9-2012

Portfolio value-at-risk optimization for asymmetrically distributed asset returns

Joel Weiqiang GOH
National University of Singapore

Kian Guan LIM
Singapore Management University, kglim@smu.edu.sg

Melvyn SIM
National University of Singapore

Weina ZHANG
National University of Singapore
DOI: https://doi.org/10.1016/j.ejor.2012.03.012

Portfolio Value-at-Risk Optimization for Asymmetrically Distributed Asset Returns

Joel Goh* Kian Guan Lim† Melvyn Sim‡ Weina Zhang§

July 2012

*Department of Decision Sciences, NUS Business School, National University of Singapore. Singapore 119245. Email: joelgoh@nus.edu.sg.

†Finance and Quantitative Finance Unit, Singapore Management University. Singapore 178899. Email: kgl@smu.edu.sg.

‡Department of Decision Sciences, NUS Business School, National University of Singapore. Singapore 119245. Affiliated with NUS Risk Management Institute and Singapore-MIT Alliance. Email: melvynsim@nus.edu.sg. The research of the author is supported by Singapore-MIT Alliance and NUS academic research grant R-314-000-068-122.

§Corresponding Author. Department of Finance, NUS Business School, National University of Singapore. Singapore 119245. Affiliated with NUS Risk Management Institute. Tel: +65-65168120. Fax: +65-67792083. Email: weina@nus.edu.sg.

Portfolio Value-at-Risk Optimization for Asymmetrically Distributed Asset Returns

July 2012

Forthcoming in *European Journal of Operational Research*

Abstract

We propose a new approach to portfolio optimization by separating asset return distributions into positive and negative half-spaces. The approach minimizes a newly-defined Partitioned Value-at-Risk (PVaR) risk measure by using half-space statistical information. Using simulated data, the PVaR approach always generates better risk-return tradeoffs in the optimal portfolios when compared to traditional Markowitz mean-variance approach. When using real financial data, our approach also outperforms the Markowitz approach in the risk-return tradeoff. Given that the PVaR measure is also a robust risk measure, our new approach can be very useful for optimal portfolio allocations when asset return distributions are asymmetrical.

Keywords: Risk Management · Asymmetric Distributions · Partitioned Value-at-Risk · Portfolio Optimization · Robust Risk Measures

Published in European Journal of Operational Research, Volume 221, Issue 2, September 2012, Pages 397-406
https://doi.org/10.1016/j.ejor.2012.03.012

1 Introduction

In the pioneering work by Markowitz (1952) [33], an optimal portfolio allocation is obtained by mean-variance optimization. However, researchers have found that two key assumptions made in the Markowitz approach are often violated in real data and experience. First, Tobin (1958) [43] and Chamberlain (1983) [12] show that the mean-variance optimization is appropriate to capture the tradeoff between risk and return only if the distribution of returns is elliptically symmetric. Once this assumption is violated, the mean and variance will not be sufficient statistics for investors to make optimal asset allocation decisions. Many empirical studies have reported evidence of asymmetries and large kurtosis in asset return distributions. For example, Mandelbrot (1963) [34], Fama (1965) [22] and Kon (1984) [31] find that extreme returns occur more frequently than would be under normal distribution. Simkowitz and Beedles (1980) [41] find that individual stock returns are positively skewed. Alles and Kling (1994) [1] show that both equity and bond indices have negative skewness. Theodossiou (1998) [42] reports skewness in international stock returns, foreign exchange rates, and commodity returns. Bekaert, Erb, Harvey and Viskanta (1998) [7] and Erb, Harvey and Viskanta (1999) [21] also find that emerging markets equities and bonds display skewness. More recently, Ang and Chen (2002)[2] document that equity portfolios are more correlated for downside moves than upside moves and the asymmetries in the data reject the null hypothesis of multivariate normal distributions.

Second, the Markowitz approach assumes that people do not differentiate positive (upside risk) and negative (downside risk) deviations from the mean. There is a growing body of research evidence that investors have asymmetrical risk attitudes towards upside and downside risks. Arditti (1967) [3] and Scott and Horvath (1980) [39] show that investors prefer positive skewness to negative skewness. Investors such as fund managers might derive very different utility from meeting versus falling-short of the target returns. Mitton and Vorkink (2007) [35] find that investors in a sample of discount brokerage transactions hold less optimal portfolios with lower Sharpe ratios but positive skewness. Given the findings, Barberis and Huang (2007) [6] and Brunnermeier, Gollier and Parker (2007) [11] propose a theoretical framework for the tradeoff between diversification and skewness in portfolio allocation, in a dramatic contrast to the initial Markowitz approach. Kerstens, Mounir and Woestyne (2011) [30] also propose an improved Markowitz approach by solving for optimal portfolios under mean-variance-skewness efficient frontier.

Better risk-return tradeoffs in the presence of asymmetric distributions are documented by Harvey and Siddique (2000) [25]. They show that systematic skewness commands an average risk premium of

3.6 percent per year. Simkowitz and Beedles (1978) [40] and Conine and Tamarkin (1981) [16] also claim that though diversification can change skewness exposure, the remaining idiosyncratic skewness is relevant in asset pricing. Therefore, the portfolio optimization under asymmetric distribution is a significant topic for research.

Robust optimization literature is developed to improve the estimation errors and handle the unstable nature of asset return distributions. The key idea is to find a portfolio that maximizes (or minimizes) the expected utility (disutility) value in the midst of infinitely many possible ambiguous distributions of the returns fitting the given mean-variance estimations. Hence, lower partial moments, value-at-risk (VaR), conditional value-at-risk (CVaR), or worst-case mean-covariance value-at-risk (WVaR) approaches are developed and improved to find optimal portfolio allocations. However, many recent improvements of robust optimization are still focusing on the conventional mean-variance framework (e.g. Chen, He, and Zhang, 2011 [15]; Delage and Ye, 2010 [17]; DeMiguel and Nogales, 2009 [18]).

In this paper, we study how to make optimal portfolio allocations in the presence of asymmetry in asset returns. We define a new robust risk measure,termed Partitioned Value-at-Risk (PVaR), which extends the WVaR risk measure proposed by El Ghaoui, Oks, and Oustry (2003) [20], to incorporate asymmetry in the distributions of asset returns. We show that when asset returns are asymmetrically distributed, PVaR-minimized portfolio allocations require lower capital reserves for the same percentage of tail risk as compared to WVaR-minimized portfolios. This advantage is especially useful for optimal portfolio allocation during market downturns when fat-tailed and skewed return distributions are often observed.

Morever, PVaR has appealing properties as a risk measure. When augmented with the support of asset returns, PVaR becomes a coherent risk measure, satisfying the coherence axioms of Artzner et al. (1999) [4], unlike conventional risk measures such as variance and Value-at-Risk (VaR). PVaR is also a robust measure of risk, as it is defined over a family of probability distributions, whereas the classical VaR measure is defined over a single probability distribution. Finally, the PVaR measure does not require the strong assumptions made in the Asymmetry-Robust VaR (AR-VaR) (Natarajan et al., 2008 [36]) and is easy to implement empirically.

The rest of this paper is organized as the follows. Section 2 provides a background discussion of the robust portfolio optimization literature to motivate our PVaR approach. Section 3 presents the definition of PVaR and the portfolio optimization approach based on it. Section 4 reports three sets of empirical results on the performance of the PVaR approach in comparison with the conventional mean-variance approach, using both simulated and actual stock returns data that are widely available.

Section 5 concludes.

Notations: Throughout this paper, we denote a random variable, $\tilde{r}$, with the tilde sign. Bold face lower case letters such as $\boldsymbol{x}$ represent vectors. We use the prime symbol " $'$ " to denote transpose and the functionals $\mathbb{P}(\cdot)$ and $\mathrm{E}(\cdot)$ to denote the probability of an event and the expectation of a random variable respectively. In particular, $\mathrm{E}_{\mathbb{P}}(\cdot)$ denotes expectation under the probability distribution $\mathbb{P}$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space and let $\tilde{r}_1, \ldots, \tilde{r}_n$ be $\mathbb{P}$-measurable random returns of n risky assets. The support of asset returns is denoted by $\mathcal{W} = \{\tilde{\boldsymbol{r}}(\omega) \ : \ \omega \in \Omega\} \subseteq \Re^n$. The vector of mean asset returns is denoted by $\bar{\boldsymbol{r}}$. We define the universe of portfolio returns as follows

$$\mathcal{V} = \{\tilde{v} \ : \ \tilde{v} = y + x_1\tilde{r}_1 + \ldots + x_n\tilde{r}_n \quad \text{where} \quad \exists(y, x_1, \ldots, x_n) \in \Re^{n+1}\}.$$

The set $\mathcal{V}$ encompasses all possible portfolio returns generated by the n risky assets and riskless cash holding y. Any $\tilde{v} \in \mathcal{V}$ can be expressed as $\tilde{v} = y + \tilde{\boldsymbol{r}}'\boldsymbol{x}$ for some $(y, \boldsymbol{x}) \in \Re^{n+1}$. We will use the notation $\tilde{v} \geq \tilde{w}$ for $\tilde{v},\ \tilde{w} \in \mathcal{V}$ to represent state-wise dominance, i.e., $\tilde{v}(\omega) \geq \tilde{w}(\omega)$ for all $\omega \in \Omega$.

2 Existing VaR Measures in Portfolio Optimization

In this section, we first review some common approaches in portfolio optimization. Specifically, we review the concept of VaR, Conditional VaR, and Worst-Case Mean-Covariance VaR.

2.1 Classical VaR Approach

Under the international BASEL Accord, VaR has become a popular risk measure used by both regulated banks as well as investment practitioners. In the conventional setting, the VaR measure is used to gauge the overall risk of the underlying portfolios when the aggregated portfolio return is deemed to be generated by a single probability distribution. Given a random variable, $\tilde{v} \in \mathcal{V}$ denoting random portfolio returns, the VaR measure is defined as follows:

$$\mathrm{VaR}_{1-\epsilon}(\tilde{v}) \stackrel{\Delta}{=} \min\{a \ : \ \mathbb{P}(a + \tilde{v} \geq 0) \geq 1 - \epsilon\},$$

which is the $(1-\epsilon)$-quantile of the portfolio return, or alternatively, that ϵ is the confidence level. This quantity, $a > 0$, can also be interpreted as the smallest amount of capital necessary to add to $\tilde{v}$ to ensure that the augmented portfolio $\tilde{v} + a$ is positive with probability at least $1 - \epsilon$. Then, the minimum VaR portfolio optimization problem can be formulated as

$$\begin{array}{ll} \min_{\boldsymbol{x}} & \mathrm{VaR}_{1-\epsilon}\left(\tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \\ \text{s.t.} & \boldsymbol{x} \in X \end{array}$$

where X represents the feasible space of asset weights. For simplicity, we consider the set

$$X = \left\{ \boldsymbol{x} \ : \ \boldsymbol{x}'\boldsymbol{e} = 1, \bar{\boldsymbol{r}}'\boldsymbol{x} = \tau \right\},$$

where $\boldsymbol{e}$ denotes a vector of ones, X is a set of normalized portfolio allocations, and the target expected return of the portfolio is τ. The feasible set does not impose any short-selling restrictions.

The problem of minimizing over VaR measure is a classical chance-constraint optimization, which is first proposed by Charnes et al. (1958) [14]. There are some fundamental issues with optimizing over the VaR measure, the first being its computational tractability. Unlike the mean-variance approach, the optimal VaR portfolio is hard to compute unless the distribution of returns is assumed to be normal or lognormal. Duffie and Pan (1997) [19], and Jorion (2000) [29] discuss this issue in their papers. Gaivoronski and Pflug (2005) [24] and Larsen et al. (2002) [32] have proposed empirical procedures to optimizing sample VaR. The minimum empirical VaR (EVaR) portfolio can be obtained by solving the following mixed-integer (MIP) optimization problem:

$$\begin{array}{lll} (\mathrm{EVaR}) & \min_{\boldsymbol{x},\boldsymbol{y},\gamma} & \gamma \\ & \text{s.t.} & \gamma + (\boldsymbol{r}^i)'\boldsymbol{x} \geq -Ky_i, \quad i = 1, \ldots, T \\ & & \boldsymbol{y}'\boldsymbol{e} = \lfloor \epsilon T \rfloor \\ & & \boldsymbol{y} \in \{0,1\}^T \\ & & \bar{\boldsymbol{r}}'\boldsymbol{x} = \tau \\ & & \boldsymbol{e}'\boldsymbol{x} = 1 \end{array} \tag{1}$$

for some large constant K and a sample of T vectors of realized asset returns, $\boldsymbol{r}^1, \ldots, \boldsymbol{r}^T$. $\lfloor \cdot \rfloor$ denotes the integral value of its argument. To check that the optimal γ in the above problem is indeed the optimal VaR over the T empirical returns, notice that in the optimal solution, $y_i = 1$ for the highest $\lfloor \epsilon T \rfloor$ realizations of portfolio losses (negative returns), and $y_i = 0$ for the remaining realizations of portfolio losses. For the set of indices i for which $y_i = 0$, the constraints

$$\gamma \geq -(\boldsymbol{r}^i)'\boldsymbol{x}$$

ensure that the optimal γ will take the lowest value that is higher than the lowest $(1-\epsilon)T$ of all losses. This is exactly the definition of sample VaR at a confidence level of ϵ. For the remaining indices i (those for which $y_i = 1$), the constraints

$$\gamma + (\boldsymbol{r}^i)'\boldsymbol{x} \geq -Ky_i$$

are not binding.

Unfortunately, even under strong assumptions that the empirical asset returns are stationary and independent across time, the relation between the EVaR and its true VaR has not been clear. It is conceivable that the number of observations needed to achieve any meaningful confidence of reliability is of the order of $1/\epsilon$. In practice, the choice of the reliability parameter is often arbitrary, e.g. $\epsilon = 1\%$ or 5%. Indeed one may require an unrealistically large sample size of stationary and independent returns before the optimal EVaR might be reasonably close to the true VaR. As a result, the computation of EVaR may not scale well with portfolio variability. Moreover, solving EVaR in the context of mathematical programming is conceptually and computationally intractable mainly due to the lack of convexity.

2.2 Conditional VaR (CVaR) Approach

Apart from the difficulties encountered in applying VaR empirically, Artzner et al. (1999) [4] argue against the VaR measure by stating that it does not have all the desirable properties of a risk measure. They introduce the idea of "coherent risk" to characterize desirable properties of risk measures. A coherent risk measure, ρ, is a functional defined on $\mathcal{V}$ that satisfies the following four axioms.

(i) Translation invariance: For all $\tilde{v} \in \mathcal{V}$ and $a \in \Re$, $\rho(\tilde{v} + a) = \rho(\tilde{v}) - a$.

(ii) Subadditivity: For all $\tilde{v}, \tilde{w} \in \mathcal{V}$, $\rho(\tilde{v} + \tilde{w}) \leq \rho(\tilde{v}) + \rho(\tilde{w})$.

(iii) Positive homogeneity: For all $\tilde{v} \in \mathcal{V}$, and $\lambda \geq 0$, $\rho(\lambda\tilde{v}) = \lambda\rho(\tilde{v})$.

(iv) Monotonicity: For all $\tilde{v}, \tilde{w} \in \mathcal{V}$ such that $\tilde{v} \geq \tilde{w}$, $\rho(\tilde{v}) \leq \rho(\tilde{w})$.

The translation invariance axiom ensures that $\rho(\tilde{v} + \rho(\tilde{v})) = 0$, so that the risk associated with the portfolio return $\tilde{v}$ after compensation with $\rho(\tilde{v})$ of capital is nullified. The subadditivity axiom states that the risk associated with the sum of two allocations is not more than the sum of their individual risk. In other words, it is never worse off to diversify the portfolio. The classical VaR fails to satisfy this axiom. The positive homogeneity axiom implies that risk measure scales proportionally with its size. The final monotonicity axiom rules out common risk measures such as standard deviation.

A popular example of such a coherent risk measure is Conditional VaR (CVaR) as discussed in Rockafellar and Uryasev (2000, 2002) [37, 38]. The CVaR measure can be written as

$$\mathrm{CVaR}_{1-\epsilon}(\tilde{v}) \triangleq \min_{a\in\Re}\left(a + \frac{1}{\epsilon}\mathrm{E}(\max\{-\tilde{v} - a, 0\})\right).$$

It is well known that CVaR is the best coherent risk measure that approximates VaR (see for instance, Föllmer and Schied (2004) [23]) and that direct optimization of CVaR is a convex optimization problem. However, evaluation of CVaR involves precise knowledge of the underlying distributions and the need to perform multi-dimensional integral, which is typically computationally prohibitive beyond the fourth dimension. Nevertheless, the minimum empirical CVaR (ECVaR) portfolio optimization problem is a linear optimization problem as follows:

$$\begin{array}{llll} \mathrm{ECVaR} & \min\limits_{\boldsymbol{x},\boldsymbol{y},a} & a + \frac{1}{\epsilon T}\sum_{t=1}^{T} y_t & \\ & \text{s.t.} & y_t \geq -\boldsymbol{r_t}'\boldsymbol{x} - a & t = 1,\ldots,T \\ & & y_t \geq 0, & t = 1,\ldots,T \\ & & \bar{\boldsymbol{r}}'\boldsymbol{x} = \tau & \\ & & \boldsymbol{e}'\boldsymbol{x} = 1. & \end{array} \qquad (2)$$

However, as in the case of EVaR, the relation between the optimal ECVaR and its true CVaR is not well understood. The number of observations needed to achieve any meaningful confidence of reliability is of the order of $1/\epsilon$ and hence, the computation of CVaR may not scale well with portfolio variability. This observation agrees with empirical findings in Yamai and Yoshiba (2002) [44] that estimation of CVaR can be very unstable for relatively small samples. More recently, Zhu and Fukushima (2009) [45] and Huang, Zhu, Fabozzi, and Fukushima (2010) [26] propose further improvements in the CVaR approach to select optimal portfolios, but they still do not consider asymmetrical distributions explicitly.

2.3 Worst-Case Mean-Covariance VaR (WVaR) Approach

The classical VaR and CVaR measures are evaluated under an assumed probability distribution of asset returns. In reality, it is impossible to determine all the outcomes of the underlying asset returns and their associated probability distributions. Nevertheless, we could still obtain the empirical descriptive statistics of the asset returns such as the mean, covariance and so on. Motivated by developments in robust optimization in addressing the ambiguity in return distributions, we consider the worst case impact of the VaR over a family of probability distributions $\mathbb{F}$ as follows

$$\mathrm{VaR}_{1-\epsilon,\mathbb{F}}(\tilde{v}) \triangleq \min\left\{a \ : \ \inf_{\mathbb{P}\in\mathbb{F}} \mathbb{P}(a + \tilde{v} \geq 0) \geq 1 - \epsilon\right\}.$$

Such a family of distributions could, for instance, be one that generates the same descriptive statistics of asset returns observed from empirical data. The rationale of considering a family of distribution is one of robustness over the assumption of probability distributions. Instead of of assuming an underlying joint distribution of the asset returns and then estimating the parameters from empirical data, we specify a family of distributions that exhibits certain traits that we can observe from analyzing empirical asset returns.

In general, it is difficult to characterize the worst-case VaR over a family of probability distributions, as discussed in Bertsimas and Popescu (2005) [9]. Nevertheless, there are tractable instances. El Ghaoui et al. (2003) [20] considered a family of probability distributions that preserves the mean and covariance of the underlying asset returns as follows,

$$\mathbb{F}(\bar{\boldsymbol{r}}, \boldsymbol{\Sigma}) = \left\{\mathbb{P} \ : \ \mathrm{E}_{\mathbb{P}}(\tilde{\boldsymbol{r}}) = \bar{\boldsymbol{r}}, \mathrm{E}_{\mathbb{P}}((\tilde{\boldsymbol{r}} - \bar{\boldsymbol{r}})(\tilde{\boldsymbol{r}} - \bar{\boldsymbol{r}})') = \boldsymbol{\Sigma}\right\}.$$

They propose the worst-case mean-covariance VaR (WVaR) defined on $\mathcal{V}$ as follows

$$\mathrm{WVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) = -y - \bar{\boldsymbol{r}}'\boldsymbol{x} + \sqrt{\frac{1-\epsilon}{\epsilon}}\|\boldsymbol{\Sigma}^{1/2}\boldsymbol{x}\|_2.$$

Using results of Bertsimas and Popescu (2005) [9], El Ghaoui et al. (2003) [20] established that there exists a set Ω and returns $\tilde{r}_1, \ldots, \tilde{r}_n$ on Ω such that

$$\mathrm{WVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) = \mathrm{VaR}_{1-\epsilon, \mathbb{F}(\bar{\boldsymbol{r}}, \boldsymbol{\Sigma})}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right).$$

The corresponding minimum WVaR portfolio optimization problem becomes

$$\begin{array}{lll} (\mathrm{WVaR}) & \min\limits_{\boldsymbol{x}} & -\bar{\boldsymbol{r}}'\boldsymbol{x} + \sqrt{\frac{1-\epsilon}{\epsilon}}\|\boldsymbol{\Sigma}^{1/2}\boldsymbol{x}\|_2 \\ & \text{s.t.} & \bar{\boldsymbol{r}}'\boldsymbol{x} = \tau \\ & & \boldsymbol{e}'\boldsymbol{x} = 1. \end{array} \tag{3}$$

Note that the constraint $\bar{\boldsymbol{r}}'\boldsymbol{x} = \tau$ ensures that the portfolio with minimal WVaR is equivalent to the one with minimal variance, which is essentially the solution of Markowitz portfolio optimization problem. Therefore, the WVaR is criticized as a symmetric risk measure whereas VaR is a downside risk measure. More recently, Natarajan et al. (2008) [36] propose a computationally tractable modified VaR measure - Asymmetry-Robust VaR (AR-VaR), that is also coherent and addresses distributional asymmetry. However, AR-VaR requires asset returns to be affinely dependent on a set of stochastically independent random factors, which may not be easily identified from empirical data.

3 Partitioned Value-at-Risk (PVaR) Approach

In this section, we propose a new tractable approach to finding optimal portfolios by extending the WVaR measure (El Ghaoui et al.(2003) [20]) in the previous section. We accomplish this by extracting more statistical information from the distribution of the assets. Essentially, we partition the space of random returns into positive and negative half-spaces. This is a natural partition because the half-spaces map directly into profit and loss respectively in the portfolio outcomes. Using the statistical information on these two half-spaces, such as the mean and covariance matrix, we define a Partitioned VaR (PVaR) risk measure. Furthermore, after adding information about the support of the asset returns, this risk measure becomes coherent.

We partition the returns, $\tilde{\boldsymbol{r}}$ into its positive and negative random vectors $(\tilde{\boldsymbol{r}}^1, \tilde{\boldsymbol{r}}^2) \in \Re^{2n}$, in which $\tilde{r}_i^1 = \max\{\tilde{r}_i, 0\}$ and $\tilde{r}_i^2 = \min\{\tilde{r}_i, 0\}$. Hence, $\tilde{\boldsymbol{r}} = \tilde{\boldsymbol{r}}^1 + \tilde{\boldsymbol{r}}^2$. Clearly, $\tilde{\boldsymbol{r}}^1$ and $\tilde{\boldsymbol{r}}^2$ isolate the statistical information whenever returns are positive and negative respectively. We let $(\bar{\boldsymbol{r}}^1, \bar{\boldsymbol{r}}^2) \in \Re^{2n}$ and $\hat{\boldsymbol{\Sigma}} \in \Re^{2n\times 2n}$ be respectively the mean and covariance of $(\tilde{\boldsymbol{r}}^1, \tilde{\boldsymbol{r}}^2)$. We call $(\bar{\boldsymbol{r}}^1, \bar{\boldsymbol{r}}^2, \hat{\boldsymbol{\Sigma}})$ the partitioned statistics of the asset returns.

We notice that $\bar{\boldsymbol{r}}^1$ (respectively $\bar{\boldsymbol{r}}^2$) is a vector with non-negative (respectively non-positive) components and we have

$$\bar{\boldsymbol{r}} = \bar{\boldsymbol{r}}^1 + \bar{\boldsymbol{r}}^2.$$

Note that the variance of $(\tilde{\boldsymbol{r}}^1, \tilde{\boldsymbol{r}}^2)$ is a $2n$ by $2n$ positive semidefinite matrix given by

$$\hat{\boldsymbol{\Sigma}} = \mathrm{var}\left(\begin{pmatrix} \tilde{\boldsymbol{r}}^1 \\ \tilde{\boldsymbol{r}}^2 \end{pmatrix}\right) = \begin{bmatrix} \mathrm{var}(\tilde{\boldsymbol{r}}^1) & \mathrm{cov}(\tilde{\boldsymbol{r}}^1, \tilde{\boldsymbol{r}}^2) \\ \mathrm{cov}(\tilde{\boldsymbol{r}}^2, \tilde{\boldsymbol{r}}^1) & \mathrm{var}(\tilde{\boldsymbol{r}}^2) \end{bmatrix},$$

where

$$\mathrm{cov}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) \stackrel{\Delta}{=} \mathrm{E}\left((\tilde{\boldsymbol{x}} - \mathrm{E}(\tilde{\boldsymbol{x}}))(\tilde{\boldsymbol{y}} - \mathrm{E}(\tilde{\boldsymbol{y}}))'\right)$$

and

$$\mathrm{var}(\tilde{\boldsymbol{x}}) \stackrel{\Delta}{=} \mathrm{cov}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{x}}).$$

Moreover,

$$\mathrm{var}(\tilde{\boldsymbol{r}}'\boldsymbol{x}) = \boldsymbol{x}'\boldsymbol{\Sigma}\boldsymbol{x} = \mathrm{var}(\tilde{\boldsymbol{r}}^{1\prime}\boldsymbol{x} + \tilde{\boldsymbol{r}}^{2\prime}\boldsymbol{x}) = (\boldsymbol{x}'\ \boldsymbol{x}')\hat{\boldsymbol{\Sigma}}\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{x} \end{pmatrix}. \tag{4}$$

Hence,

$$\boldsymbol{\Sigma} = \mathrm{var}(\tilde{\boldsymbol{r}}^1) + \mathrm{cov}(\tilde{\boldsymbol{r}}^1, \tilde{\boldsymbol{r}}^2) + \mathrm{cov}(\tilde{\boldsymbol{r}}^2, \tilde{\boldsymbol{r}}^1) + \mathrm{var}(\tilde{\boldsymbol{r}}^2).$$

Therefore, the mean and covariance of the asset returns can be derived from the partitioned statistics of the asset returns. We can therefore define the family of distributions with partitioned statistics:

$$\hat{\mathbb{F}}(\bar{\boldsymbol{r}}^1, \bar{\boldsymbol{r}}^2, \hat{\boldsymbol{\Sigma}}) \triangleq \left\{ \mathbb{P} \in \mathbb{F}(\bar{\boldsymbol{r}}, \boldsymbol{\Sigma}) \;:\; \mathrm{E}_{\mathbb{P}}\left(\begin{pmatrix} \tilde{\boldsymbol{r}}^1 \\ \tilde{\boldsymbol{r}}^2 \end{pmatrix} \right) = \begin{pmatrix} \bar{\boldsymbol{r}}^1 \\ \bar{\boldsymbol{r}}^2 \end{pmatrix}, \mathrm{E}_{\mathbb{P}}\left(\begin{pmatrix} \tilde{\boldsymbol{r}}^1 - \bar{\boldsymbol{r}}^1 \\ \tilde{\boldsymbol{r}}^2 - \bar{\boldsymbol{r}}^2 \end{pmatrix} \begin{pmatrix} \tilde{\boldsymbol{r}}^1 - \bar{\boldsymbol{r}}^1 \\ \tilde{\boldsymbol{r}}^2 - \bar{\boldsymbol{r}}^2 \end{pmatrix}' \right) = \hat{\boldsymbol{\Sigma}} \right\}.$$

We define the PVaR on $\mathcal{V}$, over the family of distributions $\hat{\mathbb{F}}$ as follows

$$\mathrm{PVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \triangleq -y - \bar{\boldsymbol{r}}'\boldsymbol{x} + \min_{\boldsymbol{s},\boldsymbol{t} \geq \boldsymbol{0}} \left\{ \sqrt{\frac{1-\epsilon}{\epsilon}} \left\| \hat{\boldsymbol{\Sigma}}^{1/2} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{s} \\ \boldsymbol{x} + \boldsymbol{t} \end{pmatrix} \right\|_2 + \bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s} - \bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t} \right\}.$$

We note that the auxilliary variables $\boldsymbol{s}$ and $\boldsymbol{t}$ are simply variables to be optimized over and do not carry any particular meaning.

Hence, the minimum-PVaR portfolio optimization problem is equivalent to solving the following quadratic optimization problem,

$$\begin{array}{lll} (\mathrm{PVaR}) & \displaystyle\min_{\boldsymbol{x},\boldsymbol{s},\boldsymbol{t}} & -\bar{\boldsymbol{r}}'\boldsymbol{x} + \sqrt{\tfrac{1-\epsilon}{\epsilon}} \left\| \hat{\boldsymbol{\Sigma}}^{1/2} \binom{\boldsymbol{x}-\boldsymbol{s}}{\boldsymbol{x}+\boldsymbol{t}} \right\|_2 + \bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s} - \bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t} \\ & \text{s.t.} & \bar{\boldsymbol{r}}'\boldsymbol{x} = \tau \\ & & \boldsymbol{e}'\boldsymbol{x} = 1 \\ & & \boldsymbol{s}, \boldsymbol{t} \geq \boldsymbol{0}. \end{array} \tag{5}$$

Levering on the fact that partitioned statistics of the asset returns carry more information in addition to the first and second moments, we next show that the PVaR measure provides a better bound to VaR than WVaR.

Theorem 1 *Suppose $\tilde{\boldsymbol{r}}$ has partitioned statistics, $(\bar{\boldsymbol{r}}^1, \bar{\boldsymbol{r}}^2, \hat{\boldsymbol{\Sigma}})$, then*

$$\mathrm{VaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \leq \mathrm{PVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \leq \mathrm{WVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right)$$

for all $y + \tilde{\boldsymbol{r}}'\boldsymbol{x} \in \mathcal{V}$.

Proof : Refer to online supplementary materials available at electronic version of this paper at *http://www.sciencedirect.com*.

Since the first inequality above holds for all returns $\tilde{\boldsymbol{r}}$ having probability distributions $\mathbb{P}$ in the family $\hat{\mathbb{F}}(\bar{\boldsymbol{r}}^1, \bar{\boldsymbol{r}}^2, \hat{\boldsymbol{\Sigma}})$, the optimized portfolio weights x^* obtained from solving Model (5) is, by design, robust against mis-specifications of $\mathbb{P}$ within this family. In contrast, the standard (non-robust) VaR measure is sensitive to the implicit or explicit specification of $\mathbb{P}$.

Comparing PVaR with the WVaR risk measure, which is also distributionally robust, PVaR provides a tighter bound to VaR than WVaR. This is because in the PVaR optimization, asymmetry is explicitly modeled and exploited, while WVaR disregards any asymmetry in distribution. Therefore, PVaR can be viewed as a tractable conic approximation of the worst case VaR using the new set of information called partitioned statistics, which contains more information than mean and covariance. It is a natural extension to El Ghaoui et al.(2003) WVaR, which only includes mean and covariance information and hence insensitive to distributional asymmetry.

Nevertheless, similar to the WVaR, the PVaR measure is not a coherent risk measure as it violates the axioms on monotonicity. We will first show in the next proposition that PVaR satisfies the other axioms of coherent risk measure.

Proposition 1 *The PVaR measure satisfies the axioms of translation invariance, subadditivity and positive homogeneity.*

Proof : Refer to online supplementary materials available at electronic version of this paper at *http://www.sciencedirect.com.*

Although the PVaR measure does not satisfy the axiom on monotonicity, we can easily fix this problem. The monotonicity axiom implicity implies that we can specify the support of asset returns, $\mathcal{W}$, so that one can ascertain whether $\tilde{v} \geq \tilde{w}$ holds. Therefore, to construct a coherent version of the PVaR, we just need the support of the asset returns, $\mathcal{W}$. Hence, we can define the Coherent Partitioned VaR (CPVaR) as follows

$$\text{CPVaR}_{1-\epsilon}\left(y+\tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \triangleq -y-\bar{\boldsymbol{r}}'\boldsymbol{x}+\min_{\boldsymbol{s},\boldsymbol{t}\geq\boldsymbol{0},\boldsymbol{w}}\left\{\sqrt{\frac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}\boldsymbol{x}-\boldsymbol{w}-\boldsymbol{s}\\ \boldsymbol{x}-\boldsymbol{w}+\boldsymbol{t}\end{pmatrix}\right\|_2+\bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s}-\bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t}+\bar{\boldsymbol{r}}'\boldsymbol{w}-\min_{\boldsymbol{r}\in\mathcal{W}}\boldsymbol{w}'\boldsymbol{r}\right\}.$$

The minimum CPVaR portfolio optimization problem is equivalent to solving the following convex optimization problem,

$$\begin{array}{llll}
(\text{CPVaR}) & \min\limits_{\boldsymbol{x},\boldsymbol{s},\boldsymbol{t},\boldsymbol{w},v} & -\bar{\boldsymbol{r}}'\boldsymbol{x}+\sqrt{\frac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}\boldsymbol{x}-\boldsymbol{w}-\boldsymbol{s}\\ \boldsymbol{x}-\boldsymbol{w}+\boldsymbol{t}\end{pmatrix}\right\|_2+\bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s}-\bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t}+\bar{\boldsymbol{r}}'\boldsymbol{w}+v & \\
& \text{s.t.} & -\boldsymbol{w}'\boldsymbol{r}\leq v \qquad \forall \boldsymbol{r}\in\mathcal{W} & \\
& & \bar{\boldsymbol{r}}'\boldsymbol{x}=\tau & (6)\\
& & \boldsymbol{e}'\boldsymbol{x}=1 & \\
& & \boldsymbol{s},\boldsymbol{t}\geq\boldsymbol{0}. &
\end{array}$$

Note that in the context of robust optimization, the first set of constraints of Model (6) is known as the *robust counterpart*. If $\mathcal{W}$ is a set whose cardinality is a small finite number, the robust counterpart can be expanded into a set of $|\mathcal{W}|$ linear constraints. However, it is conceivable that the cardinality of $\mathcal{W}$ is exponentially large or infinite. Under such circumstances, Model (6) remains a computationally tractable problem if the convex hull of $\mathcal{W}$ can be compactly represented by linear or more generally conic constraints. In this case, the *robust counterpart* can be represented by the dual conic constraints projected from higher dimensions. We refer interested readers to Ben-Tal and Nemirovski (1998) [8] or Bertsimas and Sim (2004) [10] for details. Theorem 2 shows that the CPVaR is a coherent risk measure.

Theorem 2 *The CPVaR measure satisfies the axioms of a coherent risk measure.*

Proof : Refer to online supplementary materials available at electronic version of this paper at *http://www.sciencedirect.com.*

Moreover, the CPVaR provides a tighter bound to VaR than PVaR. This is intuitive as the additional support information of asset returns under CPVaR approach helps to narrow the risk bound compared to PVaR approach.

Theorem 3 *Suppose $\tilde{\boldsymbol{r}}$ has partitioned statistics, $(\bar{\boldsymbol{r}}^1, \bar{\boldsymbol{r}}^2, \hat{\boldsymbol{\Sigma}})$ and support $\mathcal{W}$, then*

$$\mathrm{VaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \le \mathrm{CVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \le \mathrm{CPVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \le \mathrm{PVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right)$$

Proof : Refer to online supplementary materials available at electronic version of this paper at *http://www.sciencedirect.com.* Note that since CPVaR is a coherent risk measure, it must also be an upper bound to CVaR, which has been established as the coherent risk measure with the best upper bound to VaR.

Lastly, we can establish the direct links between the PVaR and CPVaR in a special case.

Corollary 1 *If $\mathcal{V}$ comprises asset returns whose support $\mathcal{W}$ is the space $\Re^n$, then PVaR = CPVaR. Moreover, PVaR is coherent on $\mathcal{V}$.*

Proof : Note that when $\mathcal{W}$ is unbounded, $\min_{\boldsymbol{r}\in\mathcal{W}} \boldsymbol{r}'\boldsymbol{w}$ takes the value of zero if $\boldsymbol{w} = \boldsymbol{0}$ and negative infinity otherwise. Hence, PVaR = CPVaR. ■

In summary, we have introduced a new portfolio optimization approach by minimizing the PVaR of the portfolio. The PVaR measure provides a tighter bound to classical VaR than the symmetric WVaR and accommodates distributional asymmetry. With support information, the PVaR measure turns into a coherent risk measure. In the next section, we empirically investigate the performance of this approach by comparing it with conventional portfolio optimization approaches.

4 Empirical Results

We compare the performance of minimizing-portfolio PVaR under our approach with minimizing portfolio variance under the classical Markowitz approach in this section. We employ three sets of data to perform the empirical analyses. Firstly, we use simulated asset returns and show that our PVaR approach performs well for negatively-skewed returns. Secondly, we compare both optimization approaches by employing a widely available data set of industry portfolio returns, “48Ind” from Kenneth French’s website. Finally, we use historical daily stock returns from NYSE and AMEX.

4.1 Simulated Data

We consider the portfolio allocation problem of 5 simulated asset returns, each distributed with skewed normal distributions (see Azzalini (1985) [5]), of increasing negative skewness. The random variables are scaled and shifted such that the means of the asset returns are

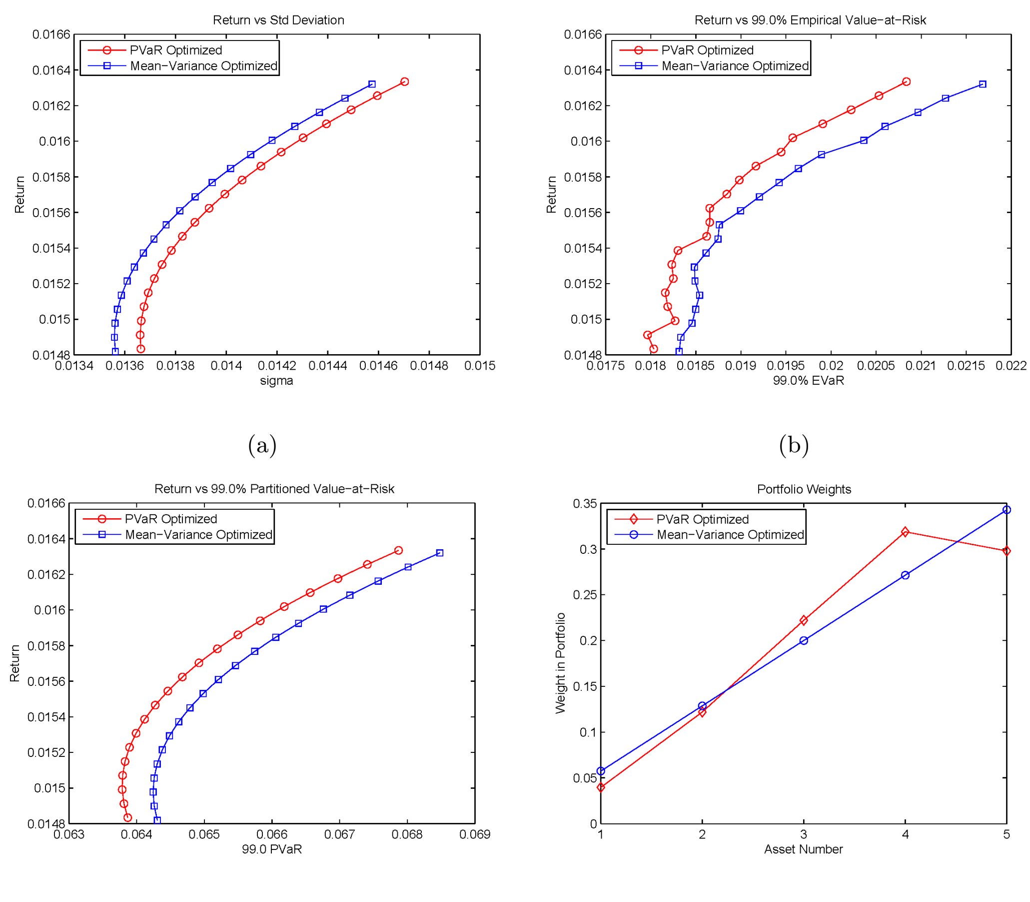
$$\mu_i = 0.01 + 0.0025i, \quad \forall i \in \{0, 1, 2, 3, 4\}$$

while their standard deviations are identical at $\sigma = 0.03$. Simulated data allow us to control for portfolio attributes such as skewness, and thus enable a more sensitive evaluation of the portfolio performances. We follow Azzalini [5] in constructing the skewness as $\gamma_i = (2 - \pi/2)\delta_i^3(\pi/2 - \delta_i^2)^{-3/2}$, and chose the skewness parameter to be $\delta_i = -0.24975i$, $\forall i \in \{0, 1, 2, 3, 4\}$ such that assets with higher mean returns have higher negative skewness to reflect the positive tradeoff between risk and return. For different target returns, we separately minimize the 99% portfolio PVaR and the portfolio variance.

Panels (a) to (c) of Figure 1 compare the mean-risk efficient frontiers of the optimized portfolios using different risk measures, such as standard deviation, empirical VaR, and PVaR. As expected, the Markowitz and PVaR approaches perform well when measured by their respective risk metrics. The Markowitz approach dominates the PVaR approach when risk is measured in standard deviation. The PVaR approach dominates the Markowitz approach when risk is measured in PVaR. More importantly, we find that the PVaR-optimized portfolio outperforms the Markowitz-optimized portfolio by having smaller empirical VaR measures. This means a smaller loss risk.

The intuition underlying these results is illustrated in Panel (d) of Figure 1, which shows the portfolio weights for a target return of 0.0165. Other targets could be constructed, but this case allows for easy comparison due to linear weights under the Markowitz approach. In this case, the Markowitz approach yields a diversified portfolio with the smallest portfolio standard deviation by having a linear weighting of the assets to achieve the target return. However, the Markowitz portfolio, being indifferent to skewness, places a large weight on Asset 5, which has a large negative skew, thereby increasing the overall portfolio empirical VaR. In contrast, the PVaR optimization heavily penalizes Asset 5, due to its large negative skewness. To achieve the target portfolio mean, the PVaR approach in compensation assigns a larger weight instead to Asset 4. The deviation from linearity increases the standard deviation in the PVaR optimal portfolio in comparison with that in the Markowitz approach, but reduces the empirical VaR.

Figure 1: Panel a to Panel c plot the efficient frontiers of portfolio return against three risk metrics for 5 skew-normal distributed assets respectively: sigma (i.e. standard deviation), 99.0% EVaR, and 99.0% PVaR. Panel d plots the weights of the optimal portfolio for a target return of 0.0165.



(a) (b)

(c) (d)

Table 1 compares various normalized or mean-to-risk metrics for the PVaR-optimized and Markowitz-optimized portfolios, with respect to different target returns. For simplicity, assuming that the risk free rate is zero, we shall henceforth interpret the μ/σ ratio as the Sharpe ratio, which is intuitively recognized as the reward to risk ratio. Thus, in addition to the Sharpe ratio, we also tabulated the ratios of mean over Empirical VaR (EVaR) and mean over PVaR, to measure the normalized tail risk of the portfolios. Similar to the Sharpe ratio, larger values of these normalized metrics indicate larger reward-risk ratios and thus better performance of the portfolio. The parameter α controls the target return τ for each optimization portfolio, and represents the percentage increase over the computed return of the equal-weighted portfolio, r_{EW}, where

$$\tau = (1 + \alpha) r_{EW}.$$

Table 1: Comparison of normalized risk metrics for different target returns between PVaR optimization and mean-variance (M-V) optimization, for the simulated dataset. The third column, Δ, records the difference of each risk metric between the PVaR optimization and M-V optimization, where Δ = PVaR − M-V.

α	$\frac{\mu}{\sigma}$			$\frac{\mu}{EVaR_{99\%}}$			$\frac{\mu}{PVaR_{99\%}}$		
	PVaR	M-V	Δ	PVaR	M-V	Δ	PVaR	M-V	Δ
0.0000	1.0856	1.0926	(0.0070)	0.8226	0.8091	0.0135	0.2323	0.2304	0.0019
0.0053	1.0916	1.0988	(0.0072)	0.8301	0.8126	0.0175	0.2337	0.2318	0.0019
0.0105	1.0971	1.1044	(0.0073)	0.8206	0.8115	0.0091	0.2350	0.2331	0.0019
0.0158	1.1020	1.1094	(0.0074)	0.8285	0.8139	0.0146	0.2362	0.2343	0.0019
0.0211	1.1064	1.1140	(0.0076)	0.8342	0.8163	0.0179	0.2374	0.2354	0.0020
0.0263	1.1102	1.1180	(0.0078)	0.8344	0.8228	0.0116	0.2383	0.2363	0.0020
0.0316	1.1135	1.1214	(0.0079)	0.8395	0.8274	0.0121	0.2392	0.2372	0.0020
0.0368	1.1163	1.1243	(0.0080)	0.8407	0.8259	0.0148	0.2400	0.2379	0.0021
0.0421	1.1185	1.1266	(0.0081)	0.8306	0.8243	0.0063	0.2406	0.2385	0.0021
0.0474	1.1202	1.1284	(0.0082)	0.8333	0.8278	0.0055	0.2412	0.2390	0.0022
0.0526	1.1214	1.1297	(0.0083)	0.8375	0.8217	0.0158	0.2416	0.2394	0.0022
0.0579	1.1221	1.1305	(0.0084)	0.8333	0.8169	0.0164	0.2419	0.2397	0.0022
0.0632	1.1222	1.1307	(0.0085)	0.8314	0.8117	0.0197	0.2421	0.2398	0.0023
0.0684	1.1219	1.1305	(0.0086)	0.8276	0.8070	0.0206	0.2422	0.2399	0.0023
0.0737	1.1212	1.1298	(0.0086)	0.8196	0.8005	0.0191	0.2422	0.2399	0.0023
0.0789	1.1200	1.1286	(0.0086)	0.8184	0.7859	0.0325	0.2420	0.2397	0.0023
0.0842	1.1183	1.1270	(0.0087)	0.8086	0.7807	0.0279	0.2418	0.2395	0.0023
0.0895	1.1162	1.1250	(0.0088)	0.7999	0.7710	0.0289	0.2415	0.2392	0.0023
0.0947	1.1138	1.1226	(0.0088)	0.7917	0.7636	0.0281	0.2411	0.2388	0.0023
0.1000	1.1109	1.1198	(0.0089)	0.7839	0.7527	0.0312	0.2407	0.2383	0.0024

Similar to the results on the efficient frontiers shown earlier, we observe that the Markowitz portfolios perform marginally better than the PVaR portfolios when Sharpe ratio is used. However, under the normalized 99%-EVaR and 99%-PVaR risk metric, the PVaR optimized portfolios outperforms the Markowitz portfolios in every case of different target returns. This result indicates that the PVaR optimized portfolios have lower tail risks than the Markowitz portfolios.

4.2 Industry Portfolio Data

In this subsection, we report the empirical results of the out-of-sample performance of PVaR-minimized portfolios in comparison with the mean-variance optimized portfolios. The dataset comprises the daily returns of 48 industry portfolios from July 1, 1963 to August 31, 2007 (48Ind). The returns featured a significant departure from normality as the Jarque-Bera test (1980) [28] for normality on each industry portfolio yielded p-values of less than 0.001 in all cases.

Similar to many studies of portfolio strategies, we employ the rolling-window procedure to estimate the optimal portfolios. In particular, we employ 5 years of historical data, with each year being 250 trading days, to estimate the required moments and compute the optimal portfolios. We then hold this portfolio for half a year, $M = 125$ trading days, to obtain out-of-sample statistics for performance evaluation. We repeat the exercise by moving forward one half-year at a time.

Let T denote the total number of daily returns in the dataset. We let $\boldsymbol{r}^t$ denote the vector of returns on the t^{th} trading day, ordered such that the most recent day corresponds to $t = 1$. We let $L = 10M$ (5 years) denote the length of estimation window. For both portfolio optimization strategies, we use the returns $\boldsymbol{r}^{kM+1}, \ldots, \boldsymbol{r}^{kM+L}$ to compute the respective portfolios $\boldsymbol{x}_k$, for a given target return, in each period for $k = 1, \ldots, K_{max}$, where the maximum number of periods, $K_{max} = \left\lfloor \frac{T-L-M}{M} \right\rfloor$. Expressing $t = kM + \hat{t}$, the out-of-sample realized return on the t^{th} day is determined by

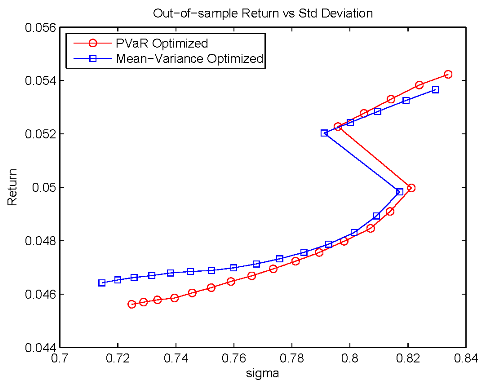
$$s_{kM+\hat{t}} = \boldsymbol{x}_{k+1}{}' \boldsymbol{r}^{kM+\hat{t}}$$

for $k = 0, \ldots, K_{max} - 1$ and $\hat{t} = 1, \ldots, M$. Denoting the actual number of days used in the experiment as $N = MK_{max}$, we obtain the out-of-sample portfolio mean, variance, and 99% Empirical VaR respectively by:

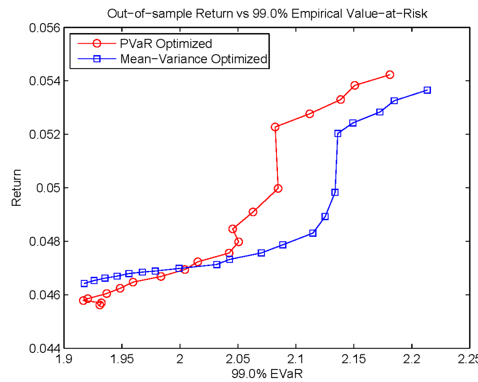
$$\begin{aligned} \hat{\mu} &= \frac{1}{N} \sum_{j=1}^{N} s_j \\ \hat{\sigma}^2 &= \frac{1}{N-1} \sum_{j=1}^{N} (s_j - \hat{\mu})^2 \\ \widehat{EVaR}_{99\%} &= -s_{(\lfloor 0.01N \rfloor)} \end{aligned}$$

where $s_{(\cdot)}$ denotes the order statistics of $s_1, \ldots, s_N$ such that $s_{(1)} \leq s_{(2)} \leq \ldots \leq s_{(N)}$. We repeat the experiment for different target returns. The target returns were selected so its average is the equal-weighted portfolio return.

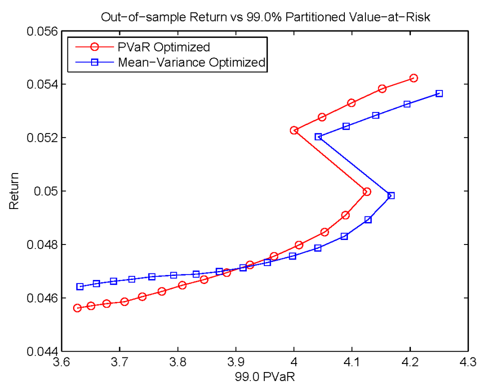
Figure 2: Panel a to Panel c plot the efficient frontiers of portfolio return against three risk metrics for 48 Industry dataset respectively: sigma, 99.0% EVaR, and 99.0% PVaR.



(a)



(b)



(c)

Figure 2 shows the efficient frontiers of the optimized portfolios under different risk measures. Similar to the simulated experiment, we observe that under both the 99%-PVaR and the 99%-Empirical VaR risk metrics, the out-of-sample PVaR-optimized portfolios perform better in the sense of smaller risk than the Markowitz portfolios for a large range of target returns. Even under the standard deviation risk metric, there is a range of target returns in which the PVaR-optimized portfolios actually have smaller standard deviations in the out-of-sample performance compared to the Markowitz portfolios.

Table 2 details the out-of-sample empirical results of the normalized risk measures of the PVaR and Markowitz portfolios over different target returns. From the table, we observe that the Sharpe ratio for

Table 2: Out-of-Sample comparison of normalized risk metrics for different target returns between PVaR optimization and mean-variance (M-V) optimization, for the 48 Industry portfolio returns (48Ind dataset). The third column, Δ, records the difference of each risk metric between the PVaR optimization and M-V optimization, where $\Delta = \text{PVaR} - \text{M-V}$.

α	$\frac{\mu}{\sigma}$			$\frac{\mu}{EVaR_{99\%}}$			$\frac{\mu}{PVaR_{99\%}}$		
	PVaR	M-V	Δ	PVaR	M-V	Δ	PVaR	M-V	Δ
0.0000	0.0629	0.0650	(0.0021)	0.0236	0.0242	(0.0006)	0.0126	0.0126	0.0000
0.0263	0.0627	0.0646	(0.0019)	0.0236	0.0242	(0.0006)	0.0126	0.0125	0.0001
0.0526	0.0624	0.0642	(0.0018)	0.0239	0.0241	(0.0002)	0.0126	0.0125	0.0001
0.0789	0.0620	0.0638	(0.0018)	0.0239	0.0240	(0.0001)	0.0125	0.0125	0.0000
0.1053	0.0618	0.0634	(0.0016)	0.0238	0.0239	(0.0001)	0.0125	0.0124	0.0001
0.1316	0.0615	0.0629	(0.0014)	0.0237	0.0238	(0.0001)	0.0125	0.0124	0.0001
0.1579	0.0612	0.0623	(0.0011)	0.0237	0.0237	0.0000	0.0125	0.0123	0.0002
0.1842	0.0609	0.0618	(0.0009)	0.0235	0.0235	0.0000	0.0124	0.0122	0.0002
0.2105	0.0607	0.0614	(0.0007)	0.0234	0.0232	0.0002	0.0124	0.0122	0.0002
0.2368	0.0605	0.0610	(0.0005)	0.0234	0.0232	0.0002	0.0124	0.0121	0.0003
0.2632	0.0602	0.0607	(0.0005)	0.0233	0.0230	0.0003	0.0123	0.0121	0.0002
0.2895	0.0601	0.0604	(0.0003)	0.0234	0.0229	0.0005	0.0123	0.0120	0.0003
0.3158	0.0600	0.0603	(0.0003)	0.0237	0.0228	0.0009	0.0123	0.0120	0.0003
0.3421	0.0603	0.0605	(0.0002)	0.0238	0.0230	0.0008	0.0124	0.0121	0.0003
0.3684	0.0609	0.0610	(0.0001)	0.0240	0.0234	0.0006	0.0124	0.0121	0.0003
0.3947	0.0657	0.0658	(0.0001)	0.0251	0.0244	0.0007	0.0129	0.0126	0.0003
0.4211	0.0656	0.0655	0.0001	0.0250	0.0244	0.0006	0.0129	0.0125	0.0004
0.4474	0.0655	0.0653	0.0002	0.0249	0.0243	0.0006	0.0128	0.0125	0.0003
0.4737	0.0653	0.0650	0.0003	0.0250	0.0244	0.0006	0.0128	0.0124	0.0004
0.5000	0.0650	0.0647	0.0003	0.0249	0.0242	0.0007	0.0128	0.0124	0.0004

mean-variance optimized portfolios generally outperforms the PVaR-optimized portfolios. However, the PVaR optimized portfolios generally perform better with higher return-risk ratios under the normalized 99%-EVaR and 99%-PVaR risk metrics. This finding again, reconfirms that the PVaR approach is an improvement from the classical Markowitz approach for minimizing tail risk, especially when the return distributions are not normal.

4.3 Daily Stock Returns

Intuitively, our PVaR approach should perform better than the Markowitz approach in optimizing over assets which exhibit strong negative skewness. We use the standard random sample selection procedure as Chan et al. (1999) [13] and Jagannathan and Ma (2003) [27], employing historical daily returns of stocks traded on the NYSE and the AMEX from 1962 to 2006. 500 stocks were randomly selected for each year during the sampling period. The stocks had prices greater than 5 dollars and market capitalization more than the 80th percentile of the size distribution of NYSE firms. A particular stock was also selected provided that it had 6 years of subsequent daily returns in the year in which it was selected. Similar to the rolling window approach used in the previous subsection, we consider the first 5 years of daily returns as in-sample data, and use this data for moment estimation and optimization. The next one year of daily returns are used as out-of-sample testing data to evaluate the performance of the optimized portfolios. When a daily return is missing, the equally weighted market return of that day is used instead.

To prepare the data, we began by rank-ordering the stocks in order of skewness (using the first 5 years of training data). In the computation of skewness for rank ordering, we omitted the maximal and minimal daily return from each stock over the 5 years to avoid outliers due to data entry errors. We next grouped these 500 stocks into 20 groups of 25 stocks each, still in order of increasing skewness. We then created 20 assets from the groups, by performing an equal-weighted portfolio return in each group. Finally, we sorted the 20 assets according to skewness, and as in the simulation case, chose the most negatively skewed 5 assets for our performance evaluations.

We now use the first 5 years of data to estimate the statistics of the 5 assets, and we apply both PVaR and Markowitz approaches to separately find the optimal portfolios, with a target mean equal to that of an equal-weighted portfolio of assets. For the PVaR optimization, we optimize at a level of 99%. As in the previous empirical procedure, we hold the optimized portfolios in the next or 6^{th} year in an out-of-sample test to evaluate various normalized risk metrics. The entire experiment was repeated 40 times for different samples of 500 stocks since the sample period (1962-2006) allows us to break the datasets into 40 over-lapping sub-samples with 6 years data in each of them. Table 3 shows the results of each of the 40 runs of the experiment, as well as the exact number of data points used for in-sample (L) and out-of-sample (M) measurements respectively.

To evaluate the performance of the respective optimized portfolios, we take the difference between the normalized risk metrics of the PVaR-optimized portfolios and the Markowitz portfolios, and compare the fraction of runs in which the PVaR-optimized portfolios outperform the the Markowitz portfolios. We further conduct a Z-test to test whether this fraction is significantly larger than 0.5, and report the p-values for the tests. As expected, we observe that for the dataset of negatively-skewed assets, the PVaR portfolios performed significantly better than the Markowitz portfolios. More specifically, the PVaR approach generates a higher risk-return tradeoff than the Makovitz approach about 65.0% of the time when we use the most negative skewed five portfolios for test.

The reason we focus on the improvement of negatively-skewed portfolio is because according to Mitton and Vorkink (2007) [35], investors prefer positive skewness over Sharpe ratio. Therefore, we are more concerned about the the optimal portfolio when underlying return distributions are negatively skewed, especially during market downturns. Hence, it is important to show that our new approach works well for negatively-skewed portfolios.

5 Conclusion

In this study, we present a new PVaR approach to portfolio optimization when underlying return distributions are ambiguous and asymmetrical. Our theoretical results show that the proposed PVaR is a very useful risk measure. Not only does the PVaR provides a tighter bound to theoretical VaR than other existing robust risk metrics, such as WVaR, it is also a coherent risk measure when the distributional support is given, whereas conventional risk measures such as variance and VaR fail to do so. The ordering of the magnitudes of CPVaR, PVaR, and WVaR also show that given more information from the partitioned statistics of the asset returns, we could avoid unnecessarily large capital allocation due to robust estimations of VaR. More importantly, our PVaR measure is an asymmetric risk measure whereby the conventional symmetric risk measures such as standard deviation do not take into consideration of the distinct difference between investors' preferences for upside risk versus the aversion to downside risk.

Empirically, we use both simulated and real financial data to demonstrates that our approach improves on existing mean-variance portfolio optimization. When return distributions are not elliptically

Table 3: Comparison of out-of-sample normalized risk metrics of a portfolio of 5 most negatively skewed assets. The summary table shows the percentage of experiments in which the respective risk metrics for the PVaR optimization outperforms the M-V optimization.

summary statistics

S/No	L	M	$\frac{\mu}{\sigma}$			$\frac{\mu}{EVaR_{99\%}}$			$\frac{\mu}{PVaR_{99\%}}$		
			PVaR	M-V	Δ	PVaR	M-V	Δ	PVaR	M-V	Δ
1	1259	251	0.02858	0.02682	0.00176	0.00939	0.00877	0.00062	0.00495	0.00466	0.00029
2	1259	224	0.23502	0.23461	0.00041	0.09576	0.09490	0.00086	0.04433	0.04404	0.00029
3	1232	251	(0.08254)	(0.08125)	(0.00130)	(0.03776)	(0.03782)	0.00006	(0.01489)	(0.01474)	(0.00015)
4	1230	255	0.06286	0.06189	0.00097	0.02164	0.02123	0.00041	0.01104	0.01087	0.00017
5	1232	254	0.03158	0.03292	(0.00134)	0.01419	0.01456	(0.00037)	0.00608	0.00633	(0.00025)
6	1235	250	(0.08629)	(0.08779)	0.00151	(0.02892)	(0.02958)	0.00066	(0.01588)	(0.01620)	0.00032
7	1234	252	(0.04641)	(0.04459)	(0.00182)	(0.01515)	(0.01434)	(0.00081)	(0.00768)	(0.00735)	(0.00033)
8	1262	252	0.00592	0.00326	0.00266	0.00253	0.00139	0.00114	0.00121	0.00067	0.00054
9	1263	255	0.14891	0.14661	0.00230	0.06308	0.06284	0.00024	0.03010	0.02954	0.00056
10	1263	253	0.05437	0.05419	0.00017	0.02144	0.02112	0.00031	0.01011	0.01005	0.00006
11	1262	251	0.01576	0.01389	0.00188	0.00597	0.00523	0.00074	0.00290	0.00255	0.00035
12	1263	253	0.08064	0.07985	0.00079	0.02691	0.02665	0.00026	0.01327	0.01317	0.00011
13	1264	253	0.02491	0.02342	0.00148	0.00537	0.00505	0.00032	0.00347	0.00326	0.00021
14	1265	252	0.24830	0.24817	0.00013	0.07847	0.07846	0.00001	0.04438	0.04432	0.00006
15	1262	253	(0.07450)	(0.06976)	(0.00474)	(0.02002)	(0.01877)	(0.00125)	(0.01192)	(0.01115)	(0.00077)
16	1262	254	0.21667	0.21651	0.00016	0.07360	0.07348	0.00013	0.04517	0.04516	0.00001
17	1265	253	0.06566	0.06757	(0.00191)	0.02806	0.02829	(0.00023)	0.01286	0.01324	(0.00038)
18	1265	252	0.06426	0.07005	(0.00580)	0.03238	0.03307	(0.00068)	0.01338	0.01487	(0.00149)
19	1264	251	0.17600	0.16933	0.00667	0.07760	0.07305	0.00454	0.03763	0.03592	0.00171
20	1263	254	0.08899	0.08955	(0.00056)	0.02881	0.02721	0.00160	0.01456	0.01467	(0.00012)
21	1264	254	(0.01588)	(0.01643)	0.00055	(0.00397)	(0.00412)	0.00015	(0.00208)	(0.00216)	0.00008
22	1264	252	0.12260	0.11884	0.00376	0.04848	0.04522	0.00327	0.02188	0.02067	0.00121
23	1263	253	0.05120	0.05339	(0.00220)	0.01408	0.01503	(0.00095)	0.00718	0.00753	(0.00035)
24	1264	251	0.03732	0.03696	0.00035	0.01285	0.01279	0.00006	0.00627	0.00621	0.00006
25	1264	255	0.05832	0.06994	(0.01162)	0.01922	0.02183	(0.00262)	0.01079	0.01307	(0.00229)
26	1265	253	0.16549	0.14521	0.02029	0.05370	0.04559	0.00811	0.03156	0.02741	0.00415
27	1264	254	0.04751	0.04716	0.00035	0.01386	0.01339	0.00047	0.00761	0.00753	0.00009
28	1266	252	0.09517	0.09601	(0.00084)	0.03853	0.03918	(0.00065)	0.01692	0.01712	(0.00020)
29	1265	252	0.18845	0.18755	0.00091	0.05482	0.05464	0.00018	0.03009	0.02995	0.00014
30	1266	252	0.09607	0.09458	0.00149	0.02937	0.02842	0.00095	0.01665	0.01634	0.00031
31	1263	253	0.21022	0.21313	(0.00291)	0.05978	0.06043	(0.00064)	0.03227	0.03281	(0.00054)
32	1263	252	(0.04366)	(0.04509)	0.00143	(0.01073)	(0.01110)	0.00037	(0.00693)	(0.00716)	0.00023
33	1261	254	0.03597	0.04093	(0.00496)	0.01491	0.01690	(0.00198)	0.00711	0.00813	(0.00102)
34	1263	251	0.06622	0.06392	0.00229	0.02481	0.02408	0.00073	0.01207	0.01156	0.00050
35	1262	246	0.09647	0.09940	(0.00293)	0.02593	0.02686	(0.00094)	0.01669	0.01722	(0.00052)
36	1256	253	(0.04926)	(0.04921)	(0.00005)	(0.01891)	(0.01889)	(0.00001)	(0.00857)	(0.00857)	(0.00001)
37	1256	253	0.26106	0.25607	0.00499	0.11330	0.10867	0.00463	0.05407	0.05271	0.00135
38	1257	251	0.08723	0.08700	0.00023	0.02791	0.02787	0.00004	0.01461	0.01456	0.00004
39	1254	253	0.13885	0.13883	0.00001	0.05566	0.05566	0.00000	0.02592	0.02592	0.00000
40	1256	250	0.09569	0.09067	0.00503	0.03336	0.03164	0.00172	0.01636	0.01557	0.00079

Summary Table

Risk Metric	% (PVaR > M-V)	p-value
$\frac{\mu}{\sigma}$	65.00%	0.02889
$\frac{\mu}{EVaR_{99\%}}$	70.00%	0.00571
$\frac{\mu}{PVaR_{99\%}}$	65.00%	0.02889

symmetrical or multivariate normal, and/or are negatively skewed, the PVaR approach offers significant empirical improvements compared to conventional approaches in terms of higher return-to-risk ratios measured in conventional empirical VaR and/or our PVaR risk metrics.

References

[1] Alles, L. A., & J. L. Kling (1994). Regularities in the variation of skewness in asset returns. *Journal of Financial Research*, 17, 427-438.

[2] Ang, Andrew, & J. Chen (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics*, 63(3), 443-494.

[3] Arditti, F. D. (1967). Risk and the required return on equity. *Journal of Finance*, 22, 19-36.

[4] Artzner, P., F. Delbaen, J. M. Eber, & D. Heath (1999). Coherent measures of risk. *Mathematical Finance*, 9, 203-228.

[5] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, 12, 171-178.

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

[7] Bekaert, G., C. B. Erb, C. R. Harvey, & T. E. Viskanta (1998). Distributional characteristics of emerging market returns and asset Allocation. *Journal of Portfolio Management*, 24, Winter, 102-116.

[8] Ben-Tal, A., & A. Nemirovski (1998). Robust convex optimization. *Mathematics of Operations Research*, 23, 769-805.

[9] Bertsimas, D., & I. Popescu (2005). Optimal inequalities in probability theory: A convex optimization approach. *SIAM Journal of Optimization*, 15(3), 780-804.

[10] Bertsimas, D., & M. Sim (2004). Price of robustness. *Operations Research*, 52, 35-53.

[11] Brunnermeier, M., C. Gollier, & J. Parker (2007). Optimal beliefs, asset prices and the preference for skewed returns. *American Economic Review Papers and Proceedings*, 97, 159-165.

[12] Chamberlain, G. (1983). A characterization of the distributions that imply mean-variance utility functions. *Journal of Economic Theory*, 29, 185-201.

[13] Chan, L. K. C., J. Karceski, & J. Lakonishok (1999). On portfolio optimization: forecasting covariances and choosing the risk model. *Review of Financial Studies*, 12, 937-974.

[14] Charnes, A., W. W. Cooper, & G. H. Symonds (1958). Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science*, 4, 235-263.

[15] Chen, L., S. He, & S. Zhang (2011). Tight Bounds for Some Risk Measures, with Applications to Robust Portfolio Selection. *Operations Research*, 59, 847-865.

[16] Conine, T. E. Jr., & M. J. Tamarkin (1981). On diversification given asymmetry in returns. *Journal of Finance*, 36, 1143-1155.

[17] Delage, E., & Y. Ye (2010). Distributionally Robust Optimization Under Moment Uncertainty with Application to Data-Driven Problems. *Operations Research*, 58, 595-612.

[18] DeMiguel, V., & F. J. Nogales (2009). Portfolio Selection with Robust Estimation. *Operations Research*, 57, 560-577.

[19] Duffie, D., & J. Pan (1997). An Overview of value at risk. *Journal of Derivatives*, 4(3), 7-49.

[20] El Ghaoui, L., M. Oks, & F. Oustry (2003).Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operation Research*, 51(4), 543-556.

[21] Erb, C. B., C. R. Harvey, & T. E. Viskanta (1999). New perspective on emerging market bonds. *Journal of Portfolio Management*, 25, Winter, 83-92.

[22] Fama, E. (1965). The behavior of stock prices. *Journal of Business*, 47, 244-280.

[23] Föllmer, H., A. Schied (2004): *Stochastic Finance: An Introduction in Discrete Time.*, Walter de Gruyter, Berlin.

[24] Gaivoronski, A., & G. Pflug (2005). Value-at-risk in portfolio optimization: properties and computational approach. *Journal of Risk*, 7(2), 1-31.

[25] Harvey, C. R., & A. Siddique (2000). Conditional skewness in asset pricing tests. *Journal of Finance*, 55(3), 1263-1295.

[26] Huang, D., S. Zhu, F. J. Fabozzi, & M. Fukushima (2010). Portfolio Selection under Distributional Uncertainty: A Relative Robust CVaR Approach. *European Journal of Operational Research*, 203, 185-194.

[27] Jagannathan, R., & T. Ma (2003). Risk reduction in large portfolios: why imposing the wrong constraints helps. *Journal of Finance*, 58(4), 1651-1683.

[28] Jarque, C. M., & A. K. Bera (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economic Letters*, 6(3), 255-259.

[29] Jorion, P. (2000). Value at risk: the new benchmark for managing financial risk. McGraw-Hill.

[30] Kerstens, K., A. Mounir, & I. V. Woestyne (2011). Geometric Representation of the Mean-Variance-Skewness Portfolio Frontier based upon the Shortage Function. *European Journal of Operational Research*, 210, 81-94.

[31] Kon, S. J. (1984). Models of stock returns - a comparison. *Journal of Finance*, 39, 147-165.

[32] Larsen, N., H. Mausser, & S. Uryasev (2002). Algorithm for Optimization of Value-at-Risk, in P. Pardalos and V.K. Tsitsiringos, (Eds.). Financial Engineering, e-Commerce and Supply Chain. Kluwer Academic Publishers, 129-157.

[33] Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance*, 7, 77-91.

[34] Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *Journal of Business*, 36, 394-419.

[35] Mitton, T., & K. Vorkink (2007). Equilibrium underdiversification and the preference for skewness. *Review of Financial Studies*, 20, 1255-1288.

[36] Natarajan, K., D. Pachamanova, & M. Sim (2008). Incorporating asymmetric distributional information in robust value-at-risk optimization. *Management Science*, 54(3), 573-585.

[37] Rockafellar R.T., & S. Uryasev (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2(3), 21-41.

[38] Rockafellar, R. T., & S. Uryasev (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, 26(7), 1443-1471.

[39] Scott, R. C., & P. A. Horvath (1980). On the direction of preference for moments of higher order than the variance. *Journal of Finance*, 35, 915-919.

[40] Simkowitz, M. A., & W. L. Beedles (1978). Diversification in a three-moment world. *Journal of Financial and Quantitative Analysis*, 13, 927-941.

[41] Simkowitz, M. A., & W. L. Beedles (1980). Asymmetric stable distributed security return. *Journal of the American Statistical Association*, 75, 306-312.

[42] Theodossiou, P. (1998). Financial Data and the Skewed Generalized t Distribution. *Management Science*, 44, 1650-1661.

[43] Tobin, J. (1958). Liquidity preference as behavior toward risk. *Review of Economic Studies*, 25, 65-85.

[44] Yamai, Y., & T. Yohiba (2002). On the validity of VaR: comparative analyses with expected shortfall. *Monetary and Economic Studies*, 20(1), 57-85.

[45] Zhu, S., & M. Fukushima (2009). Worst-Case Conditional Value-at-Risk with Application to Robust Portfolio Management. *Operations Research*, 57, 1155-1168.

Online Supplementary Materials

Proof of Theorem 1: Observe that

$$\begin{aligned}
& \mathrm{PVaR}_{1-\epsilon}\left(y+\tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \\
= & -y-\bar{\boldsymbol{r}}'\boldsymbol{x}+\min_{\boldsymbol{s},\boldsymbol{t}\geq\boldsymbol{0}}\left\{\sqrt{\tfrac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\binom{\boldsymbol{x}-\boldsymbol{s}}{\boldsymbol{x}+\boldsymbol{t}}\right\|_2+\bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s}-\bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t}\right\} \\
\leq & -y-\bar{\boldsymbol{r}}'\boldsymbol{x}+\sqrt{\tfrac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\binom{\boldsymbol{x}}{\boldsymbol{x}}\right\|_2 \\
= & -y-\bar{\boldsymbol{r}}'\boldsymbol{x}+\sqrt{\tfrac{1-\epsilon}{\epsilon}}\|\boldsymbol{\Sigma}^{1/2}\boldsymbol{x}\|_2 \\
= & \mathrm{WVaR}_{1-\epsilon}\left(y+\tilde{\boldsymbol{r}}'\boldsymbol{x}\right),
\end{aligned}$$

where the first inequality follows from choosing $\boldsymbol{s}=\boldsymbol{t}=\boldsymbol{0}$ and the second equality follows from Equality (4). To show that $\mathrm{VaR}_{1-\epsilon}\left(y+\tilde{\boldsymbol{r}}'\boldsymbol{x}\right)\leq\mathrm{PVaR}_{1-\epsilon}\left(y+\tilde{\boldsymbol{r}}'\boldsymbol{x}\right)$, it suffices to show that for all $r\in\Re$,

$$\mathrm{PVaR}_{1-\epsilon}\left(y+\tilde{\boldsymbol{r}}'\boldsymbol{x}\right)\leq r\Rightarrow\mathrm{VaR}_{1-\epsilon}\left(y+\tilde{\boldsymbol{r}}'\boldsymbol{x}\right)\leq r,$$

which, from the definition of the VaR measure, is equivalent to showing

$$\mathbb{P}(-(\tilde{\boldsymbol{r}}'\boldsymbol{x}+y)>r)\leq\epsilon.$$

We first consider the inner minimization in the PVaR definition, of the function $f(\boldsymbol{s},\boldsymbol{t})$, on the domain $\boldsymbol{s},\boldsymbol{t}\in\Re^n_+$, where

$$f(\boldsymbol{s},\boldsymbol{t})=\sqrt{\frac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}\boldsymbol{x}-\boldsymbol{s}\\ \boldsymbol{x}+\boldsymbol{t}\end{pmatrix}\right\|_2+\bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s}-\bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t}$$

we notice that $f(\boldsymbol{s},\boldsymbol{t})$ diverges to $+\infty$ when $(\boldsymbol{s},\boldsymbol{t})$ is unbounded, since $\bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s}-\bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t}\geq 0$ on its domain. Further, $f(\boldsymbol{s},\boldsymbol{t})$ is finite for $(\boldsymbol{s},\boldsymbol{t})=(\boldsymbol{0},\boldsymbol{0})$. This guarantees us the existence of finite optimizers of $f(\boldsymbol{s},\boldsymbol{t})$, denoted by $\boldsymbol{s}^*,\boldsymbol{t}^*$. Now we let

$$\mathrm{PVaR}_{1-\epsilon}\left(y+\tilde{\boldsymbol{r}}'\boldsymbol{x}\right)=-y-\bar{\boldsymbol{r}}'\boldsymbol{x}+\sqrt{\frac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}\boldsymbol{x}-\boldsymbol{s}^*\\ \boldsymbol{x}+\boldsymbol{t}^*\end{pmatrix}\right\|_2+\bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s}^*-\bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t}^*$$

Suppose $\mathrm{PVaR}_{1-\epsilon}\left(y+\tilde{\boldsymbol{r}}'\boldsymbol{x}\right)\leq r$, we have

$$-y-\bar{\boldsymbol{r}}'\boldsymbol{x}+\sqrt{\frac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}\boldsymbol{x}-\boldsymbol{s}^*\\ \boldsymbol{x}+\boldsymbol{t}^*\end{pmatrix}\right\|_2+\bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s}^*-\bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t}^*\leq r. \tag{7}$$

Hence,

$$\begin{aligned}
& \mathbb{P}(-(\tilde{\boldsymbol{r}}'\boldsymbol{x}+y)>r) \\
\leq & \mathbb{P}\left(-(\tilde{\boldsymbol{r}}'\boldsymbol{x}+y)>-y-\bar{\boldsymbol{r}}'\boldsymbol{x}+\sqrt{\tfrac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\binom{\boldsymbol{x}-\boldsymbol{s}^*}{\boldsymbol{x}+\boldsymbol{t}^*}\right\|_2+\bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s}^*-\bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t}^*\right) \quad \text{using Inequality (7)} \\
= & \mathbb{P}\left(-\tilde{\boldsymbol{r}}'\boldsymbol{x}+\bar{\boldsymbol{r}}^{1\prime}(\boldsymbol{x}-\boldsymbol{s}^*)+\bar{\boldsymbol{r}}^{2\prime}(\boldsymbol{x}+\boldsymbol{t}^*)>\sqrt{\tfrac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\binom{\boldsymbol{x}-\boldsymbol{s}^*}{\boldsymbol{x}+\boldsymbol{t}^*}\right\|_2\right) \quad \text{since } \bar{\boldsymbol{r}}=\bar{\boldsymbol{r}}^1+\bar{\boldsymbol{r}}^2.
\end{aligned}$$

Observe that since $\tilde{\boldsymbol{r}}^1 \geq \mathbf{0}$ and that $\tilde{\boldsymbol{r}}^2 \leq \mathbf{0}$, we have $\tilde{\boldsymbol{r}}^{1\prime}\boldsymbol{s}^* \geq 0$ and $-\tilde{\boldsymbol{r}}^{2\prime}\boldsymbol{t}^* \geq 0$. Therefore,

$$\begin{aligned}
& \mathbb{P}\left(-\tilde{\boldsymbol{r}}'\boldsymbol{x} + \bar{\boldsymbol{r}}^{1\prime}(\boldsymbol{x}-\boldsymbol{s}^*) + \bar{\boldsymbol{r}}^{2\prime}(\boldsymbol{x}+\boldsymbol{t}^*) > \sqrt{\frac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}\boldsymbol{x}-\boldsymbol{s}^*\\ \boldsymbol{x}+\boldsymbol{t}^*\end{pmatrix}\right\|_2\right) \\
\leq\ & \mathbb{P}\left(-\tilde{\boldsymbol{r}}'\boldsymbol{x} + \tilde{\boldsymbol{r}}^{1\prime}\boldsymbol{s}^* - \tilde{\boldsymbol{r}}^{2\prime}\boldsymbol{t}^* + \bar{\boldsymbol{r}}^{1\prime}(\boldsymbol{x}-\boldsymbol{s}^*) + \bar{\boldsymbol{r}}^{2\prime}(\boldsymbol{x}+\boldsymbol{s}^*) > \sqrt{\frac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}\boldsymbol{x}-\boldsymbol{s}^*\\ \boldsymbol{x}+\boldsymbol{t}^*\end{pmatrix}\right\|_2\right) \\
=\ & \mathbb{P}\left(-\tilde{\boldsymbol{r}}^{1\prime}(\boldsymbol{x}-\boldsymbol{s}^*) - \tilde{\boldsymbol{r}}^{2\prime}(\boldsymbol{x}+\boldsymbol{t}^*) + \bar{\boldsymbol{r}}^{1\prime}(\boldsymbol{x}-\boldsymbol{s}^*) + \bar{\boldsymbol{r}}^{2\prime}(\boldsymbol{x}+\boldsymbol{t}^*) > \sqrt{\frac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}\boldsymbol{x}-\boldsymbol{s}^*\\ \boldsymbol{x}+\boldsymbol{t}^*\end{pmatrix}\right\|_2\right) \\
=\ & \mathbb{P}\left(-(\tilde{\boldsymbol{r}}^1-\bar{\boldsymbol{r}}^1)'(\boldsymbol{x}-\boldsymbol{s}^*) - (\tilde{\boldsymbol{r}}^2-\bar{\boldsymbol{r}}^2)'(\boldsymbol{x}+\boldsymbol{t}^*) > \sqrt{\frac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}\boldsymbol{x}-\boldsymbol{s}^*\\ \boldsymbol{x}+\boldsymbol{t}^*\end{pmatrix}\right\|_2\right),
\end{aligned}$$

where the first equality is due to $\tilde{\boldsymbol{r}} = \tilde{\boldsymbol{r}}^1 + \tilde{\boldsymbol{r}}^2$. Finally, we use a well known one sided Tchebychev inequality that for any random variable $\tilde{z}$ with zero mean and standard deviation σ,

$$\mathbb{P}(\tilde{z} > k\sigma) \leq \frac{1}{1+k^2}.$$

Clearly, the random variable $-(\tilde{\boldsymbol{r}}^1-\bar{\boldsymbol{r}}^1)'(\boldsymbol{x}-\boldsymbol{s}^*) - (\tilde{\boldsymbol{r}}^2-\bar{\boldsymbol{r}}^2)'(\boldsymbol{x}+\boldsymbol{t}^*)$ has zero mean and standard deviation, $\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}\boldsymbol{x}-\boldsymbol{s}^*\\ \boldsymbol{x}+\boldsymbol{t}^*\end{pmatrix}\right\|_2$. Hence,

$$\begin{aligned}
& \mathbb{P}(-(\tilde{\boldsymbol{r}}'\boldsymbol{x}+y) > r) \\
\leq\ & \mathbb{P}\left(-(\tilde{\boldsymbol{r}}^1-\bar{\boldsymbol{r}}^1)'(\boldsymbol{x}-\boldsymbol{s}^*) - (\tilde{\boldsymbol{r}}^2-\bar{\boldsymbol{r}}^2)'(\boldsymbol{x}+\boldsymbol{t}^*) > \sqrt{\frac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}\boldsymbol{x}-\boldsymbol{s}^*\\ \boldsymbol{x}+\boldsymbol{t}^*\end{pmatrix}\right\|_2\right) \\
\leq\ & \epsilon.
\end{aligned}$$

■

Proof of Proposition 1: It is trivial to show translation invariance. Note that $\mathrm{PVaR}_{1-\epsilon}(0) = 0$. To show positive homogeneity, we observe that for all $k > 0$

$$\begin{aligned}
& \mathrm{PVaR}_{1-\epsilon}\left(k(y+\tilde{\boldsymbol{r}}'\boldsymbol{x})\right) \\
=\ & -ky - k\bar{\boldsymbol{r}}'\boldsymbol{x} + \min_{\boldsymbol{s},\boldsymbol{t}\geq\mathbf{0}}\left\{\sqrt{\frac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}k\boldsymbol{x}-\boldsymbol{s}\\ k\boldsymbol{x}+\boldsymbol{t}\end{pmatrix}\right\|_2 + \bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s} - \bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t}\right\} \\
=\ & -ky - k\bar{\boldsymbol{r}}'\boldsymbol{x} + \min_{k\boldsymbol{s},k\boldsymbol{t}\geq\mathbf{0}}\left\{\sqrt{\frac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}k\boldsymbol{x}-k\boldsymbol{s}\\ k\boldsymbol{x}+k\boldsymbol{t}\end{pmatrix}\right\|_2 + k\bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s} - k\bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t}\right\} \\
=\ & -ky - k\bar{\boldsymbol{r}}'\boldsymbol{x} + k\min_{k\boldsymbol{s},k\boldsymbol{t}\geq\mathbf{0}}\left\{\sqrt{\frac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}\boldsymbol{x}-\boldsymbol{s}\\ \boldsymbol{x}+\boldsymbol{t}\end{pmatrix}\right\|_2 + \bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s} - \bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t}\right\} \\
=\ & -ky - k\bar{\boldsymbol{r}}'\boldsymbol{x} + k\min_{\boldsymbol{s},\boldsymbol{t}\geq\mathbf{0}}\left\{\sqrt{\frac{1-\epsilon}{\epsilon}}\left\|\hat{\boldsymbol{\Sigma}}^{1/2}\begin{pmatrix}\boldsymbol{x}-\boldsymbol{s}\\ \boldsymbol{x}+\boldsymbol{t}\end{pmatrix}\right\|_2 + \bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s} - \bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t}\right\} \\
=\ & k\mathrm{PVaR}_{1-\epsilon}\left(y+\tilde{\boldsymbol{r}}'\boldsymbol{x}\right)
\end{aligned}$$

The second and forth equalities follow from substituting the nonnegative variables $\boldsymbol{s}, \boldsymbol{t}$ with $k\boldsymbol{s}, k\boldsymbol{t}$.

Finally, to show subadditivity, we note that the epigraph of the risk measure, $\text{PVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \leq r$ is a classical conic quadratic feasible space, which is convex in $(y, \boldsymbol{x}, r)$. Therefore, $\text{PVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right)$ is convex over $(y, \boldsymbol{x})$. Hence, for all $y^1 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^1, y^2 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^2 \in \mathcal{V}$, we have

$$\begin{aligned}
& \text{PVaR}_{1-\epsilon}\left(y^1 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^1 + y^2 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^2\right) & \\
= \; & 2\text{PVaR}_{1-\epsilon}\left(\tfrac{1}{2}(y^1 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^1) + \tfrac{1}{2}(y^2 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^2)\right) & \text{Positive homogeneity} \\
\leq \; & \text{PVaR}_{1-\epsilon}\left(y^1 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^1\right) + \text{PVaR}_{1-\epsilon}\left(y^2 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^2\right) & \text{Convexity.}
\end{aligned}$$

■

Proof of Theorem 2: It is trivial to show translation invariance, and the proof of that positive homogeneity and subadditivity is similar to the exposition of proposition (1). Finally, to show monotonicity, we first show that for any $(y, \boldsymbol{x})$ satisfying $y + \tilde{\boldsymbol{r}}'\boldsymbol{x} \geq 0$ or equivalently,

$$y + \min_{\boldsymbol{r} \in \mathcal{W}} \boldsymbol{r}'\boldsymbol{x} \geq 0$$

we have $\text{CPVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \leq 0$. Indeed,

$$\begin{aligned}
& \text{CPVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \\
= \; & -y - \bar{\boldsymbol{r}}'\boldsymbol{x} + \min_{\boldsymbol{s},\boldsymbol{t} \geq \boldsymbol{0}, \boldsymbol{w}} \left\{ \sqrt{\frac{1-\epsilon}{\epsilon}} \left\| \hat{\boldsymbol{\Sigma}}^{1/2} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{w} - \boldsymbol{s} \\ \boldsymbol{x} - \boldsymbol{w} + \boldsymbol{t} \end{pmatrix} \right\|_2 + \bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s} - \bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t} + \bar{\boldsymbol{r}}'\boldsymbol{w} - \min_{\boldsymbol{r} \in \mathcal{W}} \boldsymbol{w}'\boldsymbol{r} \right\} \\
\leq \; & -y - \bar{\boldsymbol{r}}'\boldsymbol{x} + \min_{\boldsymbol{s},\boldsymbol{t} \geq \boldsymbol{0}, \boldsymbol{w}} \left\{ \sqrt{\frac{1-\epsilon}{\epsilon}} \left\| \hat{\boldsymbol{\Sigma}}^{1/2} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{w} - \boldsymbol{s} \\ \boldsymbol{x} - \boldsymbol{w} + \boldsymbol{t} \end{pmatrix} \right\|_2 + \bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s} - \bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t} + \bar{\boldsymbol{r}}'\boldsymbol{w} - \min_{\boldsymbol{r} \in \mathcal{W}} \boldsymbol{w}'\boldsymbol{r} \right. \\
& \qquad \left. : \; \boldsymbol{w} = \boldsymbol{x}, \boldsymbol{s} = \boldsymbol{t} = \boldsymbol{0} \right\} \\
= \; & -y - \min_{\boldsymbol{r} \in \mathcal{W}} \boldsymbol{r}'\boldsymbol{x} \leq 0
\end{aligned}$$

Finally, given $y_1 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^1, y_2 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^2 \in \mathcal{V}$ such that $y_1 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^1 \geq y_2 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^2$ we have

$$\text{CPVaR}_{1-\epsilon}\left(y_1 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^1 - y_2 - \tilde{\boldsymbol{r}}'\boldsymbol{x}^2\right) \leq 0.$$

Hence,

$$\begin{aligned}
& \text{CPVaR}_{1-\epsilon}\left(y_1 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^1\right) & \\
\leq \; & \text{CPVaR}_{1-\epsilon}\left(y_1 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^1 - y_2 - \tilde{\boldsymbol{r}}'\boldsymbol{x}^2\right) + \text{CPVaR}_{1-\epsilon}\left(y_2 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^2\right) & \text{Subadditivity} \\
\leq \; & \text{CPVaR}_{1-\epsilon}\left(y_2 + \tilde{\boldsymbol{r}}'\boldsymbol{x}^2\right). &
\end{aligned}$$

■

Proof of Theorem 3: For all $(y, \boldsymbol{x}) \in \Re^{n+1}$, we have

$$
\begin{aligned}
& \mathrm{CPVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \\
= \;& -y - \bar{\boldsymbol{r}}'\boldsymbol{x} + \min_{\boldsymbol{s},\boldsymbol{t} \geq \boldsymbol{0}, \boldsymbol{w}} \left\{ \sqrt{\frac{1-\epsilon}{\epsilon}} \left\| \hat{\boldsymbol{\Sigma}}^{1/2} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{w} - \boldsymbol{s} \\ \boldsymbol{x} - \boldsymbol{w} + \boldsymbol{t} \end{pmatrix} \right\|_2 + \bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s} - \bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t} + \bar{\boldsymbol{r}}'\boldsymbol{w} - \min_{\boldsymbol{r} \in \mathcal{W}} \boldsymbol{w}'\boldsymbol{r} \right\} \\
\leq \;& -y - \bar{\boldsymbol{r}}'\boldsymbol{x} + \min_{\boldsymbol{s},\boldsymbol{t} \geq \boldsymbol{0}, \boldsymbol{w}} \left\{ \sqrt{\frac{1-\epsilon}{\epsilon}} \left\| \hat{\boldsymbol{\Sigma}}^{1/2} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{w} - \boldsymbol{s} \\ \boldsymbol{x} - \boldsymbol{w} + \boldsymbol{t} \end{pmatrix} \right\|_2 + \bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s} - \bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t} + \bar{\boldsymbol{r}}'\boldsymbol{w} - \min_{\boldsymbol{r} \in \mathcal{W}} \boldsymbol{w}'\boldsymbol{r} \;:\; \boldsymbol{w} = \boldsymbol{0} \right\} \\
= \;& \mathrm{PVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right).
\end{aligned}
$$

Similar to the proof of Theorem 1, to show that $\mathrm{VaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \leq \mathrm{CPVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right)$, it suffices to show that for all $r \in \Re$, $\mathrm{CPVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \leq r$, implies

$$
\mathbb{P}(-(\tilde{\boldsymbol{r}}'\boldsymbol{x} + y) > r) \leq \epsilon.
$$

We consider the inner minimization problem of the function $g(\boldsymbol{s}, \boldsymbol{t}, \boldsymbol{w})$, on the domain $\boldsymbol{s}, \boldsymbol{t} \in \Re^n_+, \boldsymbol{w} \in \Re^n$, where

$$
g(\boldsymbol{s}, \boldsymbol{t}, \boldsymbol{w}) = \sqrt{\frac{1-\epsilon}{\epsilon}} \left\| \hat{\boldsymbol{\Sigma}}^{1/2} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{w} - \boldsymbol{s} \\ \boldsymbol{x} - \boldsymbol{w} + \boldsymbol{t} \end{pmatrix} \right\|_2 + \bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s} - \bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t} + h(\boldsymbol{w})
$$

and

$$
h(\boldsymbol{w}) = \bar{\boldsymbol{r}}'\boldsymbol{w} - \min_{\boldsymbol{r} \in \mathcal{W}} \boldsymbol{w}'\boldsymbol{r}
$$

Since the vector of mean returns lies within the support, i.e. $\bar{\boldsymbol{r}} \in \mathcal{W}$, we have the inequality $h(\boldsymbol{w}) \geq 0, \ \forall \boldsymbol{w} \in \Re^n$, since

$$
h(\boldsymbol{w}) = \bar{\boldsymbol{r}}'\boldsymbol{w} - \min_{\boldsymbol{r} \in \mathcal{W}} \boldsymbol{w}'\boldsymbol{r} \geq \bar{\boldsymbol{r}}'\boldsymbol{w} - \boldsymbol{w}'\bar{\boldsymbol{r}} = 0
$$

Hence, when $(\boldsymbol{s}, \boldsymbol{t}, \boldsymbol{w})$ is unbounded, $g(\boldsymbol{s}, \boldsymbol{t}, \boldsymbol{w})$ diverges to $+\infty$, since $\bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s} - \bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t} + h(\boldsymbol{w}) \geq 0$ on the domain of g. On the other hand, $g(\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{0})$ is finite. We are therefore guaranteed finite optimizers $(\boldsymbol{s}^*, \boldsymbol{t}^*, \boldsymbol{w}^*)$ of $g(\boldsymbol{s}, \boldsymbol{t}, \boldsymbol{w})$. Now we express

$$
\mathrm{CPVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) = -y - \bar{\boldsymbol{r}}'\boldsymbol{x} + \sqrt{\frac{1-\epsilon}{\epsilon}} \left\| \hat{\boldsymbol{\Sigma}}^{1/2} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{w}^* - \boldsymbol{s}^* \\ \boldsymbol{x} - \boldsymbol{w}^* + \boldsymbol{t}^* \end{pmatrix} \right\|_2 + \bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s}^* - \bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t}^* + \bar{\boldsymbol{r}}'\boldsymbol{w}^* - \min_{\boldsymbol{r} \in \mathcal{W}} \boldsymbol{w}^{*\prime}\boldsymbol{r}
$$

Suppose $\mathrm{CPVaR}_{1-\epsilon}\left(y + \tilde{\boldsymbol{r}}'\boldsymbol{x}\right) \leq r$, we have

$$
-y - \bar{\boldsymbol{r}}'\boldsymbol{x} + \sqrt{\frac{1-\epsilon}{\epsilon}} \left\| \hat{\boldsymbol{\Sigma}}^{1/2} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{w}^* - \boldsymbol{s}^* \\ \boldsymbol{x} - \boldsymbol{w}^* + \boldsymbol{t}^* \end{pmatrix} \right\|_2 + \bar{\boldsymbol{r}}^{1\prime}\boldsymbol{s}^* - \bar{\boldsymbol{r}}^{2\prime}\boldsymbol{t}^* + \bar{\boldsymbol{r}}'\boldsymbol{w}^* - \min_{\boldsymbol{r} \in \mathcal{W}} \boldsymbol{w}^{*\prime}\boldsymbol{r} \leq r. \tag{8}
$$

Hence, using Inequality (8) and substituting $\bar{\boldsymbol{r}} = \bar{\boldsymbol{r}}^1 + \bar{\boldsymbol{r}}^2$, we have

$$
\begin{aligned}
& \mathbb{P}(-(\tilde{\boldsymbol{r}}'\boldsymbol{x} + y) > r) \\
\leq \;& \mathbb{P}\left(-\tilde{\boldsymbol{r}}'\boldsymbol{x} + \min_{\boldsymbol{r} \in \mathcal{W}} \boldsymbol{w}^{*\prime}\boldsymbol{r} + \bar{\boldsymbol{r}}^{1\prime}(\boldsymbol{x} - \boldsymbol{s}^* - \boldsymbol{w}^*) + \bar{\boldsymbol{r}}^{2\prime}(\boldsymbol{x} + \boldsymbol{t}^* - \boldsymbol{w}^*) > \sqrt{\frac{1-\epsilon}{\epsilon}} \left\| \hat{\boldsymbol{\Sigma}}^{1/2} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{w}^* - \boldsymbol{s}^* \\ \boldsymbol{x} - \boldsymbol{w}^* + \boldsymbol{t}^* \end{pmatrix} \right\|_2 \right)
\end{aligned}
$$

Observe that since $\tilde{\boldsymbol{r}}^1 \geq 0$, $\tilde{\boldsymbol{r}}^2 \leq \boldsymbol{0}$ and $\min_{\boldsymbol{r}\in\mathcal{W}} \boldsymbol{w}^{*\prime}\boldsymbol{r} \leq \boldsymbol{w}^{*\prime}\tilde{\boldsymbol{r}}$ we have

$$\tilde{\boldsymbol{r}}^{1\prime}\boldsymbol{s}^* - \tilde{\boldsymbol{r}}^{2\prime}\boldsymbol{t}^* + \boldsymbol{w}^{*\prime}\tilde{\boldsymbol{r}} \geq \min_{\boldsymbol{r}\in\mathcal{W}} \boldsymbol{w}^{*\prime}\boldsymbol{r}$$

Therefore, using one sided Tchebychev inequality, we have

$$\begin{aligned}
& \mathbb{P}\left(-\tilde{\boldsymbol{r}}'\boldsymbol{x} + \min_{\boldsymbol{r}\in\mathcal{W}} \boldsymbol{w}^{*\prime}\boldsymbol{r} + \bar{\boldsymbol{r}}^{1\prime}(\boldsymbol{x} - \boldsymbol{s}^* - \boldsymbol{w}^*) + \bar{\boldsymbol{r}}^{2\prime}(\boldsymbol{x} + \boldsymbol{t}^* - \boldsymbol{w}^*) > \sqrt{\frac{1-\epsilon}{\epsilon}} \left\| \hat{\boldsymbol{\Sigma}}^{1/2} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{w}^* - \boldsymbol{s}^* \\ \boldsymbol{x} - \boldsymbol{w}^* + \boldsymbol{t}^* \end{pmatrix} \right\|_2 \right) \\
\leq \; & \mathbb{P}\left(-(\tilde{\boldsymbol{r}}^1 - \bar{\boldsymbol{r}}^1)'(\boldsymbol{x} - \boldsymbol{s}^* - \boldsymbol{w}^*) - (\tilde{\boldsymbol{r}}^2 - \bar{\boldsymbol{r}}^2)'(\boldsymbol{x} + \boldsymbol{t}^* - \boldsymbol{w}^*) > \sqrt{\frac{1-\epsilon}{\epsilon}} \left\| \hat{\boldsymbol{\Sigma}}^{1/2} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{w}^* - \boldsymbol{s}^* \\ \boldsymbol{x} - \boldsymbol{w}^* + \boldsymbol{t}^* \end{pmatrix} \right\|_2 \right) \\
\leq \; & \epsilon.
\end{aligned}$$

■