operational Research*, 286(2), 755–768. https://doi.org/10.1016/j.ejor.2010.03.042

Milgrom, P., & Shannon, C. (1994). Monotone comparative statics. *Econometrica*, 62(1), 157–180. https://doi.org/10.2307/2951579

Moitra, S. D. (1990). Skewness and the beta distribution. *Journal of the Operational Research Society*, 41(10), 953–961. https://doi.org/10.1057/jors.1990.147

Müller, A., Scarsini, M., Tsetlin, I., & Winkler, R. L. (2017). Between first-and second-order stochastic dominance. *Management Science*, 63(9), 2933–2947. https://doi.org/10.1287/mnsc.2016.2486

Pittman, J. (2009). Probability. Springer.

Post, T., & Kopa, M. (2013). General linear formulations of stochastic dominance criteria. *European Journal of Operational Research*, 230(2), 321–332. https://doi.org/10.1016/j.ejor.2013.04.015

Rinne, H. (2011). *Location-scale distributions*. Springer.

Rothschild, M., & Stiglitz, J. E. (1970). Increasing risk: I. A definition. *Journal of Economic Theory*, 2(3), 225–243. https://doi.org/10.1016/0022-0531(70)90038-4

Simkowitz, M. A., & Beedles, W. L. (1980). Asymmetric stable distributed security returns. *Journal of the American Statistical Association*, 75(370), 306–312. https://doi.org/10.1080/01621459.1980.10477467

Singleton, J. C., & Wingender, J. (1986). Skewness persistence in common stock returns. *The Journal of Financial and Quantitative Analysis*, 21(3), 335–341. https://doi.org/10.2307/2331046

Sriboonchita, S., Wong, W.-K., Dhompongsa, S., & Nguyen, H. T. (2009). Stochastic dominance and applications to finance, risk and economics. CRC Press.

Sun, Q., & Yan, Y. (2003). Skewness persistence with optimal portfolio selection. *Journal of Banking & Finance*, 27(5), 929–955.
Appendix A. Technical propositions and proofs

Proof of Proposition 1. Consider two elements $F$ and $G$ of $\mathcal{F}_\mu$. By definition, $F$ and $G$ have a unique non-trivial crossing point $x_0(F, G) := x_0 \in (0, 1)$. Assume that $F$ and $G$ are such $F(x) < G(x)$ for $x \in [0, x_0)$ and $F(x) > G(x)$ for $x \in (x_0, 1]$. It thus follows that

$$\forall x \in (0, x_0), \int_0^x F(z)dz \leq \int_0^x G(z)dz.$$  

Moreover, $F$ and $G$ have the same mean, therefore $\int_0^1 (1 - F(z))dz = \int_0^1 (1 - G(z))dz$, which, in turns gives $\int_0^1 F(z)dz = \int_0^1 G(z)dz$. It follows that

$$\forall x \in [0, 1], \int_0^1 F(z)dz \leq \int_0^1 G(z)dz. \quad (A1)$$

As a result, $G \leq_{id} F$.

Consider now $H$ in $\mathcal{F}_\mu$ and assume that $H \neq G$. Either $H \leq_{id} G$ or $G \leq_{id} H$. Without loss of generality, assume that $H \leq_{id} G$, which is equivalent to for all $x \in [0, 1]$, $\int_0^x G(z)dz \leq \int_0^x H(z)dz$. Since for all $x \in [0, 1]$, $\int_0^x F(z)dz \leq \int_0^x G(z)dz$, it thus follows that $\int_0^1 F(z)dz \leq \int_0^1 G(z)dz \leq \int_0^1 H(z)dz$ so that if $G \leq_{id} F$ and $H \leq_{id} G$, then, $H \leq_{id} G \leq_{id} F$, that is, $\leq_{id}$ is transitive.

If $F_1 \leq_{id} F_2$ and $F_2 \leq_{id} F_1$, then (A1) yields $F_1 = F_2$. Thus antisymmetry holds.

Reflexivity also holds. The set $\mathcal{F}_\mu$ is completely ordered with respect to $\leq_{id}$. \hfill \Box

Lemma 4. Define

$$C_1(M) = \frac{M^2(1-M)}{1+M} \quad \text{and} \quad C_2(M) = \frac{M(1-M)^2}{2-M}.$$  

Let $\phi$ be the mapping $(x, \beta) \mapsto (M, V)$ where $(M, V)$ is given by (13). Then

$$\phi((1, + \infty) \times (1, + \infty)) = \{(M, V) \mid M \in (0, 1) \text{ and } V < \min(C_1(M), C_2(M))\}, \quad (A2.A)$$

$$\phi((0, 1) \times (0, 1)) = \{(M, V) \mid M \in (0, 1) \text{ and } V > \max(C_1(M), C_2(M))\}, \quad (A2.U)$$

$$\phi((0, 1) \times (1, + \infty)) = \{(M, V) \mid M \in (0, 1/2) \text{ and } C_1(M) < V < C_2(M)\}, \quad (A2.D)$$

$$\phi((1, + \infty) \times (0, 1)) = \{(M, V) \mid M \in (1/2, 1) \text{ and } C_2(M) < V < C_1(M)\}. \quad (A2.I)$$

Proof of Lemma 4. From (14) we have that $\alpha > 1$ and $\beta > 1$ is equivalent to the next five assertions:

$$\frac{M(M-M^2-V)}{V} > 1 \quad \text{and} \quad \frac{(1-M)(M-M^2-V)}{V} > 1,$$

$$V < M(M-M^2-V) \text{ and } V < (1-M)(M-M^2-V),$$

$$V(1+M) < M(M-M^2) \text{ and } V(2-M) < (1-M)(M-M^2),$$

$$V < \frac{M^2(1-M)}{1+M} \text{ and } V < \frac{M(1-M)^2}{2-M}.$$  

which is equivalent $V < \min(C_1(M), C_2(M))$. This proves (A2.A). Proofs for Equations (A2.U), (A2.D), and (A2.I) are analogous by adapting sense of the inequalities. \hfill \Box

Proof of Lemma 1. Consider $(M, V) \in \mathcal{D}$. Our goal is to prove that $f_{M,V}$ is point-wise convergent to $F_{M,D(M)}$ when $V$ tends to $D(M) = M-M^2$.

For $x = 1$, we have $\forall V \in (0, D(M))$, $f_{M,V}(1) = 1 = F_{M,D(M)}(1)$, which is to be expected for a distribution function. From now on, let us consider $x \in [0, 1)$.

The Laurent series of $G$ in 0 is

$$\Gamma(z) = \frac{1}{\gamma} - \gamma + \left(\frac{\pi^2}{12} + \frac{\gamma^2}{2}\right)z + O(z^2)$$

where $\gamma$ is the Euler constant.
Using (13) and denoting \( z = (D(M) - V)/V \), which tends to 0 when \( V \) tends to \( D(M) \), we have
\[
z = Mz, \quad \beta = (1-M)z \quad \text{and} \quad z + \beta = z.
\]
Therefore
\[
\Gamma(z)\Gamma(\beta) = \frac{1}{M(1-M)z^2} - \frac{\gamma}{\Gamma(1-M)z} - \frac{1}{12M(1-M)z} + O(z),
\]
\[
\Gamma(z + \beta) = \frac{1}{z - \gamma + O(z)}.
\]
Using
\[
B(z, \beta) = \frac{\Gamma(z)\Gamma(\beta)}{\Gamma(z + \beta)}
\]
we get
\[
\frac{1}{B(z, \beta)} = M(1-M)z + O(z^2).
\]

Subsequently
\[
f_{M,V}(x) = x^{Mz-1}(1-x)^{(1-M)z-1} M(1-M) (1 + O(z)),
\]
\[
F_{M,V}(x) = \left( \int_0^x t^{Mz-1}(1-t)^{(1-M)z-1} dt \right) M(1-M) (1 + O(z)). \tag{A3}
\]

Let \( y \) be any real number in \([0, x]\) (as a reminder: \( x < 1 \))
\[F_{M,V}(x) = \left( \int_0^y t^{Mz-1}(1-t)^{(1-M)z-1} dt + \int_y^x t^{Mz-1}(1-t)^{(1-M)z-1} dt \right) M(1-M) (1 + O(z)). \tag{A4}\]

- Regarding the first integral. For \( t \in [0, y] \) we have
\[
t^{Mz-1} \leq t^{Mz-1}(1-t)^{(1-M)z-1} \leq t^{Mz-1}(1-y)^{(1-M)z-1},
\]
\[
\int_0^y t^{Mz-1} dt \leq \int_0^y t^{Mz-1}(1-t)^{(1-M)z-1} dt \leq (1-y)^{(1-M)z-1} \int_0^y t^{Mz-1} dt,
\]
\[
\frac{1}{Mz} y^z \leq \int_0^y t^{-1}(1-t)^{\beta-1} dt \leq (1-y)^{\beta-1} \frac{1}{Mz} y^z.
\]

- Regarding the second integral. For \( t \in [y, x] \) we have
\[
t^{Mz-1}(1-t)^{(1-M)z-1} \leq y^{Mz-1}(1-x)^{(1-M)z-1},
\]
\[
\int_y^x t^{Mz-1}(1-t)^{(1-M)z-1} dt \leq y^{Mz-1}(1-x)^{(1-M)z-1} (x-y).
\]

Hence
\[
(1-M)y^{Mz}(1 + O(z)) \leq f_{M,V}(x)
\]
\[
\leq (1-y)^{(1-M)z-1}(1-M)y^{Mz} + y^{Mz-1}(1-x)^{(1-M)z-1} (x-y)\beta M)(1 + O(z)).
\]

When \( V \) tends to \( D(M) \), we have \( z \) tends to 0 and we get
\[
1-M \leq \lim_{V \to D(M)} F_{M,V}(x) \leq \frac{1-M}{1-y}.
\]

Since this inequality holds for any \( y \in [0, x] \) and for any \( x \in [0, 1) \), we have
\[
\lim_{V \to D(M)} F_{M,V}(x) = 1-M = F_{M,D(M)}(x)
\]
where \( F_{M,D(M)} \) is the distribution function of \( \delta_M \).

Proof of Proposition 2. Let \( M \in (0, 1) \) and \( V_1, V_2 \) in \((0, D(M))\). Assume \( V_1 \neq V_2 \). With no loss of generality, we can assume \( V_1 < V_2 \). Consider the distribution functions \( F_{M,V_1} \) and \( F_{M,V_2} \).

First statement. Since \( V_1 \) and \( V_2 \) are in the open interval \((0, D(M))\), the distributions functions \( F_{M,V_1} \) and \( F_{M,V_2} \) have corresponding densities \( f_{M,V_1} \) and \( f_{M,V_2} \). We define \( g_{M,V_1,V_2}(x) = \ln f_{M,V_1}(x) - \ln f_{M,V_2}(x) \) for all \( x \in (0, 1) \).

Let \((x_1, \beta_1) = (M, V_1)\) and \((x_2, \beta_2) = (M, V_2)\). Consider \( C_{M,V_1, V_2} = \ln B(x_1, \beta_1) - \ln B(x_2, \beta_2) \) where \( B \) is defined in (9). By construction, \( C_{M,V_1, V_2} \) is a constant with respect to \( x \). We have
\[
g_{M,V_1,V_2}(x) = [(2x_1-1)\ln(1-x) - (2\beta_1-1)\ln(1-\beta_1)] - [(2x_1-1)\ln(1-x) + (2\beta_1-1)\ln(1-\beta_1)] + C_{M,V_1, V_2}.
\]
Using 
\[ x_2 - x_1 = M^2(1-M) \frac{V_2-V_1}{V_1V_2} \] and 
\[ \beta_2 - \beta_1 = M(1-M)^2 \frac{V_2-V_1}{V_1V_2} \]
we get 
\[ g_{M, V_1, V_2}(x) = M(1-M) \frac{V_2-V_1}{V_1V_2} [M \ln(x) + (1-M) \ln(1-x)] + C_{M, V_1, V_2}. \]

The function \( g_{M, V_1, V_2} \) is differentiable on \((0, 1)\). We have 
\[ g'_{V_1, V_2}(x) = M(1-M) \frac{V_2-V_1}{V_1V_2} \frac{M-x}{x(1-x)} \]
thus \( g_{M, V_1, V_2} \) is increases on \((0, M)\), reaches a maximum in \( M \) and decreases on \((M, 1)\). We further have 
\[ \lim_{x \to 0} g_{M, V_1, V_2}(x) = -\infty \text{ and } \lim_{x \to 1} g_{M, V_1, V_2}(x) = -\infty. \]

Thus three situations can occur:

(i) if \( g_{M, V_1, V_2}(M) < 0 \) then \( g_{M, V_1, V_2} \) has no root in \((0, 1)\);
(ii) if \( g_{M, V_1, V_2}(M) = 0 \) then \( g_{M, V_1, V_2} \) has one root in \((0, 1)\), it is \( M \);
(iii) if \( g_{M, V_1, V_2}(M) > 0 \) then \( g_{M, V_1, V_2} \) has two roots in \((0, 1)\), one in \((0, M)\) and one in \((M, 1)\).

Since \( x \mapsto \ln(x) \) is injective on \((0, 1)\), having \( g_{M, V_1, V_2}(x) = 0 \) is equivalent to having 
\[ f_{M, V_1}(x) = f_{M, V_2}(x). \] (A5)

This latter equation has, at most, two roots in \((0, 1)\).

Now, let us go back to \( F_{M, V_1} \) and \( F_{M, V_2} \). Define \( H = F_{M, V_1} - F_{M, V_2} \). We have \( H(0) = 0 \) and \( H(1) = 1 \). If \( H \) were to vanish twice on \((0, 1)\), let \( x_1 \) and \( x_2 \) be its two roots. Applying the Rolle’s theorem on \((0, x_1)\), on \((x_1, x_2)\) and on \((x_2, 1)\) shows that \( H' \) has three roots on \((0, 1)\) which is in contradiction with \( \text{(A5)} \) since \( H' = f_{M, V_1} - f_{M, V_2} \). Subsequently, \( F_{M, V_1} \) and \( F_{M, V_2} \) intersect, at most, once on \((0, 1)\).

If \( F_{M, V_1} \) and \( F_{M, V_2} \) were not to intersect at all then one would dominate the other on \((0, 1)\) which would lead the integral of one to be strictly higher than the integral of the other, contradicting the hypothesis that their means are equal. It follows that \( F_{M, V_1} \) and \( F_{M, V_2} \) intersect exactly once on \((0, 1)\).

**Second statement.** Since \( F_{M, V_1} \) and \( F_{M, V_2} \) intersect only once, it is sufficient to prove that \( F_{M, V_1}(x_0) \leq F_{M, V_2}(x_0) \) for one \( x_0 \in (0, x_1) \). To do so, let us consider the decreasing bijection function \( Z \) from \((0, D(M))\) to \((0, +\infty)\) defined by 
\[ Z(V) = (D(M)-V)/V \]

It is continuously differentiable with respect to its first and second variables. Using \( (13) \), the coefficients \( \alpha \) and \( \beta \) corresponding to \( F_{M, Z^{-1}(z)} \) are \( \alpha = Mz \) and \( \beta = (1-M)z \). Subsequently, 
\[ \frac{\partial \zeta(z, x)}{\partial z} = \frac{1}{B(Mz, (1-M)z)} \int_0^1 t^{Mz-1} (1-t)^{(1-M)z-1} \left[ M \ln t + (1-M) \ln(1-t) + \Psi(z) - (1-M) \Psi((1-M)z) - M \Psi(Mz) \right] dt \]
where \( \Psi \) is the digamma function. Since \( \lim_{z \to 0} [M \ln t + (1-M) \ln(1-t)] = -\infty \), there exists \( h_{M, z} > 0 \) such that 
\[ \forall t \in (0, h_{M, z}), \quad M \ln t + (1-M) \ln(1-t) + \Psi(z) - (1-M) \Psi((1-M)z) - M \Psi(Mz) < 0. \]

Such an \( h_{M, z} \) depends on \( M \) and \( z \) but not on \( t \). Consider \( h = \min(h_{M, Z(V_2)}), h_{M, Z(V_1)}, x_0) \), and \( x_0 \in (0, h) \). Then, for all \( t \in (0, x_0) \) and for all \( z \in [Z^{-1}(V_2), Z^{-1}(V_1)] \), 
\[ \int_0^{x_0} t^{Mz-1} (1-t)^{(1-M)z-1} [M \ln t + (1-M) \ln(1-t) + \Psi(z) - (1-M) \Psi((1-M)z) - M \Psi(Mz)] dt < 0. \]

It follows that \( z \mapsto \zeta(z, x_0) \) is decreasing on \([Z^{-1}(V_2), Z^{-1}(V_1)]\). Since \( Z \) is decreasing, it implies that 
\[ F_{M, V_1}(x_0) < F_{M, V_2}(x_0). \]

**Lemma 5.** For all \( V \in (0, D(M)) \), the mapping \( Y_{m, V} \) is convex.

**Proof of Lemma 5.** This mapping \( Y_{m, V} \) is twice differentiable and \( Y''_{m, V} = F''_{M, V} = f_{M, V} \) which is positive on \((0, 1)\).

**Lemma 6.** For \( x \in (0, 1) \), the application \( V \mapsto Y_{m, V}(x) \) is increasing.

**Proof of Lemma 6.** Let \( m \in (0, 1) \) and \( V_1 \) and \( V_2 \) be in \((0, D(m))\).
Assume that \( Y_{m,V_1} \) and \( Y_{m,V_2} \) cross in \((0, 1)\). Note \( x_0 \) such that \( Y_{m,V_1}(x_0) = Y_{m,V_2}(x_0) \). Then \( H = Y_{m,V_1} - Y_{m,V_2} \) has three roots in \([0, 1]\) that are \( 0, x_0 \) and \( 1 \). Applying the Rolle's theorem on \((0, x_0)\) and \((x_0, 1)\) shows that \( F_{m,V_1} \) and \( F_{m,V_2} \) intersect twice, which contradicts Proposition 2.

Further assume that \( V_1 < V_2 \). The ordering is a consequence of Proposition 2, part 2.

**Proof of Lemma 2.** Let \( X \) be a random variable having for distribution function \( F_{m,D(M)} \).

\[
\mathbb{E}U(1 + r + \gamma(X-r)) = (1 - m)U(1 + r - \gamma r) + mU(1 + r - \gamma r + \gamma)
\]

which can be differentiated with respect to \( \gamma \). We have

\[
\frac{\partial}{\partial \gamma} \mathbb{E}U(1 + r + \gamma(X-r)) = \frac{\lambda}{1 - \exp(-\lambda)} \exp((r - r - 1)m(1 - r)\exp(-\lambda \gamma) - (1 - m)r)
\]

(A6)

which is of the sign of \( m(1 - r)\exp(-\lambda \gamma) - (1 - m)r \). Let \( \gamma^*_r(r, m) \) be defined by (21), then \( \gamma \mapsto \mathbb{E}U(1 + r + \gamma(X-r)) \) is non-decreasing on \((0, \gamma^*_r(r, m))\) and decreasing hereafter if \( \gamma^*_r(r, m) < 1 \). This yields the result.

**Proof of Proposition 4.** Assume that \( M < M' \), with \((M, M') \in (\frac{1}{2}, 1)^2 \). Let \( V > 0 \) be such that \( D(M') = V \). Since the boundary of the \( MV \)-dome is a parabola, \( V \) exists and \( V < D(M) \) since \( M < M' \). From (16), we know that \( F_{M,V} \) is equal to \( 1 - M \) when \( x \in [0, 1] \) and is equal to 1 when \( x = 1 \). Since \( V \) is positive while \( V < D(M) \), \( F_{M,V} \) admits a (strictly) positive density \( f_{M,V} \) so that the support (of \( F_{M,V} \)) is \([0, 1]\). Moreover, the distribution function \( F_{M,V} \) is continuous and strictly increasing since the density \( f_{M,V} \) is (strictly) positive (and continuous). It thus follows that there exists a single crossing point \( x_c \in (0, 1) \) such that \( F_{M,V}(x) = F_{M,V}(x_c) \). The crossing point is such that for all \( x < x_c \), \( F_{M,V}(x) < F_{M,V}(x) \) while for all \( x > x_c \), \( F_{M,V}(x) > F_{M,V}(x) \). Clearly, for all \( x < x_c \), \( \int_0^x F_{M,V}(z)dz < \int_0^x F_{M,V}(z)dz \). However, since \( M' > M \), it thus follows that \( \int_0^1 S_{M,V}(z)dz < \int_0^1 S_{M,V}(z)dz \) (recall that \( S \) is the survival function), which is equivalent to \( \int_0^1 F_{M,V}(z)dz < \int_0^1 F_{M,V}(z)dz \). As a result, \( F_{M,V} \) and \( F_{M',V} \) are not comparable according to second-order stochastic dominance.