Upper Bounds for Concave Distortion Risk Measures on Moment Spaces

D. Cornilly[*], L. Rüschendorf[†] and S. Vanduffel[‡]

October 16, 2017

Abstract

The study of worst-case scenarios for risk measures (e.g., quantiles) when the underlying risk (or portfolio of risks) is not completely specified is a central topic in the literature on robust risk measurement. For instance, upper bounds for quantiles and stop-loss premiums under the sole knowledge of some of the moments of the underlying risk are available in the academic literature. In this paper, we tackle the open problem of deriving upper bounds for concave distortion risk measures on moment spaces. Building on results of Rustagi (1957, 1976), we show that in general this problem can be reduced to a parametric optimization problem. We obtain in explicit form the sharp upper bound when the first moment and some other higher moment are fixed.

Key-words: Value-at-Risk (VaR); Coherent risk measure; Model uncertainty; Distortion function.

[*]Dries Cornilly, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Bruxelles, Belgium. KU Leuven, Celestijnenlaan 200B, 3001 Heverlee, Belgium. (email: `dries.cornilly@kuleuven.be`). Dries Cornilly acknowledges support from the Research Foundation Flanders (PhD fellowship 1114117N).

[†]Ludger Rüschendorf, University of Freiburg, Eckerstraße 1, 79104 Freiburg, Germany. (email: `ruschen@stochastik.uni-freiburg.de`).

[‡]Steven Vanduffel, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Bruxelles, Belgium. (email: `steven.vanduffel@vub.ac.be`).

Introduction

In this paper, we study sharp upper bounds on the risk of a portfolio with respect to a (strictly) concave distortion risk measure when the underlying risk is not fully specified in that only some information on its moments is available. This problem is relevant for various reasons. *First*, a concave distortion risk measure is coherent (Artzner, Delbaen, Eber, and Heath (1999)) and thus has all properties that "good" risk measures are typically expected to have. Moreover, if in addition to the coherency of a law-invariant risk measure, one also requires comonotone additivity, then concave distortion risk measures are the only ones that are admissible (Kusuoka (2001)). *Second*, measuring the risk of portfolios is at the center of insurance activities. When the marginal distribution functions of the portfolio components as well as their dependence structure is known, the risk of the portfolio can be numerically assessed by using for instance Monte-Carlo simulation. In most cases, however, it cannot be expected that full information on the dependence structure is available and various stakeholders such as investors and regulators could be interested in the worst-case scenario for the portfolio (i.e., when the risk measure attains its highest value). In this regard, we note that there is a rich literature on finding bounds for quantiles - also called Value-at-Risk (VaR) - of a portfolio under the assumption that all marginal distribution functions are known, but the dependence is (partially) unknown[1]. In this paper, however, we do not fix the marginal distribution functions, but derive bounds under the sole knowledge of some moments of the portfolio loss (for instance based on portfolio statistics) without specification of the marginal distribution functions. Moreover, we consider the class of concave distortion risk measures and the VaR does not belong to this class.

The most well-known concave distortion risk measure is the Tail Value-at-Risk (TVaR), also called Expected Shortfall in the literature. In fact, TVaR is the smallest coherent risk measure that is greater than the Value-at-Risk (VaR), which is the most frequently used risk measure in risk management and supervision practice, but which fails to be subadditive and thus lacks coherency. Effectively, the VaR is a particular quantile of the distribution whereas TVaR[2] is more focused on the right tail of the distribution in that it measures the expected loss, conditionally on the loss being greater than VaR. Moment bounds for VaR (which are intimitely connected to distributional bounds) and TVaR have already been studied in the insurance literature by several authors including Kaas and Goovaerts (1986), Denuit et al. (1999b), De Schepper and Heijnen (2010), Hürlimann (2002, 2008), Bernard et al. (2015, 2016) and Tian (2008). Specifically, Hürlimann (2002) finds analytical bounds for VaR and TVaR under knowledge of the mean, variance, skewness and kurtosis. An elementary derivation of bounds on VaR can be found in Bernard et al. (2016). In this regard, we point out that one cannot expect that there exists a risk measure (i.e., a single number) that captures all characteristics of risk and provides a complete picture of

[1]Papers include Rüschendorf (1982), Denuit et al. (1999a), Wang and Wang (2011), Wang et al. (2013), Embrechts et al. (2013, 2014, 2015a), Puccetti et al. (2016, 2017), Bernard et al. (2015, 2017b,a), and Rüschendorf and Witting (2017), amongst others.

[2]There are various proofs that demonstrate subadditivity of TVaR; see Embrechts et al. (2015b) and the references herein.

the risky portfolio (i.e., a random variable). For example, Hürlimann (2002) studies TVaR for various two-parameter distribution functions with fixed mean and variance by varying the loss probability and argues that TVaR does not always properly reflect the increase in (tail) risk from one distribution to another. Moreover, risk measures appear in various contexts such as risk management (McNeil et al. (2015)), pricing (Wirch and Hardy (1999)), capital allocation (Dhaene et al. (2012)), and supervision (Danielsson et al. (2001)) and a risk measure that is suitable for one purpose might be not appropriate in another context; see also Dhaene et al. (2008) for a warning against the blind use of coherent risk measures as well as Belles-Sampera et al. (2014) and Frittelli et al. (2014) for recent proposals of risk measures.

In this paper, we study bounds for any strictly concave distortion risk measure when k, not necessarily consecutive, moments of the underlying risk are known. Specifically, based on results of Rustagi (1957, 1976) we provide necessary conditions that maximizing distribution functions have to satisfy. As a consequence, the optimization problem we consider reduces to a problem in which we only need to perform an optimization with respect to some parameters. The conditions that we obtain are in general not sufficient to single out the maximizing distribution (sharp bound). However, when the mean and any other higher moment are known, we show that the feasible set of maximizing distribution functions becomes a singleton and we explicitly obtain the maximizing distribution. Interestingly, unlike in the case of VaR and TVaR, the maximizing distribution is typically not a discrete one.

1 Problem formulation

In this paper, we only consider distribution functions with a bounded support. Hence, after rescaling, we consider them on the unit interval $[0, 1]$. Denote by $\mathcal{F}$ a set of distribution functions on $[0, 1]$ for which $k \in \mathbb{N}_0$ moments are given,

$$\mathcal{F} = \left\{ F \text{ is a cdf} \;\middle|\; \int_0^1 x^i dF(x) = c_i, i \in \mathcal{I} \right\}, \tag{1}$$

where $\mathcal{I} \subset \mathbb{N}_0$ and $card(\mathcal{I}) = k$. Note that in general $\mathcal{F}$ may correspond to a set of distribution functions with any k moments fixed, not necessarily the first k ones, and not necessarily starting with the mean. In the remainder of the paper, we assume that $\mathcal{F}$ contains at least two different elements (and hence infinitely many since $\mathcal{F}$ is convex).

A distortion risk measure of a random variable X having cumulative distribution F_X is defined as

$$H_g(X) = \int_0^1 g(1 - F_X(x))dx, \tag{2}$$

where g is a distortion function, i.e., an increasing function from $[0, 1]$ to $[0, 1]$ with $g(0) = 0$ and $g(1) = 1$. Note that $H_g(X)$ solely depends on the distribution function F_X (law-invariance) and in what follows we also write $H_g(F_X)$ instead of $H_g(X)$.

Furthermore, we assume that g is strictly concave and twice differentiable, implying that H_g is a coherent risk measure; see e.g., Dhaene et al. (2006). The importance of distortion risk measures with concave distortion function (henceforth called concave disortion risk measures) is highlighted by the fact that this class coincides with the class of coherent risk measures that are law-invariant and comonotone additive (Kusuoka (2001)). Examples of concave distortion risk measures are the power distortion risk measure ($g(x) = x^\alpha, \alpha \in (0, 1)$), the dual-power distortion risk measure ($g(x) = 1 - (1 - x)^\beta, \beta \in (1, \infty)$), and the Wang distortion risk measure ($g(x) = \Phi(\Phi^{-1}(x) + \Phi^{-1}(p)), p \in (0.5, 1)$).

In this paper, we focus on the problem of determining the distribution function in $\mathcal{F}$ that yields maximum (concave) distorted expection, i.e., we consider the optimization problem

$$\sup_{F \in \mathcal{F}} H_g(F). \tag{3}$$

When only one moment is specified, say the i-th one with value c_i, it is easy to show that the solution is obtained by a discrete distribution function F that is concentrated on 0 and 1 and has i-th moment equal to c_i. To see this, observe that F dominates all other admissible distribution functions in the sense of stop-loss order (since F crosses all other distribution functions exactly once from above and has the biggest possible mean, namely c_i) and it is well-known that concave distortion risk measures are consistent with stop-loss order (see e.g., Dhaene et al. (2006)). Hence, the case $k = 1$ is not interesting and moreover, since little distributional information is used in the optimization, this case is not very useful in practice in that it leads to wide bounds. Therefore, in the remainder of the paper, we only consider the case in which $k \geqslant 2$.

Problem (3) can be seen as an extended version of an optimization problem considered in Rustagi (1957, 1976). This author considers the optimization of a certain integral when the first and second moment are known and provides some necessary conditions its solution has to satisfy. In this paper, we show that optimization of concave distortion risk measures is compatible with this integral formulation and provide, for an arbitrary sequence of moments, the necessary conditions a solution has to satisfy. Rustagi (1957, 1976) claims that in certain cases the necessary conditions he derives lead to complete specification of the solution, but a proof is missing and appears to be non-trivial. In this paper, we completely specify the maximizing distribution function when the mean and a higher moment are known and provide an algorithm to obtain the solution. For a specific choice of the distortion function we also obtain an explicit analytical solution.

In what follows, let F^- and F^+ denote the left and right inverse of the function F. When $F^- = F^+$, we use the standard notation F^{-1} for the inverse.

2 Results

In this section, we provide necessary conditions for the solution, denoted as $F^\star$, to problem (3). In general, these conditions do not lead to a complete specification of the solution. However, in the particular (yet important) case where $\mathcal{F}$ contains the distribution functions having a given mean and some other higher order moment, a complete characterization of a solution $F^\star$ is obtained and we provide a constructive algorithm. Proofs are provided in Section 4.

2.1 k known moments ($k \geqslant 2$)

The following theorem provides necessary conditions that a solution to problem (3) has to satisfy.

Theorem 2.1 (Necessary conditions)**.** *When $k \geqslant 2$, then problem (3) has a unique solution $F^\star$ and $F^\star$ is continuous on $(0,1)$. Moreover, on intervals where it is not constant, $F^\star$ coincides with $F_{\boldsymbol{\eta}}$, which is defined as*

$$F_{\boldsymbol{\eta}}(x) = \begin{cases} 0 & \text{if } \sum_{i\in\mathcal{I}} \eta_i x^{i-1} < g'(1), \\ 1 - (g')^{-1}\left(\sum_{i\in\mathcal{I}} \eta_i x^{i-1}\right) & \text{else}, \\ 1 & \text{if } \sum_{i\in\mathcal{I}} \eta_i x^{i-1} \geqslant g'(0), \end{cases} \tag{4}$$

where $\boldsymbol{\eta} := (\eta_1, \ldots, \eta_k)$ is such that

$$\int_0^1 x^i dF^\star(x) = c_i, \qquad i \in \mathcal{I}, \tag{5}$$

and

$$\int_a^b \left[g'\left(1 - F^\star(x)\right) - \sum_{i\in\mathcal{I}} \eta_i x^{i-1} \right] dx = 0 \tag{6}$$

on intervals $[a,b] = [F^{\star-}(c), F^{\star+}(c)]$ where $c \in (0,1)$.

Theorem 2.1 does not completely characterize the maximizing distribution function $F^\star$. Specifically, Theorem 2.1 reduces the problem (3) to a parametric optimization problem over the parameter set of admissible $(\eta_1, \eta_2, ..., \eta_k)$ vectors. Since this set is not readily known, this parametric optimization problem is typically hard to deal with. However, hereafter we show that when the mean and some other higher order moment are known, we are able to determine the maximizing distribution function.

In this regard, we note that Rustagi (1957, 1976) claims that in the case that he considers (first two moments given) the unique solution is given by $F_{\boldsymbol{\eta}}$, as defined in Theorem 2.1, provided $F_{\boldsymbol{\eta}}$ is a distribution function; see page 104 in Rustagi (1976) and subsequent theorems as well as the corollary on page 318 in Rustagi (1957). However, there is no proof for this statement and the argument is not obvious since there could be several admissible distribution functions of the form $F_{\boldsymbol{\eta}}$.

2.2 Mean and a higher order moment known

Let $\mathcal{F} = \mathcal{F}(c_1, c_i)$ be the set of distribution functions with fixed first moment c_1 and fixed i-th moment c_i. We formulate the following theorem.

Theorem 2.2 (Characterisation of solution)**.** *If $\mathcal{F} = \mathcal{F}(c_1, c_i)$ then problem (3) has a unique solution $F^\star$ given as*

$$F^\star(x) = \begin{cases} 0 & \textit{if } x < \max\left(\left(\frac{g'(1)-\eta_1}{\eta_i}\right)^{\frac{1}{i-1}}, 0\right), \\ 1 - (g')^{-1}\left(\eta_1 + \eta_i x^{i-1}\right) & \textit{else}, \\ 1 & \textit{if } x \geqslant \min\left(\left(\frac{g'(0)-\eta_1}{\eta_i}\right)^{\frac{1}{i-1}}, 1\right), \end{cases} \tag{7}$$

where η_1, η_i are such that $\eta_i > 0$, $\eta_1 \in (g'(1) - \eta_i, g'(0))$ and

$$\int_0^1 x dF^\star(x) = c_1 \quad \textit{and} \quad \int_0^1 x^i dF^\star(x) = c_i. \tag{8}$$

Pseudocode[3] for finding the solution (η_1, η_i) for given moments (c_1, c_i) is given in Algorithm 1. First, two functions are defined that compute the moments (c_1, c_i) of $F_{\boldsymbol{\eta}}$, given (η_1, η_i). Then a nonlinear system is solved to find the values of (η_1, η_i) for which the moments are equal to (c_1, c_i). Due to Theorem 2.2 the solution is then found.

Algorithm 1 Solve for (η_1, η_i)

mu1 $\leftarrow$ function(eta1, etai)
mui $\leftarrow$ function(eta1, etai)
diffMoments $\leftarrow$ function(eta1, etai) {
    return (mu1(eta1, etai) − c1, mui(eta1, etai) − ci)
}
(eta1, etai) = solve.system(diffMoments)

Solving problem (3) for a general choice of $\mathcal{F}$ appears very difficult. However, from Theorem 2.2 one also obtains a simple upper bound on the distorted expectation when $\mathcal{F}$ contains all distribution functions with given mean and $k-1$ given higher order moments; we write $\mathcal{F} = \mathcal{F}(c_1, (c_i)_{i \in \mathcal{I}\setminus\{1\}})$

Corollary 2.3. *Let $F^\star$ be the optimal solution to* (3) *when $\mathcal{F} = \mathcal{F}(c_1, (c_i)_{i\in\mathcal{I}\setminus\{1\}})$ and denote by $F^\star_{c_1,c_i}$ the solution to* (3) *when optimizing over the set $\mathcal{F}(c_1, c_i)$, $i \in \mathcal{I}\setminus\{1\}$. Then*

$$H_g(F^\star) \leqslant \min_{i \in \mathcal{I}\setminus\{1\}} H_g(F^\star_{c_1,c_i}). \tag{9}$$

[3] R code is available on `https://github.com/cdries/ConcaveDistortionRM`. All numerical results in this paper can be replicated by the code provided.

Remark 2.4. *The bound obtained in Corollary 2.3 cannot be expected to be sharp. To show this, consider the distortion function* $g(x) = 1-(1-x)^2$ *(see also Figure 1a) and consider the optimization problem (3) over* $\mathcal{F}(c_1, c_2)$ *and* $\mathcal{F}(c_1, c_3)$ *with* $c_1 = 0.50$, $c_2 = 0.33$ *and* $c_3 = 0.24$. *The maximizing distribution functions can be obtained numerically using Algorithm 1, and are given as*

$$F^*_{c_1,c_2}(x) = \begin{cases} 0 & \text{if } x < 0.0101 \\ -0.0103 + 1.0206x & \text{if } 0.0101 \leqslant x < 0.9899 \\ 1 & \text{if } x \geqslant 0.9899, \end{cases} \tag{10}$$

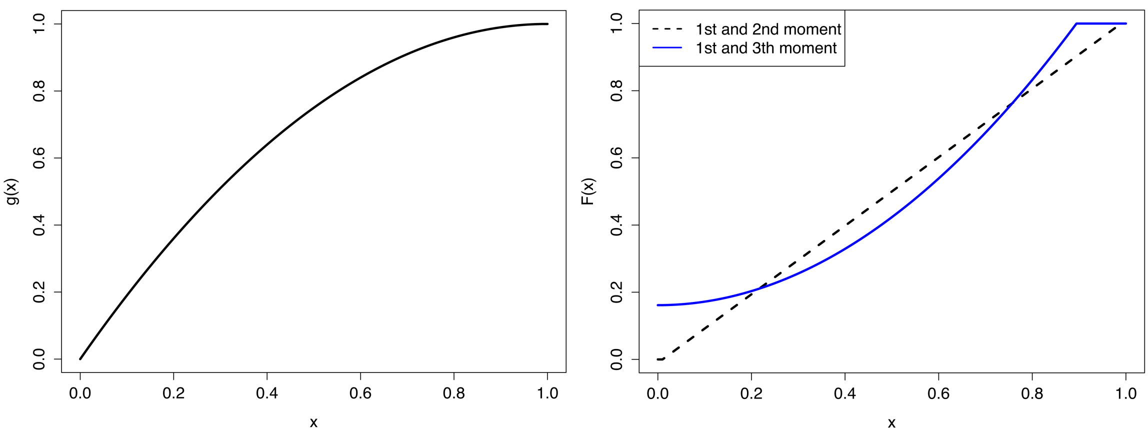
and

$$F^*_{c_1,c_3}(x) = \begin{cases} 0 & \text{if } x < 0 \\ 0.1615 + 1.0482x^2 & \text{if } 0 \leqslant x < 0.8944 \\ 1 & \text{if } x \geqslant 0.8944, \end{cases} \tag{11}$$

respectively; see Figure 1b. The values of the distorted expectation $H_g(F^*_{c_1,c_2})$ *and* $H_g(F^*_{c_1,c_3})$ *are equal to* 0.6633 *and* 0.6646, *respectively. However, since the first three moments of these distribution functions are respectively* $(0.5000, 0.3300, 0.2450)$ *and* $(0.5000, 0.3354, 0.2400)$, *both distribution functions are not even admissible with respect to the the optimization problem (3) over the set* $M(c_1, c_2, c_3)$. *Hence* F^* *can not coincide with* $F^*_{c_1,c_2}$ *nor* $F^*_{c_1,c_3}$ *and* $H_g(F^*)$ *is strictly lower than* $H_g(F^*_{c_1,c_2})$ *and* $H_g(F^*_{c_1,c_3})$.

Figure 1: Non sharpness of the bound obtained in Corollary 2.3.

(a) Distortion function (b) Maximizing distribution functions



3 Examples

In this section, we use Algorithm 1 to numerically determine the maximizing distribution function for several concave distortion risk measures. In addition, we also provide an analytical solution when the distortion function used is given by $g(x) = 1-(1-x)^2$ and the mean and variance of the distribution function are fixed.

3.1 Maximizing distribution functions for some concave distortion risk measures

We consider the power distortion function ($g(x) = x^\alpha$) with $\alpha = 0.5, 0.2, 0.1$, the dual-power distortion function ($g(x) = 1-(1-x)^\beta$) with $\beta = 2, 5, 10$, and the Wang distortion function ($g(x) = \Phi(\Phi^{-1}(x) + \Phi^{-1}(p))$) with $p = 0.8, 0.9, 0.95$. For the sake of example, we consider the moments that arise from the uniform distribution function on $[0, 1]$, i.e. $c_1 = 1/2$, $c_2 = 1/3$, $c_3 = 1/4$ and $c_4 = 1/5$.

For $\mathcal{F} = \mathcal{F}(c_1, c_2)$, $\mathcal{F} = \mathcal{F}(c_1, c_3)$ and $\mathcal{F} = \mathcal{F}(c_1, c_4)$, we determine the maximizing distribution functions numerically and display the distortion values in Table 1.

Table 1: Maximum value $H_g(F^\star)$ for several choices of distortion functions and moment spaces.

	$H_g(F^\star)$		
Distortion function	$\mathcal{F}_{(c_1,c_2)}$	$\mathcal{F}_{(c_1,c_3)}$	$\mathcal{F}_{(c_1,c_4)}$
power ($\alpha = 0.5$)	0.6754	0.6711	0.6693
power ($\alpha = 0.2$)	0.8450	0.8407	0.8382
power ($\alpha = 0.1$)	0.9175	0.9148	0.9130
dual-power ($\beta = 2$)	0.6667	0.6714	0.6782
dual-power ($\beta = 5$)	0.8660	0.8472	0.8366
dual-power ($\beta = 10$)	0.9686	0.9540	0.9404
Wang ($p = 0.80$)	0.7330	0.7276	0.7273
Wang ($p = 0.90$)	0.8360	0.8270	0.8230
Wang ($p = 0.95$)	0.9012	0.8923	0.8866

To obtain more insight in the maximizing distribution functions and their link with the distortion functions, we further focus on the power distortion ($g(x) = x^\alpha$) with parameter $\alpha = 0.2$, the dual-power distortion ($g(x) = 1 - (1-x)^\beta$) with parameter $\beta = 5$ and the Wang distortion ($g(x) = \Phi(\Phi^{-1}(x) + \Phi^{-1}(p))$) with $p = 0.8$. These distortion functions are displayed in Figure 2a. Figures 2b, 2c and 2d show the corresponding maximizing distribution functions for some of the cases considered in Table 1. Interestingly, unlike in the case of VaR and TVaR where maximizing distribution functions are discrete, the maximizing distribution functions are either continuous or appear as a mixture of a continuous and a discrete distribution function.

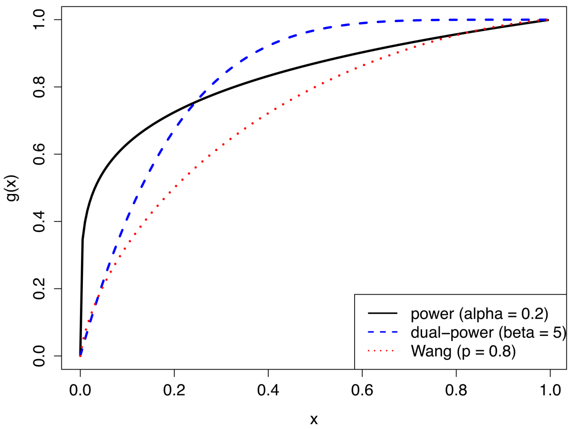
3.2 Analytic solution for the dual-power distortion risk measure

Moment spaces for distribution functions with a compact support are compact, see e.g. Karlin and Shapley (1953). However, the set E of all vectors $\boldsymbol{\eta}$ that are obtained by solving problem (3) for all possible moments, is not necessarily compact. In general, it does not seem possible to obtain an analytical description of the maximizing distribution function. In the case of the dual-power distortion function,

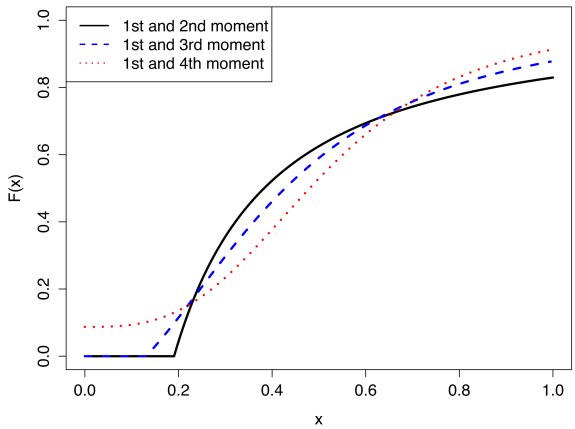
$$g(x) = 1 - (1-x)^2, \tag{12}$$

Figure 2: Maximizing distribution functions under different distortion risk measures.

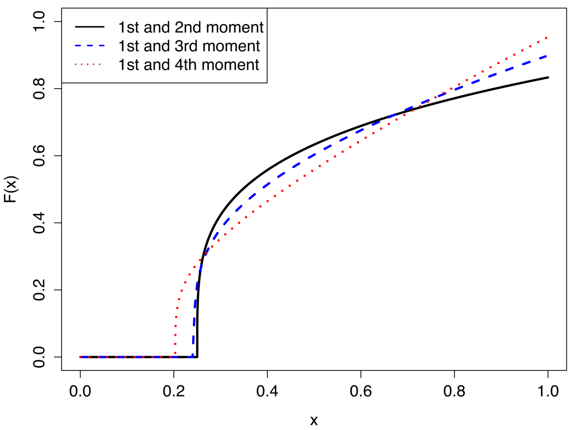
(a) Distortion functions



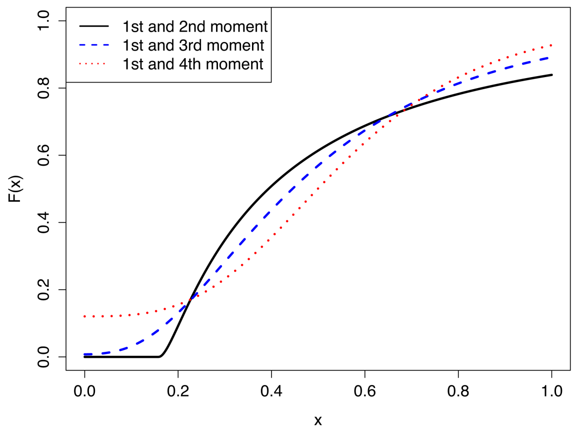
(b) Maximizing distribution functions under the dual-power distortion function ($\alpha = 0.2$)



(c) Maximizing distribution functions under the dual-power distortion function ($\beta = 5$)



(d) Maximizing distribution functions under the Wang distortion function ($p = 0.8$)



it is however possible to explicitly describe $\boldsymbol{\eta} = (\eta_1, \eta_2)'$ as a function of (c_1, c_2), see Figure 3b.

Consider the first two moments c_1 and c_2. The set of pairs (c_1, c_2) such that at least one distribution function with those moments exists, is given by the set (Karlin and Shapley (1953))

$$N = \{(x, y) | x \in [0, 1]; x^2 \leqslant y \leqslant x\}. \tag{13}$$

According to Theorem 2.2, the maximizing distribution function is of the following form:

$$F_{\boldsymbol{\eta}}(x) = \begin{cases} 0 & \text{if } x < \max\left(0, \frac{-\eta_1}{\eta_2}\right), \\ \frac{\eta_1 + \eta_2 x}{2} & \text{if else,} \\ 1 & \text{if } x \geqslant \min\left(1, \frac{2-\eta_1}{\eta_2}\right), \end{cases} \tag{14}$$

with $\eta_2 > 0$. This distribution function is uniform on a certain interval $(a, b), 0 \leqslant a < b \leqslant 1$, with possible mass points in 0 and 1.

The moment space N can be divided into four parts, the sets N_1, N_2, N_3 and N_4 defined by

$$\begin{aligned}
N_1 &= \left\{ (c_1, c_2) \in N \quad | \quad c_2 \geqslant \max\left(\frac{1}{3}(4c_1 - 1), \frac{2}{3}c_1\right) \right\}, \\
N_2 &= \left\{ (c_1, c_2) \in N \quad | \quad \frac{1}{3}(4c_1^2 - 2c_1 + 1) \leqslant c_2 < \frac{1}{3}(4c_1 - 1) \right\}, \\
N_3 &= \left\{ (c_1, c_2) \in N \quad | \quad \frac{4}{3}c_1^2 \leqslant c_2 < \frac{2}{3}c_1 \right\}, \\
N_4 &= \left\{ (c_1, c_2) \in N \quad | \quad c_2 < \min\left(\frac{4}{3}c_1^2, \frac{1}{3}(4c_1^2 - 2c_1 + 1)\right) \right\}.
\end{aligned} \tag{15}$$

These regions are shown in Figure 3a.

In each of the four regions, the values of (η_1, η_2) in $F_{\boldsymbol{\eta}}$ can be determined as a function of (c_1, c_2) as follows:

$$(\eta_1, \eta_2) = \begin{cases} (-8c_1 + 6c_2 + 2, 12(c_1 - c_2)) & \text{if } (c_1, c_2) \in N_1, \\ \left(\frac{8(1-c_1)^2(1+3c_2-4c_1)}{9(1+c_2-2c_1)^2}, \frac{16(1-c_1)^3}{9(1+c_2-2c_1)^2}\right) & \text{if } (c_1, c_2) \in N_2, \\ \left(2 - \frac{8c_1^2}{3c_2}, \frac{16c_1^3}{9c_2^2}\right) & \text{if } (c_1, c_2) \in N_3, \\ \left(1 - \frac{c_1}{\sqrt{3(c_2-c_1^2)}}, \frac{1}{\sqrt{3(c_2-c_1^2)}}\right) & \text{if } (c_1, c_2) \in N_4. \end{cases} \tag{16}$$

This transforms the sets N_1, N_2, N_3 and N_4 to the sets E_1, E_2, E_3 and E_4 defined by

$$\begin{aligned}
E_1 &= \{(\eta_1, \eta_2) \quad | \quad \eta_1 \in [0, 2); \quad 0 < \eta_2 \leqslant 2 - \eta_1\}, \\
E_2 &= \{(\eta_1, \eta_2) \quad | \quad \eta_1 \in (-\infty, 0); \quad -\eta_1 < \eta_2 \leqslant 2 - \eta_1\}, \\
E_3 &= \{(\eta_1, \eta_2) \quad | \quad \eta_1 \in [0, 2); \quad 2 - \eta_1 < \eta_2\}, \\
E_4 &= \{(\eta_1, \eta_2) \quad | \quad \eta_1 \in (-\infty, 0); \quad 2 - \eta_1 < \eta_2\}
\end{aligned} \tag{17}$$

and shown in Figure 3b. The boundaries of the set E are exactly as prescribed in Theorem 2.2.

Note that the point $(c_1, c_2) = (1/2, 1/3)$ lies on the intersection of the four different regions in the moment space. Hence, the corresponding (η_1, η_2)-values for the distribution function maximizing the distorted expectation under g as in (12) are $(\eta_1, \eta_2) = (0, 2)$ and incidentally,

$$F_{(0,2)}(x) = x, \qquad x \in [0, 1], \tag{18}$$

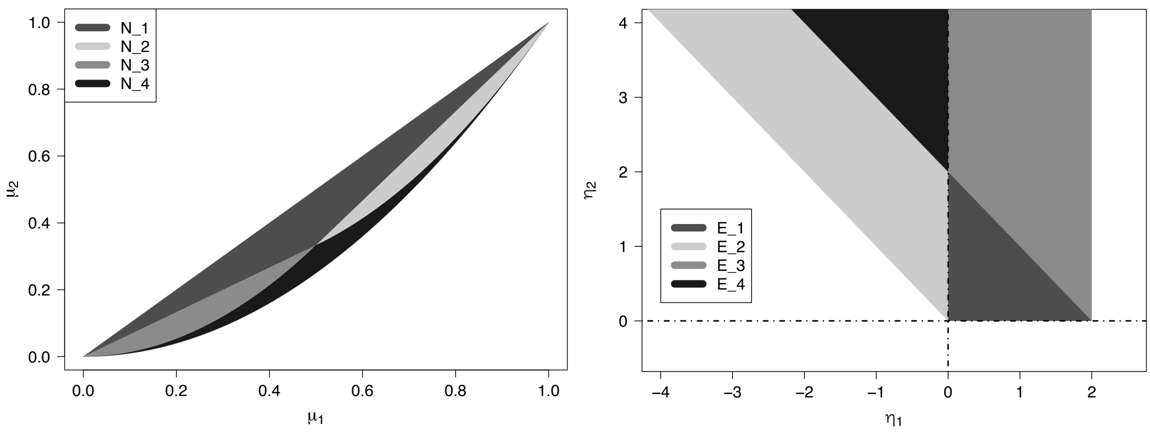
is the distribution function maximizing the distorted expectation (3). The distorted expectation equals $H_g(F_{(0,2)}) = 2/3$.

4 Proofs

In the first part of this section we prove some useful lemmas based on adaptations and extensions of the results presented in Rustagi (1957, 1976). The proofs for Theorem

Figure 3: Moment space and corresponding set for (η_1, η_2).

(a) Moment space N (b) Set E of possible (η_1, η_2)-values



2.1 and Theorem 2.2 are presented thereafter.

4.1 Some useful lemmas

In this section, let φ be a strictly convex and twice differentiable bounded function on $[0, 1]$. The results concern a solution $F^\star$ of the optimization problem

$$F^\star = \arg\min_{F\in\mathcal{F}} \int_0^1 \varphi(F(x))dx, \tag{19}$$

where $\mathcal{F}$ is given in (1).

Lemma 4.1 (Unicity)**.** *A solution $F^\star$ to* (19) *exists and is unique.*

Proof. The existence of a solution follows from Lemma 3.1 and Lemma 3.2 in Rustagi (1957). Specifically, let $\mathcal{I} = \{i_1, \ldots, i_k\}$ and observe that $\int_0^1 x^{i_j} dF(x) = c_{i_j}$ if and only if $\int_0^1 x^{i_j - 1} F(x)dx = d_{i_j}$ for the appropriate d_{i_j}. Consider the transformation

$$\begin{aligned} &T : \mathcal{F} \to \boldsymbol{R} : \\ &T(F) = \left(\int_0^1 \varphi(F(x))dx, \int_0^1 x^{i_1-1}F(x)dx, \ldots, \int_0^1 x^{i_k-1}F(x)dx\right). \end{aligned} \tag{20}$$

This transformation is continuous and linear in F and since $\mathcal{F}$ is convex and compact in the topology of convergence in distribution it maps $\mathcal{F}$ into a convex and compact set. The imposed moment restrictions on $F(x)$ thus yield a non-empty subset that is also bounded and closed and hence a solution to (19) exists.

To prove uniqueness assume that $F_0 \neq F_1$ are two admissible solutions. Let

$$M = \int_0^1 \varphi(F_0(x))dx = \int_0^1 \varphi(F_1(x))dx. \tag{21}$$

Since ϕ is strictly convex, it holds for any $\lambda \in (0,1)$ that

$$\begin{aligned}\int_0^1 &\varphi(\lambda F_0(x) + (1-\lambda)F_1(x))dx \\ &< \lambda \int_0^1 \varphi(F_0(x))dx + (1-\lambda)\int_0^1 \varphi(F_1(x))dx = M,\end{aligned} \tag{22}$$

which is a contradiction; see also page 100 in Rustagi (1976). $\square$

Define φ_y the derivative of the function φ.

Lemma 4.2 (Linearisation)**.** *If $F^\star$ solves* (19)*, then there exist $(\eta_i)_{i\in\mathcal{I}}$ such that $F^\star$ minimizes over all $F \in \mathcal{F}$ the function*

$$\int_0^1 \left[\varphi_y(F^\star(x)) + \sum_{i\in\mathcal{I}} \eta_i x^{i-1}\right] F(x)dx. \tag{23}$$

Proof. Let $F^\star$ be the solution to (19). Lemma 5.3.1 in Rustagi (1976) proves that $F^\star$ minimizes (19) if and only if

$$\int_0^1 \varphi_y(F^\star(x))F(x)dx \geqslant \int_0^1 \varphi_y(F^\star(x))F^\star(x)dx, \qquad F \in \mathcal{F}. \tag{24}$$

Denote $\mathcal{I} = \{i_1, \ldots, i_k\}$. Consider the set $\mathbf{\Gamma}$ of points $(\zeta_0, \zeta_1, \ldots, \zeta_k)$ equal to

$$\left(\int_0^1 \varphi_y(F^\star(x))F(x)dx, \int_0^1 x^{i_1-1}F(x)dx, \ldots, \int_0^1 x^{i_k-1}F(x)dx\right), \quad F \in \mathcal{F}, \tag{25}$$

where the second until last coordinates are determined by the moments of the distribution function F, obtained by integration by parts. The set $\mathbf{\Gamma}$ is closed, bounded and convex in $k+1$ dimensions. We show this based on Theorem 7.2 in Karlin and Shapley (1953) and page 313 in Rustagi (1957). Consider the transformation

$$\begin{aligned}&T : \mathcal{F} \to \mathbf{\Gamma} : \\ &\quad T(F) = \left(\int_0^1 \varphi_y(F^\star(x))F(x)dx, \int_0^1 x^{i_1-1}F(x)dx, \ldots, \int_0^1 x^{i_k-1}F(x)dx\right).\end{aligned} \tag{26}$$

The transformation T is continuous and linear in F. Since $\mathcal{F}$ is convex and compact in the topology of convergence in distribution, also its image is convex and compact.

Because the set $\mathbf{\Gamma}$ is closed, bounded and convex, there exists a boundary point $(z_0, \ldots, z_k)$ of $\mathbf{\Gamma}$ such that $F^\star$ corresponds to this boundary point. Therefore, there exists a supporting hyperplane of $\mathbf{\Gamma}$ at $(z_0, \ldots, z_k)$, i.e. there exist $\eta_{i_1}, \ldots, \eta_{i_k}, \eta$ such that

$$\eta_0 z_0 + \eta_{i_1} z_1 + \cdots + \eta_{i_k} z_k + \eta = 0, \tag{27}$$

and

$$\eta_0 \zeta_0 + \eta_{i_1} \zeta_1 + \cdots + \eta_{i_k} \zeta_k + \eta \geqslant 0, \qquad \forall \boldsymbol{\zeta} \in \mathbf{\Gamma}. \tag{28}$$

Hence, it holds that

$$\eta_0(\zeta_0 - z_0) + \cdots + \eta_{i_k}(\zeta_k - z_k) \geqslant 0, \qquad \forall \boldsymbol{\zeta} \in \boldsymbol{\Gamma}. \tag{29}$$

We now show that $\eta_0 > 0$ by eliminating the other possibilities. Let

$$\boldsymbol{\Gamma}^\star = \{(\zeta_0^\star, \zeta_1, \ldots, \zeta_k) | \zeta_0^\star \geqslant \zeta_0; (\zeta_0, \zeta_1, \ldots, \zeta_k) \in \boldsymbol{\Gamma}\}. \tag{30}$$

The set $\boldsymbol{\Gamma}^\star$ is convex and $\boldsymbol{\Gamma} \subseteq \boldsymbol{\Gamma}^\star$. The value z_0 is the minimum of ζ_0 subject to $\zeta_1 = z_1, \ldots, \zeta_k = z_k$. Hence $(z_0, \ldots, z_k)$ is also a minimum point of $\boldsymbol{\Gamma}^\star$ and thus a boundary point of $\boldsymbol{\Gamma}^\star$. Therefore there exist $(\eta_0, \ldots, \eta_{i_k}) \neq (0, \ldots, 0)$ such that

$$\eta_0(\zeta_0^\star - z_0) + \cdots + \eta_{i_k}(\zeta_k - z_k) \geqslant 0 \qquad \forall (\zeta_0^\star, \zeta_1, \ldots, \zeta_k) \in \boldsymbol{\Gamma}^\star. \tag{31}$$

Hence, since $\boldsymbol{\Gamma} \subseteq \boldsymbol{\Gamma}^\star$,

$$\eta_0(\zeta_0 - z_0) + \cdots + \eta_{i_k}(\zeta_k - z_k) \geqslant 0 \qquad \forall (\zeta_0, \zeta_1, \ldots, \zeta_k) \in \boldsymbol{\Gamma}. \tag{32}$$

Suppose $\eta_0 < 0$. Since $(z_0, \ldots, z_k)$ is a minimum point of $\boldsymbol{\Gamma}^\star$, there exists some $\alpha > 0$ such that $(z_0 + \alpha, z_1, \ldots, z_k) \in \boldsymbol{\Gamma}^\star$. By Equation (31), $\eta_0\alpha \geqslant 0$ would hold, which would imply $\eta_0 = 0$. Next, suppose that $\eta_0 = 0$. This corresponds to the boundary points of the set $\boldsymbol{\Gamma}$ where the supporting hyperplanes are parallel to the ζ_0-axis, and hence $(z_1, \ldots, z_k)$ corresponds to the boundary of the projection of $\boldsymbol{\Gamma}$ on the $(\zeta_1, \ldots, \zeta_k)$ hyperplane. But the conditions on $\mathcal{F}$ are such that the given point $(z_1, \ldots, z_k)$ will be interior to the projection set, and hence $\eta_0 \neq 0$.

Thus $\eta_0 > 0$ and we can normalize it to be one, obtaining for any $F \in \mathcal{F}$,

$$\begin{aligned} &\int_0^1 \left[\varphi_y(F^\star(x)) + \sum_{i \in \mathcal{I}} \eta_i x^{i-1}\right] F(x) dx \\ &\qquad \geqslant \int_0^1 \left[\varphi_y(F^\star(x)) + \sum_{i \in \mathcal{I}} \eta_i x^{i-1}\right] F^\star(x) dx. \end{aligned} \tag{33}$$

Equivalently, $F^\star$ minimizes (23). □

Next we provide necessary conditions that a solution to (23) should satisfy. In the next lemma, specifically define

$$\varphi : [0,1] \to [0,1] : \varphi(y) = 1 - g(1-y) \tag{34}$$

for a chosen distortion function g satisfying the required conditions.

Lemma 4.3 (Necessary conditions)**.** *Consider the optimization problem*

$$F^\star = \arg\min_{F \in \mathcal{F}} \int_0^1 \left[\varphi_y(F^\star(x)) - \sum_{i \in \mathcal{I}} \eta_i x^{i-1}\right] F(x) dx. \tag{35}$$

Then $F^\star$ is continuous on $(0,1)$ and on intervals where $F^\star$ is not constant it coincides with $F_{\boldsymbol{\eta}}$ defined as

$$F_{\boldsymbol{\eta}}(x) = \begin{cases} 0 & \textit{if } z < g'(1), \\ 1 - (g')^{-1}(z) & \textit{else}, \\ 1 & \textit{if } z \geqslant g'(0), \end{cases} \tag{36}$$

with $z = \sum_{i\in\mathcal{I}} \eta_i x^{i-1}$ and $(\eta_i)_{i\in\mathcal{I}}$ are such that

$$\int_0^1 x^i dF^\star(x) = c_i, \qquad i \in \mathcal{I}, \tag{37}$$

and

$$\int_a^b \left[g'(1 - F^\star(x)) - \sum_{i\in\mathcal{I}} \eta_i x^{i-1} \right] dx = 0 \tag{38}$$

on intervals $[a,b] = [F^{\star-}(c), F^{\star+}(c)]$ for all $c \in (0,1)$.

Proof. Properties of the solution $F^\star$ will be described in terms of $G_{\boldsymbol{\eta}}(x)$ solving

$$\varphi_y(G_{\boldsymbol{\eta}}(x)) - \sum_{i\in\mathcal{I}} \eta_i x^{i-1} = 0. \tag{39}$$

Because φ_y is differentiable, this equation readily inverts on some set I to an expression for $G_{\boldsymbol{\eta}}(x)$ in terms of the original distortion function g:

$$G_{\boldsymbol{\eta}} : I \to \mathbb{R} : G_{\boldsymbol{\eta}}(x) = 1 - (g')^{-1}\left(\sum_{i\in\mathcal{I}} \eta_i x^{i-1} \right), \tag{40}$$

with I defined as

$$I = \left\{ x \in [0,1] \middle| \sum_{i\in\mathcal{I}} \eta_i x^{i-1} \in [g'(1), g'(0)) \right\}. \tag{41}$$

We extend the function $G_{\boldsymbol{\eta}}$ to the interval $[0,1]$ in straightforward fashion:

$$F_{\boldsymbol{\eta}}(x) = \begin{cases} 0 & \text{if } z < g'(1), \\ 1 - (g')^{-1}(z) & \text{if } x \in I, \\ 1 & \text{if } z \geqslant g'(0), \end{cases} \tag{42}$$

with $z = \sum_{i\in\mathcal{I}} \eta_i x^{i-1}$

Next, some restrictions on how to transform $F_{\boldsymbol{\eta}}$ into a distribution function are derived. Define the function

$$A : (0,1) \to \mathbb{R} : A(x) = \varphi_y(F^\star(x)) - \sum_{i\in\mathcal{I}} \eta_i x^{i-1} \tag{43}$$

and the set

$$\mathcal{S} = \{x \in (0,1) | A(x) \neq 0\}. \tag{44}$$

Theorem 5.1 on page 316 of Rustagi (1957) states that if $F^\star$ minimizes (23), then the set $\mathcal{S}$ has $F^\star$-measure zero, i.e. $F^\star$ is constant when $A \neq 0$. Hence, $F^\star$ coincides with $F_{\boldsymbol{\eta}}$ on $(0,1)$ when $F^\star$ is not constant.

A corollary on page 316 Rustagi (1957) gives restrictions on $F^\star$ where it is constant. On intervals $[a,b] = [F^-(c), F^+(c)]$ where $c \in (0,1)$, it holds that

$$\int_a^b A(x)dx = 0. \tag{45}$$

Finally, the continuity of $F^\star$ on $(0,1)$ follows from its right-continuity and the fact that it has no jumps, as proven in Theorem 5.2 in Rustagi (1957). This concludes the proof. □

4.2 Proof of Theorem 2.1

First, we rewrite the maximization problem (3) as a minimization problem and consider

$$F^\star = \arg\min_{F \in \mathcal{F}} \int_0^1 (1 - g(1 - F(x)))\, dx. \tag{46}$$

Under the assumptions of the theorem, the function $\varphi(x)$ defined as

$$\varphi : [0,1] \to [0,1] : \varphi(y) = 1 - g(1-y), \tag{47}$$

is strictly convex, bounded and twice differentiable. Thus, by Lemma 4.1, there exists a unique $F^\star$ maximizing (3).

By Lemma 4.2 it holds that $F^\star$ minimizes over all $F \in \mathcal{F}$ the function

$$\int_0^1 \left[\varphi_y(F^\star(x)) + \sum_{i \in \mathcal{I}} \eta_i x^{i-1} \right] F(x)dx. \tag{48}$$

Since this optimization problem is similar to the one in Lemma 4.3, the necessary conditions in Theorem 2.1 are shown.

4.3 Proof of Theorem 2.2

Under the moment conditions it follows that $F_{\boldsymbol{\eta}}$ as stated in Theorem 2.1 is monotonic (as the composition of an increasing function and a linear one). Furthermore, $F_{\boldsymbol{\eta}}$ has then to be increasing, since otherwise $F^\star$ would be degenerated and thus not admissible ($\mathcal{F}$ contains at least two elements and the degenerated distribution function is unique with respect to its moment sequence). Hence, since $F^\star$ is continuous on $(0,1)$, it holds that the optimal solution $F^\star$ should be of following form:

$$F_{\boldsymbol{\eta}}(x) = \begin{cases} 0 & \text{if } x < \max\left(\left(\frac{g'(1)-\eta_1}{\eta_i}\right)^{\frac{1}{i-1}}, 0\right), \\ 1 - (g')^{-1}\left(\eta_1 + \eta_i x^{i-1}\right) & \text{else}, \\ 1 & \text{if } x \geqslant \min\left(\left(\frac{g'(0)-\eta_1}{\eta_i}\right)^{\frac{1}{i-1}}, 1\right), \end{cases} \tag{49}$$

for some $\boldsymbol{\eta} := (\eta_1, \eta_i)$ such that $\eta_i > 0$ and $\eta_1 \in (g'(1) - \eta_i, g'(0))$ and

$$\mu_1(\boldsymbol{\eta}) = \int_0^1 x dF_{\boldsymbol{\eta}}(x) = c_1 \quad \text{and} \quad \mu_i(\boldsymbol{\eta}) = \int_0^1 x^i dF_{\boldsymbol{\eta}}(x) = c_i. \tag{50}$$

We aim at showing that there exists only one $\boldsymbol{\eta}^\star = (\eta_1^\star, \eta_i^\star)$ such that $F_{\boldsymbol{\eta}^\star}$ satisfies the moment conditions (50).

First, we show that for any $\eta_i > 0$ there exists some η_1 such that $\mu_1(\eta_1, \eta_i) = c_1$. For any $\eta_i > 0$, the function $\mu_1(\eta_1, \eta_i)$ is strictly decreasing in η_1 on $(g'(1) - \eta_i, g'(0))$. Since $\lim_{\eta_1 \to g'(1)-\eta_i} \mu_1(\eta_1, \eta_i) = 1$ and $\lim_{\eta_1 \to g'(0)} \mu_1(\eta_1, \eta_i) = 0$, the full domain of possible mean values for any distribution function F on $[0,1]$ can be reached by varying η_1. Hence, there exists unique η_1 such that $\mu_1(\eta_1, \eta_i) = c_1$, write $\eta_1 = \eta_1((\eta_i))$.

Next, we show that there exists a unique $\eta_i^\star > 0$ such that $\mu_i(\eta_1(\eta_i), \eta_i) = c_i$. To this end, consider the function

$$\widetilde{\mu}_i : (0,\infty) \to [0,1] : \widetilde{\mu}_i(\eta_i) = \int_0^1 x^{i-1} F_{(\eta_1(\eta_i), \eta_i)}(x) dx. \tag{51}$$

. The derivative of $\widetilde{\mu}_i$ with respect to η_i (using expression (53)) is given as

$$\frac{d\widetilde{\mu}_i(\eta_i)}{d\eta_i} = \int_0^1 x^{2i-2} H'(\eta_1 + \eta_i x^{i-1}) dx + \left(\int_0^1 x^{i-1} H'(\eta_1 + \eta_i x^{i-1}) dx\right) \frac{d\eta_1(\eta_i)}{d\eta_i}, \tag{52}$$

where H' is the derivative of H on $(0,1)$ a.e., where H is defined as

$$H(y) = \begin{cases} 0 & \text{if } y < \max\left(g'(1), 0\right), \\ 1 - (g')^{-1}(y) & \text{else}, \\ 1 & \text{if } y \geqslant \min\left(g'(0), 1\right), \end{cases} \tag{53}$$

It holds that the derivative $d\eta_1(\eta_i)/d\eta_i$ equals

$$\frac{d\eta_1(\eta_i)}{d\eta_i} = -\left(\int_0^1 x^{i-1} H'(\eta_1 + \eta_i x^{i-1}) dx\right) \Big/ \left(\int_0^1 H'(\eta_1 + \eta_i x^{i-1}) dx\right). \tag{54}$$

The latter can be seen by setting the total derivative $d\widetilde{\mu}_1$ equal to zero, since we keep the mean fixed when varying η_1 and η_i accordingly.

Notice that $H'(\eta_1+\eta_i x^{i-1})/\int_0^1 H'(\eta_1+\eta_i x^{i-1})dx$ is a density on $(0,1)$. Because H' is not degenerate, the variance of X^{i-1} with respect to this density is strictly positive and hence

$$\frac{d\widetilde{\mu}_i(\eta_i)}{d\eta_i} > 0. \tag{55}$$

Since integration by parts yields the relation $i\widetilde{\mu}_i(\eta_i) = 1 - \mu_i(\eta_1(\eta_i), \eta_i)$, we thus also have that

$$\frac{d\mu_i(\eta_1(\eta_i), \eta_i)}{d\eta_i} < 0. \tag{56}$$

Thus, if $\eta^\star$ exists, it is unique due to strict monotonicity.

Consider now the limits

$$\lim_{\eta_i \to 0} \int_0^1 x^i dF_{(\eta_1(\eta_i), \eta_i)}(x) = c_1 \quad \text{and} \quad \lim_{\eta_i \to \infty} \int_0^1 x^i dF_{(\eta_1(\eta_i), \eta_i)}(x) = c_1^i, \tag{57}$$

which show that the full domain of possible values for the i-th moment can be reached by varying η_i implying that $\eta_i^\star$ exists. Hence, $\boldsymbol{\eta}^\star = (\eta_1^\star, \eta_i^\star)$ exists and is unique. Thus $F^\star = F_{\boldsymbol{\eta}^\star}$.

References

P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.

J. Belles-Sampera, M. Guillén, and M. Santolino. Beyond Value-at-Risk: GlueVaR distortion risk measures. *Risk Analysis*, 34(1):121–134, 2014.

C. Bernard, L. Rüschendorf, and S. Vanduffel. Value-at-Risk bounds with variance constraints. *Journal of Risk and Insurance*, 2015.

C. Bernard, M. Denuit, and S. Vanduffel. Measuring portfolio risk under partial dependence information. *Journal of Risk and Insurance*, 2016.

C. Bernard, L. Rüschendorf, S. Vanduffel, and R. Wang. Risk bounds for factor models. *Finance and Stochastics*, Forthcoming, 2017a.

C. Bernard, L. Rüschendorf, S. Vanduffel, and J. Yao. How robust is the Value-at-Risk of credit risk portfolios? *The European Journal of Finance*, 23(6):507–534, 2017b.

J. Danielsson, P. Embrechts, C. Goodhart, C. Keating, F. Muennich, O. Renault, H. S. Shin, et al. An academic response to Basel II. *Special Paper-LSE Financial Markets Group*, 2001.

A. De Schepper and B. Heijnen. How to estimate the value at risk under incomplete information. *Journal of Computational and Applied Mathematics*, 233(9):2213–2226, 2010.

M. Denuit, E. De Vylder, and C. Lefevre. Extremal generators and extremal distributions for the continuous s-convex stochastic orderings. *Insurance: Mathematics and Economics*, 24(3):201–217, 1999a.

M. Denuit, C. Genest, and É. Marceau. Stochastic bounds on sums of dependent risks. *Insurance: Mathematics and Economics*, 25(1):85–104, 1999b.

J. Dhaene, S. Vanduffel, M. J. Goovaerts, R. Kaas, Q. Tang, and D. Vyncke. Risk measures and comonotonicity: a review. *Stochastic models*, 22(4):573–606, 2006.

J. Dhaene, R. J. Laeven, S. Vanduffel, G. Darkiewicz, and M. J. Goovaerts. Can a coherent risk measure be too subadditive? *Journal of Risk and Insurance*, 75(2):365–386, 2008.

J. Dhaene, A. Tsanakas, E. A. Valdez, and S. Vanduffel. Optimal capital allocation principles. *Journal of Risk and Insurance*, 79(1):1–28, 2012.

P. Embrechts, G. Puccetti, and L. Rüschendorf. Model uncertainty and VaR aggregation. *Journal of Banking & Finance*, 37(8):2750–2764, 2013.

P. Embrechts, G. Puccetti, L. Rüschendorf, R. Wang, and A. Beleraj. An academic response to basel 3.5. *Risks*, 2(1):25–48, 2014.

P. Embrechts, B. Wang, and R. Wang. Aggregation-robustness and model uncertainty of regulatory risk measures. *Finance and Stochastics*, 19(4):763–790, 2015a.

P. Embrechts, R. Wang, et al. Seven proofs for the subadditivity of Expected Shortfall. *Dependence Modeling*, 3(1):126–140, 2015b.

M. Frittelli, M. Maggis, and I. Peri. Risk measures on p(r) and Value at Risk with probability/loss function. *Mathematical Finance*, 24(3):442–463, 2014.

W. Hürlimann. Analytical bounds for two Value-at-Risk functionals. *Astin Bulletin*, 32 (02):235–265, 2002.

W. Hürlimann. Extremal moment methods and stochastic orders. *Boletin de la Associacion Matematica Venezolana*, 15:153–301, 2008.

R. Kaas and M. Goovaerts. Best bounds for positive distributions with fixed moments. *Insurance: Mathematics and Economics*, 5(1):87–92, 1986.

S. Karlin and L. S. Shapley. *Geometry of moment spaces*. Number 12. American Mathematical Soc., 1953.

S. Kusuoka. On law invariant coherent risk measures. In *Advances in mathematical economics*, pages 83–95. Springer, 2001.

A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management: Concepts, techniques and tools*. Princeton university press, 2015.

G. Puccetti, L. Rüschendorf, and D. Manko. Var bounds for joint portfolios with dependence constraints. *Dependence Modeling*, 4(1), 2016.

G. Puccetti, L. Rüschendorf, D. Small, and S. Vanduffel. Reduction of Value-at-Risk bounds via independence and variance information. *Scandinavian Actuarial Journal*, 2017(3): 245–266, 2017.

L. Rüschendorf. Random variables with maximum sums. *Advances in Applied Probability*, pages 623–632, 1982.

L. Rüschendorf and J. Witting. VaR bounds in models with partial dependence information on subgroups. *Dependence Modeling*, 5(1):59–74, 2017.

J. S. Rustagi. On minimizing and maximizing a certain integral with statistical applications. *The Annals of Mathematical Statistics*, pages 309–328, 1957.

J. S. Rustagi. Variational methods in statistics. In *Variational methods in statistics*. Academic Press, 1976.

R. Tian. Moment problems with applications to value-at-risk and portfolio management. *PhD Dissertation, Georgia State University*, 2008.

B. Wang and R. Wang. The complete mixability and convex minimization problems with monotone marginal densities. *Journal of Multivariate Analysis*, 102(10):1344–1360, 2011.

R. Wang, L. Peng, and J. Yang. Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. *Finance and Stochastics*, 17(2):395–417, 2013.

J. L. Wirch and M. R. Hardy. A synthesis of risk measures for capital adequacy. *Insurance: mathematics and Economics*, 25(3):337–347, 1999.