



City Research Online

City, University of London Institutional Repository

Citation: Owadally, I. and Landsman, Z. (2013). A characterization of optimal portfolios under the tail mean-variance criterion. *Insurance: Mathematics and Economics*, 52(2), pp. 213-221. doi: 10.1016/j.insmatheco.2012.12.004

This is the submitted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/17092/>

Link to published version: <http://dx.doi.org/10.1016/j.insmatheco.2012.12.004>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

A Characterization of Optimal Portfolios under the Tail Mean-Variance Criterion

Iqbal Owadally^{a,*}, Zinoviy Landsman^{b,**}

^a*Faculty of Actuarial Science and Insurance, Cass Business School, City University London, 106 Bunhill Row, London EC1Y 8TZ, United Kingdom*

^b*Department of Statistics, University of Haifa, Mount Carmel, 31905 Haifa, Israel*

Abstract

The tail mean-variance model was recently introduced for use in risk management and portfolio choice, and involves a criterion that focuses on the risk of rare but large losses, which is particularly important when losses have heavy-tailed distributions. If returns or losses follow a multivariate elliptical distribution, the use of risk measures that satisfy certain well-known properties is equivalent to risk management in the classical mean-variance framework. The tail mean-variance criterion does not satisfy these properties, however, and the precise optimal solution typically requires the use of numerical methods. We use a convex optimization method and a mean-variance characterization to find an explicit and easily implementable solution for the tail mean-variance model. When a risk-free asset is available, the optimal portfolio is altered in a way that differs from the classical mean-variance setting. A complete solution to the optimal portfolio in the presence of a risk-free asset is also provided.

Keywords: Tail conditional expectation, Tail variance, Optimal portfolio selection, Quartic equation

JEL: G11, *Classification:* IM12 IE13 IB63

*Corresponding author. *Email:* m.i.owadally@city.ac.uk. *Phone:* +44 (0)20 7040 8478. *Fax:* +44 (0)20 7040 8572.

***Email:* landsman@stat.haifa.ac.il. *Phone:* +972 4 8249003. *Fax:* +972 4 8253849.

1. Introduction

In this paper, we consider portfolio selection under the tail mean-variance criterion that was introduced by Landsman (2010):

$$TMV(L) = \mathbb{E}[L \mid L > \text{VaR}_q(L)] + \lambda \text{Var}[L \mid L > \text{VaR}_q(L)]. \quad (1)$$

In the above, (i) $\lambda > 0$, (ii) L is a random loss, with a continuous distribution, on a portfolio, (iii) $\text{VaR}_q(L)$ is the value-at-risk on the portfolio and is defined as

$$\text{VaR}_q(L) = \inf(x \in \mathbb{R} : F_L(x) \geq q), \quad (2)$$

where $q \in (0, 1)$ and $F_L(x)$ is the cumulative distribution function of loss L . ($q = 0.95$ therefore corresponds to a 5% value-at-risk.) The objective is to find an optimal portfolio that minimizes the tail mean-variance criterion subject to a budget constraint.

Since portfolio return $R = -L$, the tail mean-variance criterion may be regarded as an analogue of the classical mean-variance criterion (see e.g. Panjer et al., 1998, p. 379):

$$MV(L) = \mathbb{E}L + \frac{1}{2}\tau\text{Var}L = -\mathbb{E}R + \frac{1}{2}\tau\text{Var}R, \quad (3)$$

with $\tau > 0$. This originates from the mean-variance portfolio theory of Markowitz (1952), of course. Unlike its classical counterpart, the tail mean-variance criterion focuses on the behaviour of the tail of portfolio returns through the q -quantile specified in the value-at-risk. This is of interest to portfolio managers whose clients may be concerned with portfolio performance in the event of extreme losses on capital markets.

We make three significant contributions in this paper. First, we use a convex optimization method to find an explicit and easily implementable solution for tail mean-variance optimization. We retain the assumption of joint elliptically distributed returns on securities but, unlike the solution of Landsman (2010), our solution is simple and avoids a sequence of matrix partitions and manipulations. The matrices may be of a large dimension for portfolios containing many hundred securities and involving a large variance-covariance matrix, so our solution has a considerable computational advantage. Secondly, our optimal solution is amenable to a simple interpretation. We provide a simple characterization

for the tail mean-variance optimal portfolio in terms of mean-variance efficiency. This facilitates comparison with optimal portfolios under other criteria, such as the mean-variance and value-at-risk criteria. Thirdly, we further extend the work of Landsman (2010) by the inclusion of risk-free lending and borrowing. A complete, closed-form optimal solution is provided.

It is convenient to introduce at this point some notation and assumptions used in the rest of the paper. \mathbb{R} , \mathbb{R}_+ and \mathbb{R}_{++} denote the sets of real numbers, real non-negative numbers and real positive numbers respectively. We assume that there are n risky securities with mean return $\boldsymbol{\mu} \in \mathbb{R}^n$ and variance-covariance matrix $\Sigma \in \mathbb{R}^{n \times n}$. Define $\mathbf{0}$ and $\mathbf{1}$ to be column vectors of zeros and ones respectively, of dimension n . As is usual, we assume that $\boldsymbol{\mu}$ is not collinear with $\mathbf{1}$, i.e. that securities do not all have the same mean return, and that Σ is a (symmetric) positive definite and non-singular matrix. Let $\mathbf{x} \in \mathbb{R}^n$ be the vector of proportions of wealth invested in a portfolio. Define \mathcal{P} as the set of feasible portfolios of risky securities only:

$$\mathcal{P} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{1}^T \mathbf{x} = 1\}. \quad (4)$$

We assume in the above that all wealth is invested. In section 6, a risk-free asset is included, in which case the proportion of wealth invested in the risk-free asset is $1 - \mathbf{1}^T \mathbf{x}$. There is no constraint other than the budget constraint that all wealth be invested and, in particular, short sales are allowed.

The plan of this paper is as follows. The motivation and literature related to this paper are described in section 2. In section 3, we briefly review some results on the tail conditional expectation, the tail variance, and on the tail mean-variance criterion. A characterization of the optimal portfolio under the tail mean-variance criterion using mean-variance efficiency is discussed in section 4. A novel solution for the tail mean-variance optimal portfolio is derived in section 5 when portfolios consist of risky securities only. It is illustrated with the help of a numerical example. When risk-free lending and borrowing are introduced, the tail mean-variance optimal portfolio is altered and this is detailed in section 6. A few proofs appear in the Appendix.

2. Background and Related Literature

There are two strands in the development of work related to portfolio optimization. One strand concerns multi-period optimal portfolio selection, initiated by Merton (1969, 1971). Security prices are modelled as continuous-time random processes based on geometric Brownian motion: see Karatzas and Shreve (1991), Yang and Zhang (2005) and Chacko and Viceira (2005) among others. A recent and significant development in this direction is made by Zhao and Rong (2012) who assume investment in multiple risky assets under the constant elasticity of variance (CEV) model with stochastic volatility. This generalizes the geometric Brownian motion model where volatility is deterministic.

In this paper, we choose to follow the second strand of portfolio optimization, namely single-period optimization originally due to Markowitz (1952). There are several reasons why we make this choice. First, the single-period mean-risk model is routinely used in the investment industry because it allows for practical trading constraints, frictional costs and similar realistic issues. We endeavor in this paper to add to the level of sophistication that may be applied to such models. Secondly, parameter and model mis-specification risks are very significant in practical portfolio management. The one-period model can accommodate these, for example through a Bayesian approach (see e.g. Garlappi et al., 2007; Tu and Zhou, 2004; Kan and Zhou, 2007), in a way which is readily implementable by practitioners. We supply a closed-form solution in the single-period model which may in future be extended to allow greater robustness. Thirdly, the effect of inter-temporal hedging is typically small, so that optimal portfolios in the dynamic setting are often (but not always) close to those in the static setting (Chacko and Viceira, 2005). Finally, both individual and institutional investors are concerned with downside and tail risks in portfolio returns, and traditional models fail to capture this. We are able to obtain explicit expressions for optimal portfolios in the one-period setting taking tail risk into account.

In order to solve for tail mean-variance optimal portfolios, we also leverage recent findings on risk measures. Indeed, the tail mean-variance criterion may be viewed as a weighted sum of two risk measures. The first term on the right hand side of equation (1)

is the tail conditional expectation of losses, which is identical to the expected shortfall risk measure since losses are assumed to be continuously distributed (see e.g. McNeil, Frey and Embrechts, 2005, p. 45). Acerbi and Tasche (2002) show that the tail conditional expectation satisfies a set of acceptable properties for risk measures and is therefore deemed to be a coherent risk measure in the sense of Artzner et al. (1999). The second term on the right hand side of equation (1) represents the tail variance, proposed by Furman and Landsman (2006).

Notice that the tail conditional expectation is the best estimate, in a least squares sense, of the worst losses on a portfolio, when losses larger than the q -quantile are considered:

$$\mathbb{E}[L \mid L > \text{VaR}_q(L)] = \arg \inf_w \mathbb{E}[(L - w)^2 \mid L > \text{VaR}_q(L)]. \quad (5)$$

On the other hand, the tail variance of Furman and Landsman (2006) gives the estimated squared deviation of the worst losses from the tail conditional expectation:

$$\text{Var}[L \mid L > \text{VaR}_q(L)] = \inf_w \mathbb{E}[(L - w)^2 \mid L > \text{VaR}_q(L)]. \quad (6)$$

The tail conditional expectation is explicitly calculated for multivariate distributions that are normal, elliptical, gamma, Pareto and for exponential dispersion models by Panjer (2002), Landsman and Valdez (2003), Furman and Landsman (2005), Chiragiev and Landsman (2007) and Landsman and Valdez (2005) respectively. Furman and Landsman (2006) calculate the tail variance in the case of multivariate normal distributions and, more generally, elliptical distributions.

3. Results on Tail Conditional Expectation and Tail Variance

We start by briefly reviewing some known results on spherical and elliptical distributions and their application to risk management.

If $\mathbf{z} : \Omega \rightarrow \mathbb{R}^n$ is spherically distributed with characteristic generator ψ , then its characteristic function is a function of the Euclidean norm of \mathbf{t} , where $\mathbf{t} \in \mathbb{R}^n$ is the argument of the characteristic function:

$$\varphi_{\mathbf{z}}(\mathbf{t}) = \mathbb{E}[\exp(it^T \mathbf{z})] = \psi\left(\frac{1}{2}\|\mathbf{t}\|^2\right). \quad (7)$$

ψ is termed the characteristic generator because it specifies different members of the spherical family of distributions. For example, the standard normal random variable has characteristic function $\exp(-t^2/2)$.

Elliptical distributions are affine transformations of spherical distributions, so that probability density contours are distorted from spheroids to ellipsoids. Let $\mathbf{r} : \Omega \rightarrow \mathbb{R}^n$ be a vector of asset returns on n securities. If \mathbf{r} is elliptically distributed with location vector $\boldsymbol{\mu} \in \mathbb{R}^n$, dispersion matrix $\Sigma \in \mathbb{R}^{n \times n}$ and characteristic generator ψ , then its characteristic function is

$$\varphi_{\mathbf{r}}(\mathbf{t}) = \mathbb{E}[\exp(it^T \mathbf{r})] = \exp(it^T \boldsymbol{\mu}) \psi\left(\frac{1}{2} \mathbf{t}^T \Sigma \mathbf{t}\right). \quad (8)$$

For further details on spherical and elliptical distributions, see Fang et al. (1990).

A key property of elliptically distributed random variables with the same characteristic generator ψ is that any linear combination of these random variables is also elliptically distributed with the same characteristic generator. This is easily established from equation (8). Thus, the return $\mathbf{x}^T \mathbf{r}$ on a portfolio \mathbf{x} is elliptically distributed with mean $\boldsymbol{\mu}^T \mathbf{x}$, variance $\mathbf{x}^T \Sigma \mathbf{x}$ and characteristic generator ψ , assuming that the mean and variance exist (Owen and Rabinovitch, 1983).

Landsman and Valdez (2003) show that, when returns are jointly elliptically distributed, the tail conditional expectation of portfolio loss, based on the q -quantile, may be written as

$$\mathbb{E}[L \mid L > \text{VaR}_q(L)] = -\boldsymbol{\mu}^T \mathbf{x} + \lambda_{1,q} \sqrt{\mathbf{x}^T \Sigma \mathbf{x}}, \quad (9)$$

since $L = -\mathbf{x}^T \mathbf{r}$. The parameter $\lambda_{1,q}$ is uniquely specified by q :

$$\lambda_{1,q} = h_{Z,Z^*}(z_q), \quad (10)$$

where $h_{Z,Z^*}(z) = f_{Z^*}(z)/\bar{F}_Z(z)$ is a distorted hazard rate evaluated at the q -quantile of the standardized random variable $Z = (r_1 - \mu_1)/\sqrt{\sigma_{11}}$. Here, $\bar{F}_Z(z) = 1 - F_Z(z)$ is the decumulative distribution function of Z whereas $F_Z(z)$ is the cumulative distribution function of Z ; the first element of \mathbf{r} is r_1 with mean $\mu_1 = \mathbb{E}r_1$ and variance $\sigma_{11} = \text{Var}(r_1)$; and $f_{Z^*}(z)$ is the density of some spherical random variable Z^* , the distribution of which

is called the distribution associated with the elliptical family (see Landsman and Valdez, 2003, for details).

The corresponding tail variance, when returns are elliptically distributed, is derived by Furman and Landsman (2006):

$$\text{Var}[L \mid L > \text{VaR}_q(L)] = \lambda_{2,q} \mathbf{x}^T \Sigma \mathbf{x}. \quad (11)$$

The parameter $\lambda_{2,q}$ is also uniquely specified by q :

$$\lambda_{2,q} = r(z_q) + h_{Z,Z^*}(z_q)(z_q - h_{Z,Z^*}(z_q)), \quad (12)$$

where $r(z) = \overline{F}_{Z^*}(z)/\overline{F}_Z(z)$ and is also evaluated at the q -quantile of Z , and $\overline{F}_{Z^*}(z)$ is the decumulative distribution function of Z^* .

In view of equations (9) and (11), the tail mean-variance criterion of equation (1) may be written as (Landsman, 2010):

$$f(\mathbf{x}; \lambda, q) = -\boldsymbol{\mu}^T \mathbf{x} + \lambda_{1,q} \sqrt{\mathbf{x}^T \Sigma \mathbf{x}} + \lambda \lambda_{2,q} \mathbf{x}^T \Sigma \mathbf{x}. \quad (13)$$

We may write the tail mean-variance criterion more succinctly as $f(\mathbf{x})$ where this causes no confusion. The two parameters of the tail mean-variance criterion describe an investor's risk preferences. The first parameter, $\lambda \in \mathbb{R}_{++}$, can be regarded as a risk aversion parameter and is akin to τ in the classical mean-variance criterion of equation (3). The larger λ , the more risk-averse an investor is, and the larger the additional return they expect as a compensation for a unit increase in the variance of their portfolio return. The second parameter, $q \in (0, 1)$, defines a certain threshold of loss on the portfolio. A tail mean-variance-optimizing investor is sensitive to losses beyond the q -quantile. These are losses which are typically rare but large.

4. Characterization of the Tail Mean-Variance Optimal Portfolio

4.1. Classical Mean-Variance Criterion

It is well-known that, when returns or losses follow a multivariate elliptical distribution, risk management under risk measures such as the value-at-risk and the tail conditional

expectation is equivalent to the use of the classical mean-variance framework. This is highlighted by McNeil, Frey and Embrechts (2005), with an emphasis on the value-at-risk measure. More specifically, Proposition 6.13 of McNeil, Frey and Embrechts (2005, p. 247) shows that an investor who wishes to achieve a target expected return will choose the same optimal portfolio whether his measure of risk is the variance, the value-at-risk or any other measure that satisfies the properties of translation invariance and positive homogeneity. (See Artzner et al. (1999) for properties of coherent risk measures.) It is clear from equation (1) that the tail mean-variance criterion does not satisfy positive homogeneity: $TMV(kL) \neq k \times TMV(L)$ for $k > 0$. Nevertheless, we shall show that the tail mean-variance criterion, under multivariate elliptical risks, does revert to the mean-variance setting.

To this end, consider the classical mean-variance criterion

$$g(\mathbf{x}; \tau) = -\boldsymbol{\mu}^T \mathbf{x} + \frac{1}{2} \tau \mathbf{x}^T \Sigma \mathbf{x}, \quad (14)$$

which corresponds to equation (3). $g(\mathbf{x}; \tau)$ expresses a risk-averse investor's tradeoff between expected portfolio return $\boldsymbol{\mu}^T \mathbf{x}$ and variance of portfolio return $\mathbf{x}^T \Sigma \mathbf{x}$, through a risk aversion parameter $\tau > 0$. Such an investor seeks to minimize $g(\mathbf{x}; \tau)$ subject to a budget constraint such as $\mathbf{1}^T \mathbf{x} = 1$.

The minimization of $g(\mathbf{x})$ wrt $\mathbf{x} \in \mathcal{P}$ is straightforward (see e.g. Panjer et al., 1998, p. 382) and results in the optimal solution

$$\bar{\mathbf{x}} = \mathbf{x}_0 + \frac{1}{\tau} \mathbf{z} \quad (15)$$

where

$$\mathbf{x}_0 = (\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1} \Sigma^{-1} \mathbf{1}, \quad (16)$$

$$\mathbf{z} = \Sigma^{-1} \boldsymbol{\mu} - (\mathbf{1}^T \Sigma^{-1} \boldsymbol{\mu}) (\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1} \Sigma^{-1} \mathbf{1}. \quad (17)$$

$\mathbf{x}_0 \in \mathcal{P}$ is the global minimum variance portfolio, i.e. the portfolio that minimizes $\mathbf{x}^T \Sigma \mathbf{x}$ subject to $\mathbf{1}^T \mathbf{x} = 1$. As for $\mathbf{z} \notin \mathcal{P}$, it is a self-financing portfolio since it clearly satisfies the property that $\mathbf{1}^T \mathbf{z} = 0$.

The concept of efficient frontiers is natural in mean-variance optimization. We employ the usual definition of mean-variance efficiency that is found in textbook treatments of portfolio theory, e.g. in Panjer et al. (1998, p. 379): a portfolio \mathbf{x} is called (mean-variance) efficient if there exists no portfolio \mathbf{x}_1 with $\boldsymbol{\mu}^T \mathbf{x}_1 \geq \boldsymbol{\mu}^T \mathbf{x}$ and $\mathbf{x}_1^T \Sigma \mathbf{x}_1 < \mathbf{x}^T \Sigma \mathbf{x}$. The mean-variance efficient frontier \mathcal{E} is then the set of portfolios defined by

$$\mathcal{E} = \{\mathbf{x}_0\} \cup \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \arg \min_{\mathbf{x} \in \mathcal{P}} g(\mathbf{x}; \tau) \text{ for each } \tau \in \mathbb{R}_{++}\}. \quad (18)$$

At this stage we collect some statistical properties about the returns of $\bar{\mathbf{x}}$, \mathbf{x}_0 and \mathbf{z} , as they will be useful later. Recall throughout that Σ is a real, positive definite, symmetric, non-singular matrix.

It is well-known (see e.g. Lee and Lee, 2006, p. 458) that the return on the global mean-variance portfolio \mathbf{x}_0 in equation (16) has the same covariance with the return on all non-self-financing portfolios $\mathbf{x} \in \mathcal{P}$ (for which $\mathbf{1}^T \mathbf{x} = 1$), including itself:

$$\mathbf{x}^T \Sigma \mathbf{x}_0 = (\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1} \mathbf{x}^T \mathbf{1} = (\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1} = \mathbf{x}_0^T \Sigma \mathbf{x}_0. \quad (19)$$

But the return on \mathbf{x}_0 also has zero covariance with the return on the self-financing portfolio \mathbf{z} of equation (17) (for which $\mathbf{1}^T \mathbf{z} = 0$):

$$\mathbf{z}^T \Sigma \mathbf{x}_0 = (\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1} \mathbf{z}^T \mathbf{1} = 0. \quad (20)$$

The mean return on the global mean-variance optimal portfolio is proportional to its variance:

$$\boldsymbol{\mu}^T \mathbf{x}_0 = (\boldsymbol{\mu}^T \Sigma^{-1} \mathbf{1}) (\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1} = (\boldsymbol{\mu}^T \Sigma^{-1} \mathbf{1}) (\mathbf{x}_0^T \Sigma \mathbf{x}_0) \quad (21)$$

in view of equation (19).

Turning to the self-financing portfolio \mathbf{z} of equation (17), and comparing it with \mathbf{x}_0 in equation (16), we find that $\mathbf{z} = \Sigma^{-1} \boldsymbol{\mu} - (\mathbf{1}^T \Sigma^{-1} \boldsymbol{\mu}) \mathbf{x}_0$, so that its mean return is

$$\boldsymbol{\mu}^T \mathbf{z} = \boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu} - (\mathbf{1}^T \Sigma^{-1} \boldsymbol{\mu}) \boldsymbol{\mu}^T \mathbf{x}_0, \quad (22)$$

which equals its variance of return since

$$\mathbf{z}^T \Sigma \mathbf{z} = \boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu} - 2(\mathbf{1}^T \Sigma^{-1} \boldsymbol{\mu}) \boldsymbol{\mu}^T \mathbf{x}_0 + (\mathbf{1}^T \Sigma^{-1} \boldsymbol{\mu})^2 \mathbf{x}_0^T \Sigma \mathbf{x}_0 = \boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu} - (\mathbf{1}^T \Sigma^{-1} \boldsymbol{\mu}) \boldsymbol{\mu}^T \mathbf{x}_0. \quad (23)$$

Replacing $\boldsymbol{\mu}^T \mathbf{x}_0$ from equation (21) into equation (22) results in an expression for the mean and variance of the return on the self-financing portfolio:

$$\boldsymbol{\mu}^T \mathbf{z} = \mathbf{z}^T \Sigma \mathbf{z} = \left[(\boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu}) (\mathbf{1}^T \Sigma^{-1} \mathbf{1}) - (\mathbf{1}^T \Sigma^{-1} \boldsymbol{\mu})^2 \right] (\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1}. \quad (24)$$

Finally, from equations (15), (19), (20) and (23), we find that the variance of return on the mean-variance optimal portfolio is:

$$\bar{\mathbf{x}}^T \Sigma \bar{\mathbf{x}} = \mathbf{x}_0^T \Sigma \mathbf{x}_0 + \frac{1}{\tau^2} \mathbf{z}^T \Sigma \mathbf{z} + \frac{2}{\tau} \mathbf{x}_0^T \Sigma \mathbf{z} = (\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1} + \frac{1}{\tau^2} \boldsymbol{\mu}^T \mathbf{z}. \quad (25)$$

4.2. Tail Mean-Variance Criterion

Landsman (2010) shows that the tail mean-variance optimal portfolio that minimizes $f(\mathbf{x})$ (in equation (13)) subject to $\mathbf{1}^T \mathbf{x} = 1$ is

$$\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x}) = \mathbf{x}_0 + \frac{w^*}{\ell} \mathbf{u}. \quad (26)$$

In the above, \mathbf{x}_0 is the global minimum variance portfolio of equation (16), and ℓ and \mathbf{u} are functions of mean $\boldsymbol{\mu}$ and covariance Σ of stock returns. The functions ℓ and \mathbf{u} are fairly complicated and involve partitioned matrices and their inverses and they are not reproduced here: see Landsman (2010) for details. As for w^* , it is the unique real root within the interval $(0, \ell/2\lambda\lambda_{2,q})$ of the following quartic equation:

$$w^4 - \frac{\ell}{\lambda\lambda_{2,q}} w^3 + \left[(\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1} + \frac{\ell^2 - \lambda_{1,q}^2}{(2\lambda\lambda_{2,q})^2} \right] w^2 - \ell \frac{(\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1}}{\lambda\lambda_{2,q}} w + \ell^2 \frac{(\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1}}{(2\lambda\lambda_{2,q})^2} = 0 \quad (27)$$

Our aim is to explain the form of \mathbf{x}^* , relate it to mean-variance efficient portfolios, and simplify its calculation. The following two lemmas are useful in proving our main theorem.

Lemma 1. *Assume $q \in (0, 1)$. Then $\lambda_{1,q} > 0$ and $\lambda_{2,q} > 0$.*

Proof. $\lambda_{1,q}$ is a ratio of probability density and decumulative probability in equation (10). $\lambda_{2,q} > 0$ follows immediately from equation (11). \square

Lemma 2. *$f(\mathbf{x})$ is strictly convex on $\mathcal{P} \subset \mathbb{R}^n$.*

See Landsman (2010) for a proof. The mean $\boldsymbol{\mu}^T \mathbf{x}$ and variance $\mathbf{x}^T \Sigma \mathbf{x}$ are continuous on \mathcal{P} , and so is $f(\mathbf{x})$. On the r.h.s. of equation (13), $\mathbf{x}^T \Sigma \mathbf{x}$ is strictly convex on \mathbb{R}^n and $-\boldsymbol{\mu}^T \mathbf{x}$ is convex (linear), while the coefficients $\lambda_{1,q}$ and $\lambda \lambda_{2,q}$ are positive by Lemma 1. So one only needs to show that $\sqrt{\mathbf{x}^T \Sigma \mathbf{x}}$ is convex on \mathbb{R}^n , rather than strictly convex.

A consequence of Lemma 2 is that $f(\mathbf{x})$ has at most one minimum on \mathcal{P} . The following lemma characterizes this minimum if it exists.

Lemma 3. *Assume that there exists a portfolio $\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x})$. Then $\mathbf{x}^* \in \mathcal{E}$.*

Proof. We wish to show that $\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x}) \Rightarrow \mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathcal{P}} g(\mathbf{x}; \tau_0)$ for some $\tau_0 > 0$. Proof by contradiction: Suppose that $\mathbf{x}^* = \arg \min_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x})$ and $\mathbf{x}^* \neq \arg \min_{\mathbf{x} \in \mathcal{P}} g(\mathbf{x}; \tau_0)$. Then there exists a portfolio $\mathbf{x}_1 \in \mathcal{P}$ with $g(\mathbf{x}_1; \tau_0) < g(\mathbf{x}^*; \tau_0)$ such that $\boldsymbol{\mu}^T \mathbf{x}_1 \geq \boldsymbol{\mu}^T \mathbf{x}^*$ and $\mathbf{x}_1^T \Sigma \mathbf{x}_1 < \mathbf{x}^{*T} \Sigma \mathbf{x}^*$. In equation (13), $\lambda > 0$ by definition and $\lambda_{1,q}, \lambda_{2,q} > 0$ by virtue of Lemma 1. Therefore, $f(\mathbf{x}_1) < f(\mathbf{x}^*)$. But then $\mathbf{x}^* \neq \arg \min_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x})$, which is a contradiction. \square

The proof of Lemma 3 is simple enough, but it is illuminating to rewrite the tail mean-variance criterion in equation (13) as follows:

$$f(\mathbf{x}) = -\boldsymbol{\mu}^T \mathbf{x} + \lambda \lambda_{2,q} \left(\sqrt{\mathbf{x}^T \Sigma \mathbf{x}} + \frac{\lambda_{1,q}}{2\lambda \lambda_{2,q}} \right)^2 - \frac{\lambda_{1,q}^2}{4\lambda \lambda_{2,q}}. \quad (28)$$

Note that $\lambda \lambda_{2,q} > 0$, $\lambda_{1,q}/2\lambda \lambda_{2,q} > 0$ by Lemma 1. For $\sqrt{\mathbf{x}^T \Sigma \mathbf{x}} > 0$, the quadratic $\left(\sqrt{\mathbf{x}^T \Sigma \mathbf{x}} + \lambda_{1,q}/2\lambda \lambda_{2,q} \right)^2$ is monotonically increasing in $\sqrt{\mathbf{x}^T \Sigma \mathbf{x}}$, and therefore in $\mathbf{x}^T \Sigma \mathbf{x}$. Thus, holding the mean of portfolio return constant, minimizing $f(\mathbf{x})$ is equivalent to minimizing the variance of portfolio return.

Based on Lemma 3, it would not be surprising if the tail mean-variance optimal portfolio solution \mathbf{x}^* has the same form as the mean-variance optimal portfolio $\bar{\mathbf{x}}$. Comparing the solution \mathbf{x}^* of Landsman (2010) in equation (26) to $\bar{\mathbf{x}}$ in equation (15), we find that they do indeed bear a strong resemblance in that they both involve an adjustment to the global minimum variance portfolio \mathbf{x}_0 .

5. Tail Mean-Variance Optimization with Risky Securities Only

5.1. Optimization

If a tail mean-variance-optimizing investor has a given target expected return, then his optimal portfolio is readily calculated using standard methods as it lies on the mean-variance efficient frontier. This is indeed also true of investors who have a target expected return and wish to minimize a translation-invariant and positive-homogeneous risk measure, as stated in Proposition 6.13 of McNeil, Frey and Embrechts (2005, p. 247). However, if there is no given target expected return, and if the investor wishes to freely optimize the tail mean-variance criterion of equation (1), the precise optimal solution is not so easily calculated and would typically be found using numerical methods. Nevertheless, Landsman (2010) obtained an analytic solution and our aim in this section is to shed some light on this solution and to simplify it.

Before proving our main theorem, we establish a simple lemma on quartic polynomials. As will become apparent later, this lemma means that we do not need to be concerned with the algebraic solution of quartic equations.

Lemma 4. *The quartic polynomial $F(x) = (x-a)^2(x^2+b) - cx^2$ with $a, b, c > 0$ has exactly one non-coincident real zero in $(0, a)$ and exactly one non-coincident real zero in (a, ∞) .*

See Appendix A.1 for a proof of the above. The following theorem contains our main result.

Theorem 1. *The minimum of $f(\mathbf{x}; \lambda, q)$ in equation (13) wrt $\mathbf{x} \in \mathcal{P}$ exists and is unique and occurs at*

$$\mathbf{x}^* = \mathbf{x}_0 + \frac{1}{\tau^*} \mathbf{z}, \quad (29)$$

where \mathbf{x}_0 and \mathbf{z} are given in equations (16) and (17) respectively, and $\tau^* \in \mathbb{R}$ is the unique root of the quartic equation

$$(\tau - 2\lambda\lambda_{2,q})^2 \left[(\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{-1} \tau^2 + \boldsymbol{\mu}^T \mathbf{z} \right] - \tau^2 \lambda_{1,q}^2 = 0 \quad (30)$$

that is located in the range $(2\lambda\lambda_{2,q}, \infty)$.

Proof. Define the Lagrangian $\mathcal{L}_f(\mathbf{x}, \gamma_f) = f(\mathbf{x}) - \gamma_f(\mathbf{1}^\top \mathbf{x} - 1)$ where $\gamma_f \in \mathbb{R}$ is a Lagrange multiplier. $\partial \mathcal{L}_f / \partial \mathbf{x} = \mathbf{0}$ and $\partial \mathcal{L}_f / \partial \gamma_f = 0$ may be written as

$$\boldsymbol{\mu} - \left(\frac{\lambda_{1,q}}{\sqrt{\mathbf{x}^\top \Sigma \mathbf{x}}} + 2\lambda\lambda_{2,q} \right) \Sigma \mathbf{x} + \gamma_f \mathbf{1} = \mathbf{0}, \quad \mathbf{1}^\top \mathbf{x} = 1. \quad (31)$$

Suppose that the solution to the equation set (31) exists and denote it by $(\mathbf{x}^*, \gamma_f^*)$.

Consider instead the mean-variance optimization problem with $g(\mathbf{x})$ defined in equation (14). Define the Lagrangian $\mathcal{L}_g(\mathbf{x}, \gamma_g) = g(\mathbf{x}) - \gamma_g(\mathbf{1}^\top \mathbf{x} - 1)$ where $\gamma_g \in \mathbb{R}$ is another Lagrange multiplier. The optimal solution $(\bar{\mathbf{x}}, \bar{\gamma}_g)$ is the solution to

$$\boldsymbol{\mu} - \tau \Sigma \mathbf{x} + \gamma_g \mathbf{1} = \mathbf{0}, \quad \mathbf{1}^\top \mathbf{x} = 1. \quad (32)$$

Now, the solution $(\mathbf{x}^*, \gamma_f^*)$ of equation system (31) coincides with the solution $(\bar{\mathbf{x}}, \bar{\gamma}_g)$ of equation system (32) provided that τ takes a specific value, denoted by τ^* , such that

$$\tau^* = \frac{\lambda_{1,q}}{\sqrt{(\mathbf{x}^*)^\top \Sigma \mathbf{x}^*}} + 2\lambda\lambda_{2,q} = \frac{\lambda_{1,q}}{\sqrt{(\bar{\mathbf{x}}^\top \Sigma \bar{\mathbf{x}})}} + 2\lambda\lambda_{2,q}. \quad (33)$$

The existence of \mathbf{x}^* depends on the existence of a solution for $\tau^* \in \mathbb{R}_{++}$ in the above equation.

Substituting $\bar{\mathbf{x}}^\top \Sigma \bar{\mathbf{x}}$ from equation (25) into equation (33) and rearranging yields

$$\tau^* - 2\lambda\lambda_{2,q} = \frac{\tau^* \lambda_{1,q}}{\sqrt{((\mathbf{1}^\top \Sigma^{-1} \mathbf{1})^{-1} + \boldsymbol{\mu}^\top \mathbf{z})}}, \quad (34)$$

which can also be rewritten as the quartic equation (30). In equation (34), $\lambda_{1,q} > 0$ and $2\lambda\lambda_{2,q} > 0$ by Lemma 1, and $\mathbf{1}^\top \Sigma^{-1} \mathbf{1} > 0$ by positive definiteness of Σ . Furthermore, $\boldsymbol{\mu}^\top \mathbf{z} > 0$ because $\boldsymbol{\mu}^\top \mathbf{z} = \mathbf{z}^\top \Sigma \mathbf{z}$ from equation (23) and Σ is positive definite and $\mathbf{z} \neq \mathbf{0}$. ($\boldsymbol{\mu}^\top \mathbf{z} > 0$ also follows from the Cauchy-Schwarz inequality on the r.h.s. of equation (24).) Since $\tau^* \in \mathbb{R}_{++}$, the r.h.s. of equation (34) is positive and real, and so should be the l.h.s. Thus, $\tau^* > 2\lambda\lambda_{2,q}$. Lemma 4 confirms the existence of a unique real root of equation (30) in the range $(2\lambda\lambda_{2,q}, \infty)$, which also serves to confirm the existence of $\mathbf{x}^* = \bar{\mathbf{x}}$ with $\tau = \tau^*$ in equation (15). Finally, uniqueness of the minimum at \mathbf{x}^* is guaranteed by Lemma 2. \square

5.2. Practical Advantages and Comparison with Previous Solution

Some remarks about Theorem 1 are in order. First, the optimal portfolio \mathbf{x}^* in equation (29) corresponds to the original solution in equation (26) derived by Landsman (2010), but has a much simpler form. In particular, the quartic equation (30) corresponds to the original quartic equation (27) but is simpler.

Secondly, Theorem 1 is markedly easier to implement than the original solution of Landsman (2010). The cumbersome functions ℓ and \mathbf{u} in equation (26) do not arise, and successive matrix partitions, concatenations and inversions are avoided. For practical investment purposes, a large number of securities is likely to be held in a portfolio. Our solution speeds up considerably the computation of optimal portfolios when there are a large number of securities, involving large covariance matrices.

Thirdly, portfolio optimization is known to be sensitive to parameter and model misspecification risks (see e.g. Tu and Zhou, 2004; Kan and Zhou, 2007). The simplified optimal portfolio in equation (29) means that sensitivity analysis is feasible and more straightforward. A portfolio manager can investigate different portfolios as parameters change. In particular, the change in the optimal portfolio under different multivariate elliptical distributions, capturing increasing or decreasing dependence in the tail, can be measured. The construction of robust portfolios, for example incorporating Bayesian priors, can also be contemplated (see e.g. Garlappi et al., 2007).

Fourthly, the original proof of Landsman (2010) proceeds by laborious substitution of the portfolio budget constraint into the tail mean-variance optimization objective of equation (13). Our proof involves an elegant use of the Lagrangian multiplier and a comparative analysis with another optimization problem. The solutions are numerically identical, but the one in Theorem 1 is more insightful in addition to being more computationally convenient.

Finally, a key difference between the proof of Theorem 1 and the original proof of Landsman (2010) is the use of Lemma 4. This obviates the need to explore the algebraic solution of quartic equations. Lemmas 2 and 4 furnish uniqueness and existence respectively. But the quartic equation (30) has four roots, of which three must be ruled out. Equation (34)

shows that roots outside the range $(2\lambda\lambda_{2,q}, \infty)$ are redundant. Showing that $h(0) > 0$ and $h(2\lambda\lambda_{2,q}) < 0$, where $h(\tau)$ is the quartic polynomial (with a positive leading coefficient) on the l.h.s. of equation (30), establishes only that there is at least one real root in the range $(0, 2\lambda\lambda_{2,q})$ and *at least* one real root in the range $(2\lambda\lambda_{2,q}, \infty)$. In order to establish that there is *exactly* one real root in the range that corresponds to $(2\lambda\lambda_{2,q}, \infty)$ in the quartic equation (27), Landsman (2010) uses the famous solution of Ferrari, involving the resolvent cubic (Neumark, 1965). The algebraic solution of cubics and quartics are equally famous for being unwieldy, and the quartic equation (30) is easily solved by numerical methods in practice, so Lemma 4 is assuredly of greater convenience.

5.3. Numerical Illustration and Comparison with Other Portfolio Optimization Models

To illustrate our portfolio optimization solution, we use the same numerical example as given by Landsman (2010), which facilitates comparison with this earlier study. The weekly returns in 2007 on 10 stocks, which are listed on Nasdaq and grouped in the Computers industrial sector, are used in this example. The stocks and their means and covariances of returns are displayed in Landsman (2010) and we do not reproduce them here, in order to save space.

Optimal portfolio weights are tabulated in Landsman (2010) when returns are multivariate normal. The optimal portfolios can be calculated more quickly using our Theorem 1, so we model the stock returns with a multivariate Student- t distribution with 6 degrees of freedom ($\nu = 6$) thereby capturing the heavy-tailed feature of asset returns. We use the mean vector and covariance matrix estimated by Landsman (2010). It is worth mentioning that comparable statistics are obtained by McNeil, Frey and Embrechts (2005, p. 85) who fit a Student- t distribution to the weekly and daily returns on 10 of the Dow Jones stocks. Tu and Zhou (2004) and Jorion (1996), among other authors, also find that the Student- t distribution cannot be rejected on U.S. stock return data over different periods and return horizons, and they report statistics commensurate to those used here.

As is well known, the dispersion matrix in the Student- t distribution is a scaled version of the covariance matrix, with the scaling factor depending on the number of degrees

Stock	ADBE	TISA	HAUP	IMMR	LOGI
Min TMV $q = 0.95$	0.3653	0.2034	0.0842	-0.0315	0.0358
Min MV	0.4931	0.3264	0.1699	-0.1176	0.0390
Min VaR $q = 0.95$	0.4042	0.2409	0.1103	-0.0577	0.0368
Min Variance	0.3298	0.1691	0.0603	-0.0075	0.0349

Stock	NVDA	OIIM	PANL	SCMM	SSYS
Min TMV $q = 0.95$	0.0138	0.1316	-0.0261	0.1836	0.0400
Min MV	-0.0778	0.1357	-0.0994	0.1480	-0.0174
Min VaR $q = 0.95$	-0.0141	0.1328	-0.0484	0.1728	0.0225
Min Variance	0.0393	0.1304	-0.0057	0.1935	0.0559

Table 1: Optimal portfolio weights obtained by minimizing (i) tail conditional expectation (Min TCE), (ii) value-at-risk (Min VaR), (iii) variance (Min Variance).

of freedom. Landsman and Valdez (2003) and Landsman (2010) introduce a generalized Student- t distribution whose kurtosis may be adjusted through a certain power parameter, while keeping its covariance matrix constant and equal to the dispersion matrix. We use the classical multivariate Student- t here, as also discussed by Landsman and Valdez (2003), since it is better known and has numerous applications in finance and other areas.

Optimal portfolio weights are shown in Table 1. In the table and subsequent figures, ‘Min TMV’, ‘Min MV’, ‘Min VaR’ and ‘Min Variance’ denote the portfolios obtained when minimizing the tail mean-variance criterion, the mean-variance criterion, the value-at-risk and the variance respectively. The optimal portfolios in Table 1 are comparable, but not identical, with those tabulated in Landsman (2010) since we use the multivariate Student- t distribution and therefore capture heavy tails.

The optimal tail mean-variance portfolio (‘Min TMV’), calculated at the confidence level $q = 0.95$, is displayed in mean-standard deviation space in Figure 1. The mean-variance frontier is also displayed. We note that ‘Min TMV’ lies on the efficient frontier,

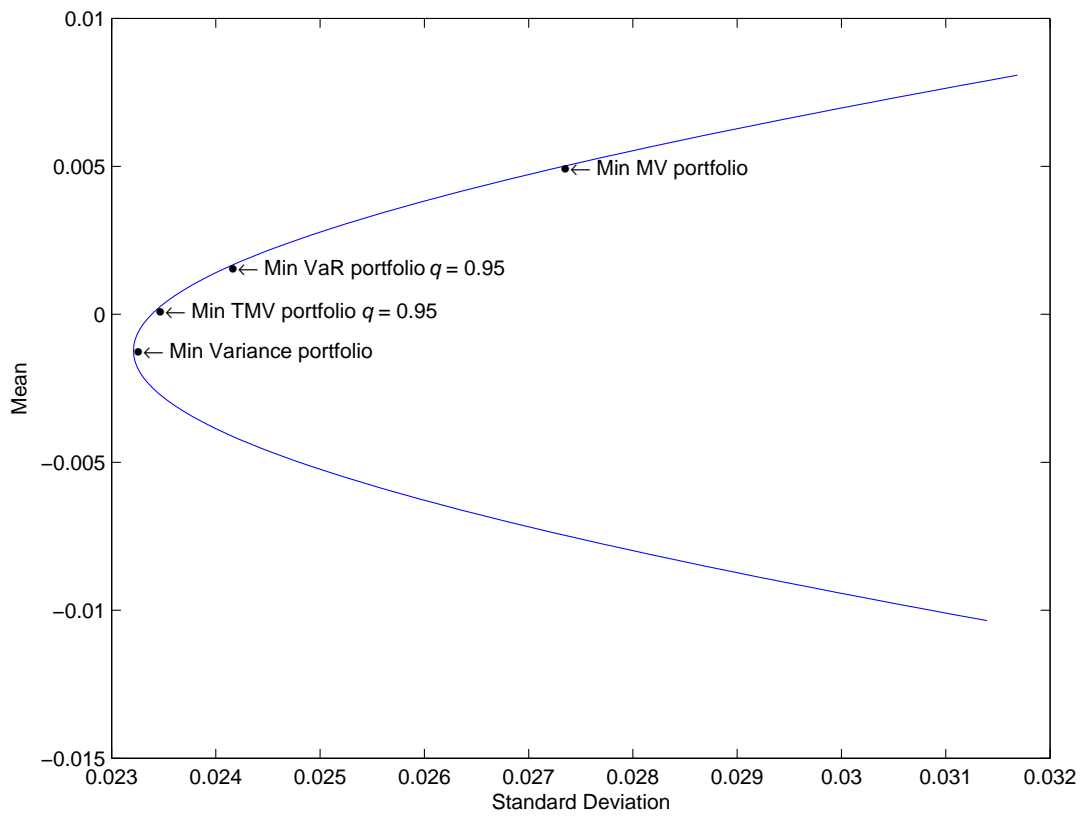


Figure 1: Portfolio frontier showing optimal portfolios obtained by minimizing (i) variance (Min Variance), (ii) tail mean-variance criterion (Min TMV), (iii) value-at-risk (Min VaR), (iv) mean-variance criterion (Min MV).

defined as \mathcal{E} in equation (18), which consists of the concave segment of the frontier in Figure 1. This is consistent with Lemma 3, of course.

In Figure 1, it is natural to compare the ‘Min TMV’ portfolio, which is obtained by minimizing the criterion in equation (1), with the corresponding mean-variance optimal portfolio (‘Min MV’), which is obtained by minimizing the criterion in equation (3). We set $\lambda = \tau/2$ for consistency between equations (1) and (3). The ‘Min MV’ portfolio also lies on the efficient frontier in Figure 1, of course. In fact, a comparison of the tail mean-variance optimal portfolio \mathbf{x}^* in Theorem 1 and the classical mean-variance optimal portfolio $\bar{\mathbf{x}}$ of equation (15) immediately yields the following corollary.

Corollary 1. $\arg \min_{\mathbf{x} \in \mathcal{P}} f(\mathbf{x}; \lambda, q) = \arg \min_{\mathbf{x} \in \mathcal{P}} g(\mathbf{x}; \tau^*)$, where τ^* is given in Theorem 1.

Corollary 1 says that optimization of the tail mean-variance criterion can be achieved by optimizing the classical mean-variance criterion with the risk aversion parameter τ^* being evaluated as the root of a quartic equation according to Theorem 1. This is also consistent with Lemma 3, but it goes further because, through the calculation of τ^* , it gives a simple method for a portfolio manager to allow for leptokurtic asset returns and aversion to tail risk, through the tail mean-variance model.

In Figure 1, the ‘Min TMV’ portfolio is more conservative than the ‘Min MV’ portfolio, that is, its return has a lower standard deviation than that of the ‘Min MV’ portfolio. This is not surprising since avoiding large losses is a key objective in the tail mean-variance criterion.

It is also natural to compare the ‘Min TMV’ portfolio with the portfolio that minimizes value-at-risk (denoted by ‘Min VaR’ in Figure 1). Recall that q , in equations (1) and (2), represents the quantile threshold beyond which an investor is sensitive to losses. Our numerical experiments, as illustrated by Figure 1, show that the optimal tail mean-variance portfolio appears to be more conservative than the minimum value-at-risk portfolio at the same confidence level q , in the sense that it consists of higher weights in the less volatile stocks.

Figure 2 shows that the tail mean-variance criterion affords greater control than the

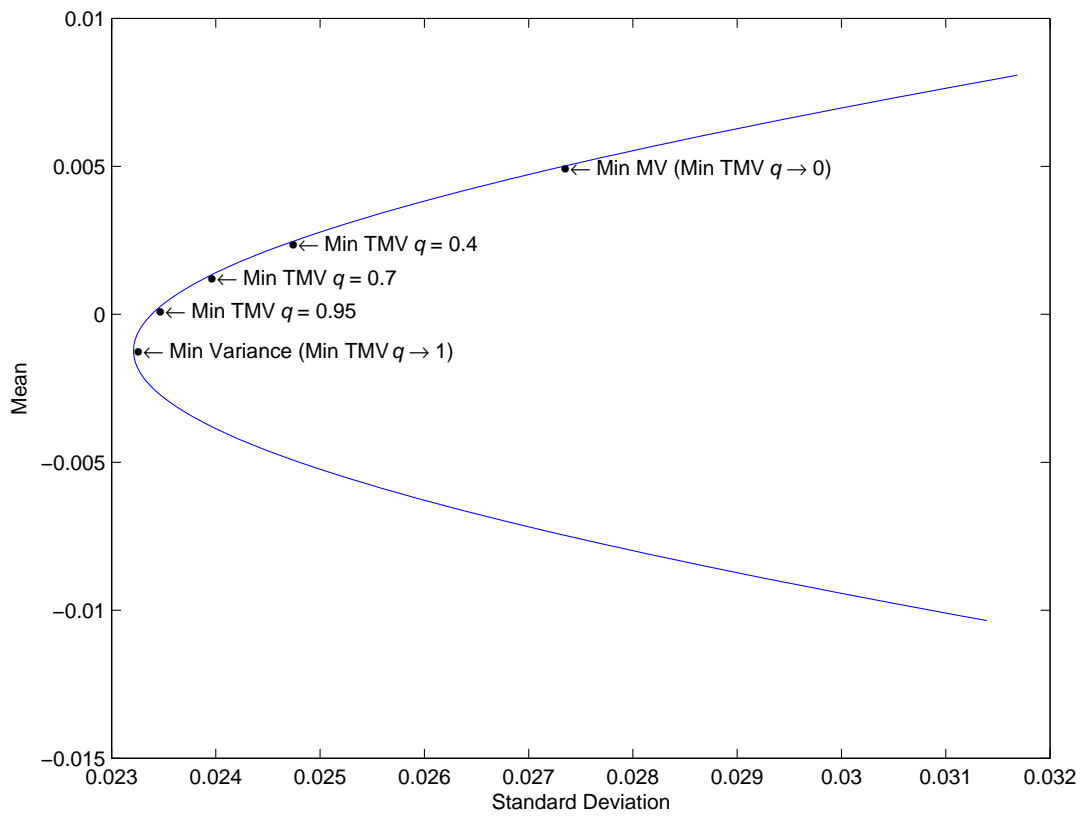


Figure 2: Portfolio frontier showing optimal portfolios obtained by minimizing (i) variance (Min Variance), (ii) tail mean-variance criterion (Min TMV) at different confidence levels $q = 0.4, 0.7, 0.95$, (iii) mean-variance criterion (Min MV).

mean-variance criterion in that an investor can account for his aversion to tail risk. As demonstrated by Landsman (2010) and depicted in Figure 2, the ‘Min TMV’ portfolio tends to the ‘Min MV’ portfolio as $q \rightarrow 0$; on the other hand, it tends to the global minimum variance portfolio as $q \rightarrow 1$. An investor can specify the q -quantile of loss beyond which he is sensitive to losses. The greater this threshold is, the more sensitive he is to large losses, and the more conservative his optimal portfolio becomes.

6. Tail Mean-Variance Optimization with a Risk-Free Asset

Suppose that a risk-free asset earning a non-random rate r is introduced to the portfolio opportunity set. The overall portfolio of an investor then consists of holding a proportion, say $y \in \mathbb{R}$, of total wealth in a sub-portfolio of risky securities and the remainder $(1 - y)$ in the risk-free asset. We assume that there is no constraint on the risk-free asset, i.e. that both risk-free lending and borrowing at rate r are possible.

6.1. Classical Mean-Variance Optimization with Risk-Free Lending and Borrowing

We first consider classical mean-variance optimization only. It is well-known that the efficient frontier of the overall portfolio, in mean-standard deviation space, is a straight line with a positive slope and a vertical intercept at r . As emphasized by Merton (1972), three cases can arise, depending on the level of the risk-free rate r relative to the mean return of risky assets. The three cases depend, more specifically, on the relationship between the risk-free rate r and the mean return on the minimum variance portfolio \mathbf{x}_0 . The first case is economically significant and tends to be the only one dealt with in elementary texts (Sharpe et al., 1995; Elton et al., 2011), but for completeness we consider all three possible cases below.

$$\text{Case 1. } r < \boldsymbol{\mu}^T \mathbf{x}_0 \Leftrightarrow \mathbf{1}^T \Sigma^{-1} (\boldsymbol{\mu} - r \mathbf{1}) > 0.$$

In mean-standard deviation space, the efficient frontier is a straight line with a positive slope starting from the risk-free asset r_f and going through a “tangency” portfolio \mathbf{x}_t :

$$\mathbf{x}_t = \frac{\Sigma^{-1} (\boldsymbol{\mu} - r \mathbf{1})}{\mathbf{1}^T \Sigma^{-1} (\boldsymbol{\mu} - r \mathbf{1})}. \quad (35)$$

See e.g. Kennedy (2010, p. 16). Note that $\mathbf{x}_t \in \mathcal{E}$, where \mathcal{E} is defined in equation (18) and represents the original efficient frontier of risky securities only. In this case, \mathbf{x}_t is never shorted, i.e. $y \geq 0$. If $y > 1$, then the investor borrows at rate r ; if $y < 1$, the investor lends at rate r ; and if $y = 1$, there is neither lending nor borrowing.

Case 2. $r > \boldsymbol{\mu}^\top \mathbf{x}_0 \Leftrightarrow \mathbf{1}^\top \Sigma^{-1}(\boldsymbol{\mu} - r\mathbf{1}) < 0$.

The efficient frontier is again a straight line with a positive slope from r , but this time it consists of short positions in \mathbf{x}_t financing long positions in the risk-free asset. Thus, $y \leq 0$.

Notice that the denominator of \mathbf{x}_t in equation (35) is negative in Case 2. In fact, \mathbf{x}_t now lies in the lower half of the hyperbolic mean-variance frontier of risky securities only, so the straight line efficient frontier is *not* tangential to \mathcal{E} in this case, but rather lies above it. For details, see Merton (1972) and Huang and Litzenberger (1988, p. 79).

Case 3. $r = \boldsymbol{\mu}^\top \mathbf{x}_0 \Leftrightarrow \mathbf{1}^\top \Sigma^{-1}(\boldsymbol{\mu} - r\mathbf{1}) = 0$.

The efficient frontier is again a straight line with a positive slope from r , but now consists of all wealth invested in the risk-free asset, along with investment, of a proportion $y \geq 0$ of wealth, in the self-financing portfolio \mathbf{z} introduced in equation (17). While we do not show this explicitly here, it is not difficult to show that the ‘‘arbitrage portfolio’’ derived by Huang and Litzenberger (1988, p. 80) is identical to \mathbf{z} .

6.2. Tail Mean-Variance Optimization with Risk-Free Lending and Borrowing

We now return to tail mean-variance optimization.

Theorem 2. *Suppose that both risk-free lending and borrowing at rate r are available.*

(a) *If $r \neq \boldsymbol{\mu}^\top \mathbf{x}_0$, then the optimal tail mean-variance portfolio consists of holding a proportion y^* of wealth in tangency portfolio \mathbf{x}_t and the rest in the risk-free asset, where*

$$y^* = \begin{cases} \frac{A(B - \lambda_{1,q})}{2\lambda\lambda_{2,q}B} & \text{if } \lambda_{1,q} \leq B \\ 0 & \text{if } \lambda_{1,q} \geq B \end{cases} \quad (36)$$

and \mathbf{x}_t is given in equation (35). In the above, $A = \mathbf{1}^T \Sigma^{-1}(\boldsymbol{\mu} - r\mathbf{1})$ and

$$B = \sqrt{(\boldsymbol{\mu} - r\mathbf{1})^T \Sigma^{-1}(\boldsymbol{\mu} - r\mathbf{1})}.$$

(b) If $r = \boldsymbol{\mu}^T \mathbf{x}_0$, then the optimal tail mean-variance portfolio consists of investing all wealth in the risk-free asset and investing a proportion y^* of wealth in the self-financing portfolio \mathbf{z} , where

$$y^* = \begin{cases} \frac{C - \lambda_{1,q}}{2\lambda\lambda_{2,q}C} & \text{if } \lambda_{1,q} \leq C \\ 0 & \text{if } \lambda_{1,q} \geq C \end{cases} \quad (37)$$

and \mathbf{z} is given in equation (17). In the above, $C = \sqrt{\boldsymbol{\mu}^T \Sigma^{-1}(\boldsymbol{\mu} - r\mathbf{1})}$.

Refer to Appendix A.2 for a proof of Theorem 2.

In the most common situation (described by case 1 in section 6.1), the risk-free rate is lower than the average return on the global minimum variance portfolio of risky securities. A situation where this did not hold would be unsustainable in the long term as it would result in a flight from risky assets into short-term Treasury bills, depressing risky security prices to the point where their expected return would rise relative to the risk-free rate.

Equation (36) in Theorem 2 is then seen to be a sensible proposition from a portfolio management point of view. The more risk-averse a tail mean-variance-minimizing investor is, the larger λ or $\lambda_{1,q}$ or $\lambda_{2,q}$ in equation (13) is likely to be, and therefore the more of his wealth, in relative terms, he will invest in the risk-free asset and the lower y^* is in equation (36). If he is so risk-averse that $\lambda_{1,q}$ is greater than the threshold represented by B , then he will invest only in the risk-free asset ($y^* = 0$).

The fact that there is such a threshold is peculiar to the tail mean-variance criterion. Indeed, there is no such threshold in classical mean-variance portfolio theory. To explain this feature, we resort to a visual interpretation in μ - σ space, where μ is the mean and σ is the standard deviation of portfolio returns. From equations (1) and (13), $\mu = \lambda\lambda_{2,q}\sigma^2 + \lambda_{1,q}\sigma - TMV$ and contours of equal tail mean-variance criterion are convex in μ - σ space. Contours of lower TMV lie above contours of higher TMV . An optimal portfolio is located at the point where the contour of lowest TMV is tangential to the mean-variance efficient frontier.

Now, iso- TMV contours have a slope that is not less than $\lambda_{1,q}$, since $d\mu/d\sigma = 2\lambda\lambda_{2,q}\sigma + \lambda_{1,q}$. In fact, at their point of intersection with the vertical axis ($\sigma = 0$), they have a slope of $\lambda_{1,q}$. On the other hand, the straight line mean-variance efficient frontier, joining the risk-free asset r to the tangency portfolio \mathbf{x}_t , has a slope of $(\boldsymbol{\mu}^T \mathbf{x}_t - r)/\sqrt{\mathbf{x}_t^T \boldsymbol{\Sigma} \mathbf{x}_t} = B$ (in case 1). Consequently, when $\lambda_{1,q} > B$, the iso- TMV contours are always steeper than the straight line efficient frontier, so that no point of tangency occurs. The portfolio that yields the lowest TMV occurs at the point where the highest iso- TMV contour intersects the efficient frontier, and is therefore the risk-free asset ($y^* = 0$). The lower $\lambda_{1,q}$ is, the less steep the contours of level TMV are; the tangency point between iso- TMV contours and efficient frontier then moves higher up along the straight line efficient frontier.

This is in sharp contrast to classical mean-variance optimization with the MV criterion ($g(\mathbf{x})$) in equation (14), where there is always a tangency point between contours of equal MV and the straight line efficient frontier (at least for investors who are not perfectly risk-averse and for whom $\tau < \infty$ in equation (14)). These contours are also convex but have a slope of zero when $\sigma = 0$, compared to a positive slope of $\lambda_{1,q}$ in the TMV case. In a loose sense, one might state that the TMV criterion generally implies a higher degree of risk aversion than the MV criterion.

Part (b) of Theorem 2 is also deserving of comment. This refers to the unusual case where the risk-free rate coincides exactly with the mean return on the minimum variance portfolio (case 3 in section 6.1), whereupon no tangency portfolio exists. This result was first derived by Merton (1972) in the classical mean-variance context. The efficient frontier of risky securities only, defined as \mathcal{E} in equation (18), has an asymptote that goes through the risk-free asset and has a slope of C (Huang and Litzenberger, 1988, p. 79). In the classical mean-variance case, optimal portfolios then lie on this asymptote, and consist of all wealth being lent risk-free, together with a self-financing investment in portfolio \mathbf{z} requiring no additional wealth. Again, if a TMV -minimizing investor is sufficiently risk-averse, the slope of his contours of level TMV will be high enough that they are never tangential to the asymptote, so that investment in \mathbf{z} does not occur ($y^* = 0$).

7. Conclusion

Asset return distributions are known to be heavy-tailed. Investors are also sensitive to extreme events in the capital and insurance markets, which occur in the tail of return distributions. The tail mean-variance criterion, which is considered in this paper, enables investors to select portfolios that take into account the risk of large but rare losses. It uses both the tail conditional expectation and the tail variance risk measures. If returns or losses are jointly elliptically distributed, then an analytic solution for the optimal portfolio under the tail mean-variance criterion is available. We characterized this optimal portfolio as a mean-variance efficient portfolio, even though the tail mean-variance criterion is not a translation-invariant positive-homogeneous risk measure. We also derived an explicit solution for the optimal portfolio, which improves on previous work by providing both insight and computational convenience, avoiding the partition, inversion and concatenation of large matrices. In the presence of a risk-free asset, the optimal portfolio remains mean-variance efficient. For a range of parameter values in the tail mean-variance criterion that imply greater aversion to losses beyond a certain threshold, it turns out to be optimal to invest only in the risk-free asset. The optimal portfolios that we calculated are simple enough that risk managers and portfolio managers can readily compute and implement them.

Appendix A. Proofs

Appendix A.1. Proof of Lemma 4

$F(0) > 0$ and $F(a) < 0$ and $F(x) \rightarrow +\infty$ as $x \rightarrow \pm\infty$, hence $F(x)$ has at least one real zero in $(0, a)$ and at least one real zero in (a, ∞) . Denote these two zeros by x_1 and x_2 respectively. There are two possible cases concerning the remaining two zeros, say x_3 and x_4 .

First, x_3 and x_4 may be complex conjugates, in which case x_1 and x_2 are indeed non-repeated real zeros. Secondly, x_3 and x_4 may both be real zeros, either of which could be

coincident with x_1 or x_2 .

$$F(x) = x^4 - 2x^3 + (a^2 + b - c)x^2 - 2bx + a^2b. \quad (\text{A.1})$$

The coefficient of x^3 in the above is negative $\Rightarrow \sum x_i < 0$. Therefore, at least one of x_3 and x_4 is negative. Further, the coefficient of x is also negative $\Rightarrow F'(0) < 0$. Since $F(0) > 0$, it follows that both of x_3 and x_4 are negative (possibly coincident). Hence, x_1 and x_2 are non-repeated real zeros.

This completes the proof of Lemma 4.

As an aside, we note that $a^2 + b - c < 0 \Rightarrow x_1 \in (0, a)$, $x_2 \in (a, \infty)$, and x_3 and x_4 are complex conjugate roots, by virtue of Descartes' rule of signs applied to the coefficients of $F(x)$ in equation (A.1). In fact, were the sign of $a^2 + b - c$ to be specified, then Sturm's theorem (Itô, 1993, p. 36) would be applicable.

Appendix A.2. Proof of Theorem 2

Let μ be the mean return on an overall portfolio (of risky and risk-free assets) and σ be the standard deviation of return on this overall portfolio. Then the tail mean-variance criterion may be written as

$$f(\mathbf{x}) = -\mu + \lambda_{1,q}\sigma + \lambda\lambda_{2,q}\sigma^2. \quad (\text{A.2})$$

Note that the proportion invested in the risk-free asset is $1 - y = 1 - \mathbf{1}^T \mathbf{x}$.

By an argument similar to the proof of Lemma 3, the portfolio that minimizes the tail mean-variance criterion in equation (A.2) must lie on the straight line efficient frontier: denote the mean return on the tail mean-variance optimal portfolio by μ^* , its standard deviation by σ^* , and the minimal value of the criterion by φ^* ; suppose that the tail mean-variance optimal portfolio does not lie on the efficient frontier; then there exists an overall portfolio, with mean return μ_1 and standard deviation σ_1 and tail mean-variance criterion φ_1 , that satisfies $\mu_1 \geq \mu^*$ and $\sigma_1 < \sigma^*$; but, by equation (A.2), this would imply that $\varphi_1 < \varphi^*$, which is a contradiction.

The tail mean-variance optimal portfolio is therefore a mean-variance efficient portfolio belonging to one of the three cases described in section 6.1. Given $\text{sgn}(\boldsymbol{\mu}^\top \mathbf{x}_0 - r)$, the optimal portfolio is fully specified by y .

Case 1. $y \geq 0$ and $A > 0$. Define μ_t to be the mean return on the tangency portfolio \mathbf{x}_t and σ_t to be its standard deviation of return. The mean return on an efficient portfolio in case 1 is $\mu = y\mu_t + (1 - y)r$ and the standard deviation of its return is $\sigma = y\sigma_t$, with $y \geq 0$. The tail mean-variance criterion of equation (A.2) is therefore $f(y) = -r - (\mu_t - r)y + \lambda_{1,q}\sigma_t y + \lambda\lambda_{2,q}\sigma_t^2 y^2$ and has at most one minimum, since $f''(y) = 2\lambda\lambda_{2,q}\sigma_t^2 > 0$.

From the tangency portfolio \mathbf{x}_t in equation (35), it is straightforward to find that $\mu_t = C^2/A$ and $\mu_t - r = B^2/A$ and $\sigma_t^2 = B^2/A^2$. Furthermore, $\sigma_t = B/|A| = B/A$, since $A > 0$ in case 1. Note that $B^2 = (\boldsymbol{\mu} - r\mathbf{1})^\top \Sigma^{-1}(\boldsymbol{\mu} - r\mathbf{1}) > 0$ and $B \in \mathbb{R}_{++}$ by virtue of positive definiteness of Σ (and hence of Σ^{-1}), and by the assumption of linear independence of $\boldsymbol{\mu}$ and $\mathbf{1}$. Therefore, $f(y) = -r - (B^2/A)y + \lambda_{1,q}(B/A)y + \lambda\lambda_{2,q}(B/A)^2 y^2$

The Kuhn-Tucker conditions are $y \geq 0$, $f'(y) \geq 0$ and $yf'(y) = 0$. The complementary slackness condition is satisfied *either* by $f'(y^*) = 0 \Leftrightarrow y^* = A(B - \lambda_{1,q})/2\lambda\lambda_{2,q}B$, with $y^* \geq 0$ provided that $\lambda_{1,q} \leq B$; *or* by $y^* = 0$, in which case $f'(0) \geq 0$ provided $\lambda_{1,q} \geq B$.

Case 2. $y \leq 0$ and $A < 0$. Constrained minimization proceeds as in case 1 above. The standard deviation of return on the tangency portfolio is now $\sigma_t = B/|A| = -B/A$, since $A < 0$ in case 2. But the standard deviation of return on the overall portfolio is $\sigma = -y\sigma_t = yB/A$ (since $y \leq 0$). Therefore, $f(y)$ is unchanged from case 1.

The Kuhn-Tucker conditions do change because of the non-positivity constraint: they are $y \leq 0$, $f'(y) \leq 0$ and $yf'(y) = 0$, and are satisfied *either* by $f'(y^*) = 0 \Leftrightarrow y^* = A(B - \lambda_{1,q})/2\lambda\lambda_{2,q}B$, with $y^* \leq 0$ provided that $\lambda_{1,q} \leq B$; *or* by $y^* = 0$, in which case $f'(0) \leq 0$ provided $\lambda_{1,q} \geq B$. (Recall that $A < 0$ in case 2.)

y^* for cases 1 and 2 may therefore be combined in equation (36), which captures both long and short positions in the tangency portfolio.

Case 3. Finally, in case 3, $y \geq 0$ but $A = 0$. All wealth is invested in the risk-free asset but the self-financing portfolio \mathbf{z} is also held and is worth a proportion y of wealth.

Define μ_z to be the mean return on portfolio \mathbf{z} and σ_z to be its standard deviation of return. Substituting $r = \boldsymbol{\mu}^T \mathbf{x}_0$ into equation (22) gives $\mu_z = \boldsymbol{\mu}^T \Sigma^{-1}(\boldsymbol{\mu} - r\mathbf{1}) = C^2$. From equation (24), $\sigma_z^2 = \mu_z = C^2$. Note that $C^2 = \boldsymbol{\mu}^T \Sigma^{-1}(\boldsymbol{\mu} - r\mathbf{1}) > 0$ and $C \in \mathbb{R}_{++}$ since $\mu_z = \sigma_z^2 > 0$. This is also confirmed by an application of the Cauchy-Schwarz inequality to the term in square brackets on the r.h.s. of equation (24).

The mean return on an efficient portfolio in case 3 is therefore $\mu = r + \mu_z y = r + C^2 z$, and the standard deviation of its return is $\sigma = y\sigma_t = yC$, with $y \geq 0$ and $C > 0$. The tail mean-variance criterion of equation (A.2) is therefore $f(y) = -r - C^2 y + \lambda_{1,q} C y + \lambda \lambda_{2,q} C^2 y^2$ and has at most one minimum, since $f''(y) = 2\lambda \lambda_{2,q} C^2 > 0$.

The Kuhn-Tucker conditions for a minimum in $f(y)$ wrt y subject to the non-negativity constraint on y lead again to two possible solutions: either $f'(y^*) = 0 \Leftrightarrow y^* = (C - \lambda_{1,q})/2\lambda \lambda_{2,q} C$ which is non-negative provided $\lambda_{1,q} \leq C$; or $y^* = 0$ with $f'(0) \geq 0$ provided that $\lambda_{1,q} \geq C$.

This completes the proof of Theorem 2.

References

- Acerbi, C., Tasche, D., 2002. On the coherence of expected shortfall. *Journal of Banking and Finance* 26(7), 1487–1503.
- Artzner, P., Delbaen, F., Eber, J.M., Heath, D., 1999. Coherent measures of risk. *Mathematical Finance* 9, 203–228.
- Chacko, G., Viceira, L.M., 2005. Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets. *Review of Financial Studies* 18(4), 1369–1402.
- Chiragiev, A. and Landsman, Z., 2007. Multivariate Pareto portfolios: TCE-based capital allocation and divided differences. *Scandinavian Actuarial Journal* 4, 261–280.
- Elton, E.J., Gruber, M.J., Brown, S.J., Goetzmann, W.N., 2011. *Modern Portfolio Theory and Investment Analysis*, 8th ed. John Wiley and Sons.

- Fang, K.T., Kotz, S., Ng, K.W., 1990. Symmetric Multivariate and Related Distributions. Chapman and Hall, London.
- Furman, E., Landsman, Z., 2005. Risk capital decomposition for a multivariate dependent gamma portfolio. *Insurance: Mathematics and Economics* 37(3), 635–649.
- Furman, E., Landsman, Z., 2006. Tail variance premium with applications for elliptical portfolio. *ASTIN Bulletin* 36(2), 433–462.
- Garlappi, L., Uppal, R., Wang, T. 2007. Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies* 20(1), 41–81.
- Huang, C., Litzenberger, R.H., 1988. Foundations for Financial Economics. Prentice Hall, Englewood Cliffs, New Jersey.
- Itô, K. (ed.), 1993. Encyclopedic Dictionary of Mathematics, 2nd ed. MIT Press, Massachusetts.
- Jorion, P., 1996. Risk²: Measuring the risk in Value at Risk. *Financial Analysts Journal* 52(6), 47–56.
- Kan, R., Zhou, G., 2007. Optimal Portfolio Choice with Parameter Uncertainty. *Journal Of Financial And Quantitative Analysis* 42(3), 621-656.
- Karatzas, I., Shreve, S.E., 1991. Brownian Motion and Stochastic Calculus. Springer, Berlin.
- Kennedy, D., 2010. Stochastic Financial Models. Chapman and Hall/CRC Press, Boca Raton, Florida.
- Landsman, Z.M., 2010. On the Tail Mean-Variance optimal portfolio selection. *Insurance: Mathematics and Economics* 46, 547–553.
- Landsman, Z.M., Valdez, E.A., 2003. Tail conditional expectations for elliptical distributions. *North American Actuarial Journal* 7(4), 55–71.

- Landsman, Z.M., Valdez, E.A., 2005. Tail conditional expectations for exponential dispersion models. *ASTIN Bulletin* 35(1), 189–209.
- Lee, C.-F., Lee, C. (ed.), 2006. *Encyclopedia of Finance*. Springer.
- Markowitz, H.M., 1952. Portfolio selection. *Journal of Finance* 7, 77–91.
- McNeil, A.J., Frey, R., Embrechts, P., 2005. *Quantitative Risk Management*. Princeton University Press.
- Merton, R.C., 1969. Lifetime portfolio selection under uncertainty: the continuous-time case. *Review of Economics and Statistics* 51, 247-257.
- Merton, R.C., 1971. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory* 3, 373-413.
- Merton, R.C., 1972. An analytic derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis* 7(4), 1851–1872.
- Neumark, S., 1965. *Solution of Cubic and Quartic Equations*. Pergamon Press, Oxford.
- Owen, J., Rabinovitch, R., 1983. On the class of elliptical distributions and their applications to the theory of portfolio choice. *Journal of Finance* 38(3), 745–752.
- Panjer, H.H., 2002. *Measurement of risk, solvency requirements, and allocation of capital within financial conglomerates*. University of Waterloo, Canada.
- Panjer, H.H. et al., 1998. *Financial Economics*. The Actuarial Foundation, Schaumburg, Illinois.
- Sharpe, W.F., Alexander, G.J, Bailey, J.V., 1995. *Investments*, 5th ed. Prentice Hall, Englewood Cliffs, New Jersey.
- Tu, J., Zhou, G., 2004. Data-generating process uncertainty: what difference does it make in portfolio decisions? *Journal of Financial Economics* 72, 385-421.

Yang, H., Zhang, L., 2005. Optimal investment for insurer with jump-diffusion risk process. *Insurance: Mathematics and Economics* 37, 615-634.

Zhao, H., Rong, X., 2012. Portfolio selection problem with multiple risky assets under the constant elasticity of variance model. *Insurance: Mathematics and Economics* 50(1), 179–190.