Robust Multi-period Portfolio Selection Based on Downside Risk with Asymmetrically Distributed Uncertainty Set

Aifan Ling[1]*, Jie Sun[2]† Meihua Wang[3]‡

1. School of Finance, Jiangxi University of Finance & Economics, Nanchang, 330013, China
2. Department of Mathematics and Statistics, Curtin University, Perth, Australia
3. School of Economics and Management, Xidian University, Xi'an 710071, China.

December 3, 2018

Abstract. Motivated by the asymmetrical attitudes of investors towards downside losses and upside gains, this paper proposes a robust multi-period portfolio selection model based on downside risk with asymmetrically distributed uncertainty set, in which the downside losses of a portfolio are controlled by the lower partial moment (LPM). A computationally tractable approximation approach based on second-order cone optimization is used for solving the proposed model. We show in theory that the optimal solution of the robust model can generate a given probability guarantee for individual and joint stochastic constraints. The effect of the asymmetrically distributed uncertainty set on performance of the optimal solution is analyzed by the usual comparative static method. Comprehensive numerical comparisons with real market data are reported and indicate that the proposed model can obtain the smaller standard deviation and turnover ratios which reduce the Sharpe ratios of optimal portfolio, compared with some well-known models in the literature.

Keywords. Risk management, Robust portfolio optimization, Lower partial moment, Asymmetric uncertainty set, Multi-period portfolio selection.

JEL. C61, G11

1 Introduction

Behaviorial finance research finds that the attitudes of investors towards downside losses and upside gains are asymmetric (Kahneman and Tversky, 1979) and that investors who are sensitive to downside losses, relative to upside gains, require a compensation for holding assets

---

*The corresponding author, Telephone: +86+791-83816792, E-mail: aifanling@jxufe.edu.cn
†E-mail: jie.sun@curtin.edu.au
‡The co-corresponding author: yuyu0504@163.com

that covary strongly with the market when the market declines. Recent empirical results support this opinion and show that stocks with high tail risk (downside loss risk) can earn higher returns than stocks with low tail risk (see, e.g. Ang, Chen and Xing, 2006; Kelly and Jiang, 2014)

The lower partial moment (LPM for short) introduced by Bawa and Lindenberg (1977) is an intuitive downside risk measure, and is more effective in controlling portfolio losses (Chen, He and Zhang, 2011). Bali, Cakici and Whitelaw (2014) apply the properties of LPM to construct hybrid tail covariance risk (H-TCR) and find a positive and significant relation between H-TCR and expected stock returns. Hence, how to optimally obtain the premium or compensation of downside losses is an important and worthy problem in portfolio selection, but it is difficult due to the features of 'sharp peak and heavy tail' of risky assets which can lead to large losses for investors. These results motivate our interest in using LPM to establish our portfolio model since LPM is a simple and intuitive downside loss risk measure that can capture the left tail loss of portfolios.

Formally, let $\widetilde{\xi}$ be a random variable and a be constant. The LPM of random variable $\widetilde{\xi}$ with respect to a target a can be expressed as

$$\mathrm{LPM}_k(a; \widetilde{\xi}) = \mathbb{E}\left[\left((a - \widetilde{\xi})_+\right)^k\right],$$

where $k \geq 0$ is a constant that reflects the attitude of investor towards risk and $(\cdot)_+ = \max(\cdot, 0)$. In order to get higher compensation for tail risk, we naturally consider maximizing the excess return of portfolio for each unit of LPM, that is, maximizing the ratio

$$LPMratio = \frac{\text{Excess returns of portfolio}}{\mathrm{LPM}_k}, \quad \text{for } k \geq 1,$$

can obtain the best premium, or compensation for downside losses, measured by LPM. However, the maximization of the ratio is not easy to solve. Fortunately, noting the relationship between the Sharpe ratio and the mean-variance model, we can prove that the maximization of the $LPMratio$ is equivalent to the minimization of the mean-LPM (MLPM for short) risk measure defined as[1]

$$\rho_k(a, \lambda; \widetilde{\xi}) = -\mathbb{E}[\widetilde{\xi}] + \lambda \cdot \mathrm{LPM}_k(a; \widetilde{\xi}), \quad k \geq 0, \tag{1.1}$$

where $\lambda \geq 0$ is a constant. This is another motivation for using the the mean-LPM portfolio model —— it can reflect a trade-off between returns and downside losses.

Ruszczyński and Shapiro (2006) find that $\rho_k(a; \widetilde{\xi})$ with $k \geq 1$ is a coherent risk measure, that is, $\rho_k(a; \widetilde{\xi})$ satisfies convexity, monotonicity, translation invariance, and positive homogeneity

---

[1]The proof of equivalence of two problems is straightforward and omitted. The readers interested in the proof can request it by email from the corresponding author.

(see, e.g., Artzner et al. (1999) for details of coherent measures). In particular, when $k = 1$, $\mathrm{LPM}_k(a; \widetilde{\xi})$ can be computed by

$$\mathrm{LPM}(a; \widetilde{\xi}) = \mathbb{E}\left[(a - \widetilde{\xi})_+\right], \tag{1.2}$$

which is the expected shortfall of $\widetilde{\xi}$ falling below the target a. Hereafter we will drop the subscripts of LPM and ρ when $k = 1$.

Two difficulties must be faced by an investor if he/she uses MLPM to control the risk of the portfolio. One is that the exact expression of LPM is hard unless the distribution of the random variable $\widetilde{\xi}$ is known. Another is that estimations errors of the mean value and variance of $\widetilde{\xi}$ always exist even if $\widetilde{\xi}$ is assumed to follow a normal distribution. To overcome these difficulties, the worst-case risk measure (e.g., EL Ghaoui et al., 2003; Zhu et al., 2009; Zhu and Fukushima, 2009, Zymler et al., 2011, 2013; Zhu et al., 2016) and robust portfolio optimization models (e.g., Huang et al., 2008, 2010; Goldfarb and Iyengar, 2003; Ling, Sun and Yang, 2014; Ling et al., 2017) have been recently developed in the literature. Robust models can generally deal with distribution uncertainty and parameter estimation error, but the performances of robust models are dependent on the construction of uncertainty sets. Most importantly, many robust models share a common feature: the uncertainty structures are symmetric and therefore inconsistent with the asymmetric uncertainty structures in real markets. Recently, based on the forward and backward deviations of a random variable, Chen, Sim and Sun (2007) proposed an asymmetric uncertainty set and considered the probability guarantee that a single stochastic linear constraint was satisfied when the coefficients were affine functions of a random vector belonging to the asymmetric uncertainty set. Cases of joint stochastic linear constraints with asymmetric uncertainty sets were subsequently considered by Chen et al. (2010). An application in robust portfolio optimization has been proposed by Natarajan Pachamanova and Sim (2008), where the authors investigated a single-period portfolio selection problem with worst-case VaR and asymmetric uncertainty sets (see also Natarajan, Pachamanova and Sim (2009) for further results).

This paper studies a robust multi-period mean-LPM portfolio optimization model with an asymmetrically distributed uncertainty set. We attempt to analyze the impact of such an uncertainty set. A significant characteristic of our model is that the objective function includes not only the terminal wealth of the portfolio, but also the loss risk of the portfolio. The loss risk of a portfolio has become a major concern of investors since the 2008 financial crisis. To capture the 'market uncertainty' of real market portfolios, we assume that the market is driven by a linear risk factor and then construct an asymmetric uncertainty set of the linear risk factor. We also include a proportional transaction cost in our model and explore a computationally tractable approximation. Based on a second-order cone optimization, we obtain the solution methodology

3

of the proposed robust model. We also prove the joint probability guarantee that the optimal solution satisfies all stochastic constraints. We present several numerical comparisons with worst-case value-at-risk (WVaR), conditional VaR (CVaR), worst-case CVaR (WCVaR) and the equally-weighted strategy. With real market data, we show that the proposed models can obtain portfolios with better expected returns and Sharpe ratios and, at the same time, smaller standard deviations and lower turnover ratios.

There is much research to motivate our interests on this problem. Multi-period portfolio selection problems have been investigated for a long time, for example, Li and Ng (2000), Zhang, Liu and Xu (2012), Dokuchaev (2010), Bodnar, Parolya and Schmid (2015), Cui, Li and Li (2017), and Liu and Chen (2018). Recently, Gârleanu and Pedersen (2013) considered a new model of multi-period portfolio selection in which a closed-form optimal dynamic portfolio policy was obtained when the trading was costly and security returns were predictable by signals with different mean-reversion speeds. The dynamic model with general transaction costs was extended by Mei, DeMiguel and Nogales (2016) and the effect of general costs on optimal strategy was analyzed.

Our work is an important part of a growing literature on robust multi-period portfolio problems which have become a new focus of portfolio selection research and have been considered by Ben-Tal, Margelit and Nemirovski (2000), Bertsimas and Pachamanova (2008), Shen and Zhang (2008). Complexity and numerical difficulties are the main barriers to robust multi-period portfolio selection problems. These are also the main reasons why only some simple cases have been solved in this area. Our work extends the results of Natarajan Pachamanova and Sim (2008) in two directions. The first extension is that the asymmetrically distributed uncertainty set is used in a multi-period model rather than in a single-period model. The second is that MLPM is used as the risk measure rather than VaR. As is well known, VaR is not coherent and in fact can be viewed as the zero-order moment of a random variable, whereas LPM is a first-order moment and MLPM is a coherent risk measure[2].

Our work is related to Gulpinar and Pachamanova (2013), Gulpinar, Canakoglu and Pachamanova (2014) and Gulpinar, Pachamanova and Canakoglu (2016), where the multi-period Asset-Liability Management under uncertainty are considered . However, there exist some important differences between these work and our studies. The first key difference is that our paper considers a risk aversion objective (downside risk measure), whereas their papers (e.g. see, Gulpinar and Pachamanova, 2013 and Gulpinar, Pachamanova and Canakogl, 2016) consider risk neural objectives (expected wealth). The second key difference is that our model involves joint chance constraints and an asymmetric uncertainty set, which is different from the symmetric uncer-

---

[2]Even though VaR is not coherent, Natarajan, Pachamanova and Sim (2008) show that an approximation to VaR based on an asymmetric uncertainty set is coherent.

tainty set in Gulpinar and Pachamanova (2013). The third key difference is that our model can obtain the probability guarantee of individual (joint) stochastic constraint(s) for the robust multi-period portfolio model, which is particularly important in theory for robust portfolio optimization problems.

The paper is also related to Rujeerapaiboon, Kuhn and Wiesemann (2016), which also considers a robust multi-period portfolio selection problem. However, Rujeerapaiboon, Kuhn and Wiesemann (2016) discuss the accumulated terminal returns measured by VaR, while symmetrically distributed uncertainty of risky assets returns is assumed in that paper. Besides, they consider only minimizing the probability of the losses without discussion of the maximization of expected terminal returns. Unlike the model of Rujeerapaiboon, Kuhn and Wiesemann (2016), our model uses the coherent MLPM risk measure to control the loss of portfolios. Asymmetrically distributed uncertainty plays a key role in our model and we can obtain a high probability guarantee by choosing the size of the uncertainty set. The transaction cost is considered in our multi-period model, helping to reduce turnover ratios when the optimal strategy needs to be updated. Our model can be solved numerically by second-order cone programming techniques, which generally need less CPU memory and less computing time than the semi-definite programming techniques used in Rujeerapaiboon, Kuhn and Wiesemann (2016).

This paper is organized as follows. In Section 2, we describe the multi-period portfolio selection model with transaction costs. Section 3 recalls the ideas of robust optimization with stochastic constraints and constructs the asymmetric uncertainty set of a random vector. In Section 4, we formulate the robust counterpart of the multi-period portfolio selection model under the asymmetric uncertainty set and develop a computationally tractable second-order cone optimization approach. Section 5 reports our numerical results with comparisons based on real market data. The variables, parameters and symbols are listed in Appendix.

2 Multi-period portfolio selection model

Before introducing the multi-period portfolio selection model, it is necessary to describe the basic idea of robust portfolio selection that we will consider. Let $g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})$ be a gain (or wealth) function associated with the decision vector $\boldsymbol{x} \in \mathcal{X} \subseteq \mathbb{R}^n$ and the random vector $\widetilde{\boldsymbol{\xi}} \in \mathbb{R}^m$. Denote the joint distribution function of $\widetilde{\boldsymbol{\xi}}$ by F, which can be a continuous or a discrete distribution. Instead of assuming precise knowledge of the distribution F, we assume in this paper that the distribution F is only known to belong to an uncertainty set $\mathcal{F}$, that is, $F \in \mathcal{F}$. We assume that $\mathbb{E}\left[|g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})|\right] < \infty$ holds for all $\boldsymbol{x} \in \mathcal{X}$, so that the downside risk LPM will be properly defined. Then, the mean-LPM risk measure of the gain (or wealth) function $g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})$ with respect to a

constant target a is given by

$$
\begin{aligned}
\rho(a, \lambda; g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})) &= -\mathbb{E}[g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})] + \lambda \cdot \text{LPM}(a; g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})) \\
&= -\mathbb{E}[g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})] + \lambda \cdot \mathbb{E}[((a - g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}))_+)].
\end{aligned}
\tag{2.1}
$$

We consider mainly the following portfolio optimization problem in this paper,

$$
\min_{\boldsymbol{x} \in \mathcal{X}, G(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \geq 0.} \rho(a, \lambda; g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})),
\tag{2.2}
$$

where $G(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \geq 0$ is a (or some) constraint(s) involving the random vector $\widetilde{\boldsymbol{\xi}}$, which is called a stochastic constraint. Two difficulties must be faced by the decision maker. One is that the exact expression of ρ is difficult due to the unknown distribution of $\widetilde{\boldsymbol{\xi}}$. Another is that $G(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \geq 0$, in general, does not hold truely for all realizations of $\widetilde{\boldsymbol{\xi}}$.

We now introduce the background of the problem. Assume that an investor can invest in n risky assets and one risk-free asset at $t = 0$, the beginning of the 1st period, and then the investor will consider his/her terminal wealth and risk of the portfolio at $t = T$, the end of the last period. Denote the random returns of $n + 1$ assets at time $t(t = 0, 1, \cdots, T)$ by $(r_0^t; \widetilde{\boldsymbol{r}}^t) = (r_0^t, \widetilde{r}_1^t, \cdots, \widetilde{r}_n^t)' \in \mathbb{R}^{n+1}$, where r_0^t is the deterministic risk-free return and $\widetilde{r}_i^t$ is the random return of the ith risky asset, $\widetilde{\boldsymbol{r}}^0 = \boldsymbol{r}^0$ is realized and known by the investor at time $t = 0$. Let $(x_0^t; \boldsymbol{x}^t) = (x_0^t, x_1^t, \cdots, x_n^t)' \in \mathbb{R}^{n+1}$ $(t = 1, \cdots, T)$ be the portfolio weights at the tth period, where x_0^t is the cash amount and $\boldsymbol{x}^t = (x_1^t, \cdots, x_n^t)'$ with $x_i^t(i \geq 1)$, the dollar amount invested in the ith risky asset. For simplicity, assume that the initial budget for the investment is one, so that the budget constraint of the first period is

$$
x_0^1 + \boldsymbol{e}'\boldsymbol{x}^1 = 1, \boldsymbol{x} \geq 0,
\tag{2.3}
$$

where $\boldsymbol{e}$ is a vector of ones with suitable dimension. At time t (or at the beginning of the $(t + 1)^{th}$ period) $(t = 1, \cdots, T - 1)$, the investor will rebalance his/her t^{th}-period portfolio position $\boldsymbol{x}^t$. After rebalancing, recall that the assets are measured by their "dollar value", then, under self-finance and proportional transaction costs, we have for $t = 1, \cdots, T - 1$,

$$
x_i^{t+1} = x_i^t(1 + \widetilde{r}_i^t) + \Delta_i^{b,t+1} - \Delta_i^{s,t+1}, i = 1, \cdots, n,
\tag{2.4}
$$

and

$$
x_0^{t+1} = (1 + r_0^t)x_0^t + (1 - \theta) \sum_{i=1}^{n} \Delta_i^{s,t+1} - (1 + \theta) \sum_{i=1}^{n} \Delta_i^{b,t+1},
\tag{2.5}
$$

where θ is the proportional transaction cost of the traded one dollar, and $\Delta_i^{b,t+1}$ and $\Delta_i^{s,t+1}$

6

are respectively the dollar amounts of buying and selling the ith risky asset at the beginning of $(t+1)^{th}$ period. The wealth of portfolio, denoted by W^T, at the end of the Tth period is

$$W^T = x_0^T(1 + r_0^T) + (\boldsymbol{e} + \widetilde{\boldsymbol{r}}^T)'\boldsymbol{x}^T.$$

The objective of a risk-averse investor is to maximize expected wealth at the end of the final period (the Tth period) and minimize simultaneously the downside risk measured by the portfolio LPM, that is, the portfolio optimization of the investor can be expressed as,

$$\min_{\boldsymbol{x}^t, t=1,\cdots,T} \left\{\rho(a, \lambda; W^T)\right\} = \min_{\boldsymbol{x}^t, t=1,\cdots,T} \left\{-\mathbb{E}[W^T] + \lambda \cdot \mathbb{E}[(a - W^T)_+]\right\}, \qquad (2.6)$$

subject to (2.3), (2.4) and (2.5), where $\lambda \geq 0$ is the coefficient of risk aversion corresponding to the Tth period, which can reflect the trade-off between risk and terminal wealth.

We return to the rebalancing constraints (2.4) and (2.5). Due to the random parameter $\widetilde{r}_i^t$, these random equalities can not be satisfied with positive probability. Using the transformation of variables of Ben-Tal, Margalit and Nemirovski (2000), we denote

$$\begin{cases} R_0^{t+1} = (1 + r_0^0)(1 + r_0^1)\cdots(1 + r_0^t); \\ R_i^1 = 1 + r_i^0, \quad \widetilde{R}_i^2 = (1 + r_i^0)(1 + \widetilde{r}_i^1), \\ \cdots\cdots, \\ \widetilde{R}_i^{t+1} = \widetilde{R}_i^t(1 + \widetilde{r}_i^t) = (1 + r_i^0)(1 + \widetilde{r}_i^1)\cdots(1 + \widetilde{r}_i^t), \quad i = 1, \cdots, n, \quad t = 0, 1, \cdots, T. \end{cases} \qquad (2.7)$$

and $\widetilde{\boldsymbol{R}}^t = (\widetilde{R}_1^t, \cdots, \widetilde{R}_n^t)'(t = 1, 2, \cdots, T+1)$, where $\widetilde{\boldsymbol{R}}^1 = \boldsymbol{R}^1$ is certain and known at the beginning of the 1st period. Let

$$w_i^t = \frac{x_i^t}{\widetilde{R}_i^t}, \quad t = 1, 2, \cdots, T, \quad i = 0, 1, \cdots, n, \qquad (2.8)$$

and

$$z_i^{b,t} = \frac{\Delta_i^{b,t}}{\widetilde{R}_i^t}, \quad z_i^{s,t} = \frac{\Delta_i^{s,t}}{\widetilde{R}_i^t}, \quad t = 1, 2, \cdots, T, \quad i = 0, 1, \cdots, n, \qquad (2.9)$$

and denote $\boldsymbol{w}^t = (w_1^t, \cdots, w_n^t)' \in \mathbb{R}^n (t = 1, 2, \cdots, T)$. Then, the equality constraint (2.3) is

$$R_0^1 w_0^1 + (\boldsymbol{R}^1)'\boldsymbol{w}^1 = R_0^1 w_0^1 + \sum_{i=1}^{n} R_i^1 w_i^1 = 1. \qquad (2.10)$$

Especially, the equality constraint (2.4) can be changed into

$$w_i^{t+1} = w_i^t + z_i^{b,t+1} - z_i^{s,t+1} \quad i = 1, \cdots, n, \quad t = 1, 2, \cdots, T-1 \qquad (2.11)$$

which eliminates the random parameters. The equality constraint (2.5) is, for $t = 1, 2, \cdots, T-1$,

$$
\begin{aligned}
w_0^{t+1} &= w_0^t + (1 - \theta) \sum_{i=1}^n \frac{\widetilde{R}_i^{t+1}}{R_0^{t+1}} z_i^{s,t+1} - (1 + \theta) \sum_{i=1}^n \frac{\widetilde{R}_i^{t+1}}{R_0^{t+1}} z_i^{b,t+1} \\
&= w_0^t + \frac{(1 - \theta)}{R_0^{t+1}} \left(\widetilde{\boldsymbol{R}}^{t+1} \right)' \boldsymbol{z}^{s,t+1} - \frac{(1 + \theta)}{R_0^{t+1}} \left(\widetilde{\boldsymbol{R}}^{t+1} \right)' \boldsymbol{z}^{b,t+1}.
\end{aligned}
\tag{2.12}
$$

(2.12) is different from (2.10) and (2.11) since the constraint (2.12) contains the random parameters $\widetilde{\boldsymbol{R}}^t$ that will lead to the infeasibility. Thus, we relax (2.12) into a convex inequality constraint as follows.

$$
w_0^{t+1} \le w_0^t + \frac{(1 - \theta)}{R_0^{t+1}} \left(\widetilde{\boldsymbol{R}}^{t+1} \right)' \boldsymbol{z}^{s,t+1} - \frac{(1 + \theta)}{R_0^{t+1}} \left(\widetilde{\boldsymbol{R}}^{t+1} \right)' \boldsymbol{z}^{b,t+1}.
\tag{2.13}
$$

Under (2.7)—(2.9), the terminal wealth $W^T = w_0^T R_0^{T+1} + (\widetilde{\boldsymbol{R}}^{T+1})' \boldsymbol{w}^T$. Then, the multi-period portfolio selection problem can be reformulated as

$$
\begin{aligned}
\min_{w_0^t, \boldsymbol{w}^t, \boldsymbol{z}^{b,t} \boldsymbol{z}^{s,t}} \quad & \rho(a, \lambda; W^T) \\
&= -\mathbb{E}\left[w_0^T R_0^{T+1} + (\widetilde{\boldsymbol{R}}^{T+1})' \boldsymbol{w}^T \right] + \lambda \cdot \mathbb{E}\left[\left(a - w_0^T R_0^{T+1} - (\widetilde{\boldsymbol{R}}^{T+1})' \boldsymbol{w}^T \right)_+ \right] \\
\text{s.t.} \quad & (2.10), (2.11) \text{ and } (2.13); \\
& w_i^t \ge 0, \quad i = 0, 1, \cdots, n, t = 1, \cdots, T; \\
& z_i^{b,t} \ge 0, z_i^{s,t} \ge 0, i = 1, \cdots, n, t = 1, \cdots, T.
\end{aligned}
\tag{2.14}
$$

Noting that the objective function of problem (2.14) is related to nonlinear expectation and therefore is difficult to calculate without the distribution information of the random cumulative return $\widetilde{\boldsymbol{R}}^{t+1}$. On the other hand, the constraints (2.13) involving the random cumulative return $\widetilde{\boldsymbol{R}}^{t+1}$ are joint stochastic constraints. In theory, we need to find the optimal decisions $(w_0^t, \boldsymbol{w}^t, \boldsymbol{z}^{b,t} \boldsymbol{z}^{s,t})$, such that the inequality in (2.13) holds for all periods and all realizations of $\widetilde{\boldsymbol{R}}^{t+1}$. This is clearly not practical. An alternative approach is to transfer these stochastic constraints into deterministic constraints, and then find the optimal decisions $(w_0^t, \boldsymbol{w}^t, \boldsymbol{z}^{b,t} \boldsymbol{z}^{s,t})$, such that stochastic constraints (2.13) hold with a high probability. We will adopt the second approach to solve problem (2.14) in this paper and will introduce the robust optimization problem with asymmetric uncertainty set in the next section.

3 Robust optimization and asymmetric distribution

Consider a stochastic program with a general stochastic constraint,

$$\min_{\boldsymbol{x}} \left\{ \boldsymbol{c}'\boldsymbol{x} : G(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \geq 0, \boldsymbol{x} \in \mathcal{X} \right\}, \tag{3.1}$$

where $\widetilde{\boldsymbol{\xi}} \in \mathbb{R}^m$ is a random vector with the support set $\mathbb{S}_\xi \subseteq \mathbb{R}^m$. A solution $\boldsymbol{x}$ of (3.1) satisfying the stochastic constraint for all realizations of $\widetilde{\boldsymbol{\xi}} \in \mathbb{S}_\xi$ is desired, but in practice is generally impossible. The steps of robust optimization are to construct an uncertainty set, $\mathcal{U}_\Omega$, that is compact where the size is dependent on a constant Ω, and then to find a solution $\boldsymbol{x}$, such that (i) $\boldsymbol{x}$ is an optimal solution of (3.1) for all $\boldsymbol{\xi} \in \mathcal{U}_\Omega$, and (ii) the probability that the solution $\boldsymbol{x}$ of the resulting robust counterpart satisfies the stochastic constraint is a tractable function of Ω. The robust counterpart of stochastic problem (3.1) is

$$\min_{\boldsymbol{x}} \left\{ \boldsymbol{c}'\boldsymbol{x} : G(\boldsymbol{x}, \boldsymbol{\xi}) \geq 0, \quad \boldsymbol{\xi} \in \mathcal{U}_\Omega, \boldsymbol{x} \in \mathcal{X} \right\} \tag{3.2}$$

which is a determinate (or non-stochastic) optimization problem. If we can find the upper bound u_Ω dependent on Ω, such that the solution $\boldsymbol{x}$ of robust counterpart (3.2) satisfies

$$\mathbb{P} \left\{ G(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \geq 0 \right\} \geq u_\Omega,$$

then, we can use a constant Ω, such that u_Ω is close to one. This gives a probability guarantee of the solution $\boldsymbol{x}$ of robust counterpart (3.2).

A widely-used uncertainty set is the ellipsoidal set (e.g., EL Ghaoui, 2003; Gulpinar and Pachamanova, 2013, and Gulpinar et al., 2016):

$$\mathcal{U}_\Omega = \left\{ \boldsymbol{\xi} : \|\boldsymbol{\xi}\|_2 \leq \Omega \right\}. \tag{3.3}$$

It has been proved that if $G(\boldsymbol{x}, \boldsymbol{\xi})$ is linear in $\boldsymbol{\xi}$, the robust counterpart (3.2) with $\mathcal{U}_\Omega$ given by (3.3) can be converted into a convex cone constraint that can be solved using interior point algorithms, where the probability guarantee is not less than $u_\Omega = 1 - \exp(-\Omega^2/2)$. Hence, robust optimization ideas can also be used to deal with the following probability constraint

$$\mathbb{P} \left\{ G(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \geq 0 \right\} \geq 1 - \epsilon, \tag{3.4}$$

if we choose an appropriate constant Ω, such that $u_\Omega = 1 - \epsilon$.

In order to obtain the robust counterpart of problem (2.14) using robust optimization ideas, we introduce now the asymmetric distribution uncertainty set proposed by Chen et al. (2007).

To this end, let the function $G(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})$ have the following bilinear form in $\boldsymbol{x}$ and $\widetilde{\boldsymbol{\xi}}$:

$$G(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) = h_0(\boldsymbol{x}) + \sum_{i=1}^{m} h_i(\boldsymbol{x})\widetilde{\xi}_i. \tag{3.5}$$

Here $h_i(i = 0, 1, \cdots, m)$ is also a linear function of $\boldsymbol{x}$, and $\widetilde{\xi}_i(i = 0, 1, \cdots, m)$ is called the primitive uncertainty. Without loss of generality, we assume that $\widetilde{\boldsymbol{\xi}}$ satisfies the standard conditions[3],

$$\mathbb{E}[\widetilde{\boldsymbol{\xi}}] = \boldsymbol{0}, \quad and \quad \mathbb{E}[\widetilde{\boldsymbol{\xi}}\widetilde{\boldsymbol{\xi}}'] = \boldsymbol{I}, \tag{3.6}$$

and has support $\mathbb{S} = [-\boldsymbol{l}, \boldsymbol{u}]$, where $\boldsymbol{l} = (l_1, \cdots, l_m)'$, $\boldsymbol{u} = (u_1, \cdots, u_m)'$, $-\infty \le -l_i \le u_i \le \infty$, and $\boldsymbol{I}$ is an identity matrix. Consider the uncertainty set introduced by Chen, Sim and Sun (2007) as follows,

$$\mathcal{F}_\Omega = \left\{ \boldsymbol{\xi} : \exists \ \underline{\boldsymbol{u}}, \overline{\boldsymbol{u}} \in \mathbb{R}_+^m, \ \boldsymbol{\xi} = \overline{\boldsymbol{u}} - \underline{\boldsymbol{u}}, \ \|\boldsymbol{P}^{-1}\overline{\boldsymbol{u}} + \boldsymbol{Q}^{-1}\underline{\boldsymbol{u}}\|_2 \le \Omega, \boldsymbol{\xi} \in [-\boldsymbol{l}, \boldsymbol{u}]. \right\}, \tag{3.7}$$

where $\boldsymbol{P} = \mathrm{diag}(p_1, \cdots, p_m)$ and $\boldsymbol{Q} = \mathrm{diag}(q_1, \cdots, q_m)$, and the parameters $p_i = \sigma_f(\widetilde{\xi}_i) > 0$ and $q_i = \sigma_b(\widetilde{\xi}_i) > 0$ are the 'forward' and 'backward' deviations of random variable ξ_i, respectively. Clearly, uncertainty set $\mathcal{F}_\Omega$ is convex and compact with size controlled by the constant Ω.

As illustrated in Chen et al. (2007), if $\boldsymbol{P} = \boldsymbol{Q} = \boldsymbol{I}$ and $\boldsymbol{l} = \boldsymbol{u} = \infty$, then $\mathcal{U}_\Omega = \mathcal{F}_\Omega$. Hence, the uncertainty set $\mathcal{F}_\Omega$ generalizes the ellipsoidal uncertainty set $\mathcal{U}_\Omega$ used by El Ghaoui et al. (2003) in their worst-case VaR formulation. Intuitively, to capture distributional asymmetries, we decompose the primitive data uncertainty, $\widetilde{\boldsymbol{\xi}}$ into two random variables, $\overline{\widetilde{\boldsymbol{u}}} = \max\{\widetilde{\boldsymbol{\xi}}; 0\}$ and $\underline{\widetilde{\boldsymbol{u}}} = \max\{-\widetilde{\boldsymbol{\xi}}; 0\}$, such that $\widetilde{\boldsymbol{\xi}} = \overline{\widetilde{\boldsymbol{u}}} - \underline{\widetilde{\boldsymbol{u}}}$. The multipliers $\boldsymbol{P}^{-1}$ and $\boldsymbol{Q}^{-1}$ normalize the effective perturbation contributed by both $\overline{\widetilde{\boldsymbol{u}}}$ and $\underline{\widetilde{\boldsymbol{u}}}$ such that the norm of the aggregated values falls within the budget of uncertainty.

The 'forward' and the 'backward' deviations of a random variable $\widetilde{\xi}$ with mean value zero are defined by

$$p(\widetilde{\xi}) = \inf \left\{ \alpha_p : \alpha_p \ge 0, \mathbb{E}\left[\exp\left(\frac{\phi}{\alpha_p}\widetilde{\xi}\right)\right] \le \exp\left(\frac{\phi^2}{2}\right), \forall \ \phi \ge 0 \right\}, \tag{3.8}$$

$$q(\widetilde{\xi}) = \inf \left\{ \beta_q : \beta_q \ge 0, \mathbb{E}\left[\exp\left(-\frac{\phi}{\beta_q}\widetilde{\xi}\right)\right] \le \exp\left(\frac{\phi^2}{2}\right), \forall \ \phi \ge 0 \right\}. \tag{3.9}$$

Deviation measures $p(\widetilde{\xi})$ and $q(\widetilde{\xi})$ capture distributional asymmetry. It can be shown (Chen

---

[3]If $\widetilde{\boldsymbol{\xi}}$ has mean value $\boldsymbol{\mu}_\xi$ and non-diagonal covariance matrix $\boldsymbol{\Sigma}_\xi$, let $\widetilde{\boldsymbol{f}} = \boldsymbol{\Sigma}_\xi^{-1/2}(\widetilde{\boldsymbol{\xi}} - \boldsymbol{\mu}_\xi)$. Then $\widetilde{\boldsymbol{f}}$ satisfies the standard conditions (3.6). The conditions are also used by Chen, Sim and Sun (2007) and Natarajan, Pachamanova and Sim (2008).

et al. 2007) that for a random variable $\widetilde{\xi}$ with zero mean, $p(\widetilde{\xi})$ and $q(\widetilde{\xi})$ are always greater than or equal to the standard deviation of the distribution. In general, $p(\widetilde{\xi})$ and $q(\widetilde{\xi})$ are finite if the support $[-l, u]$ of the distribution for $\widetilde{\xi}$ is finite. If the support is infinite, $p(\widetilde{\xi})$ and $q(\widetilde{\xi})$ are not guaranteed to be finite. However, in the important case of a normally distributed random variable $\widetilde{\xi}$, $p(\widetilde{\xi})$ and $q(\widetilde{\xi})$ are finite, and equal to the standard deviation (Chen et al. 2007). The exact distribution of $\widetilde{\xi}$ is unknown in many practical problems. It is therefore important to be able to estimate $p(\widetilde{\xi})$ and $q(\widetilde{\xi})$ from the data.

Figure 1 gives two 2-dimension examples for the uncertainty set $\mathcal{F}_\Omega$. One can find from Figure 1 that $\mathcal{F}_\Omega$ is asymmetric with respect to the central point $(0, 0)$. Then the robust
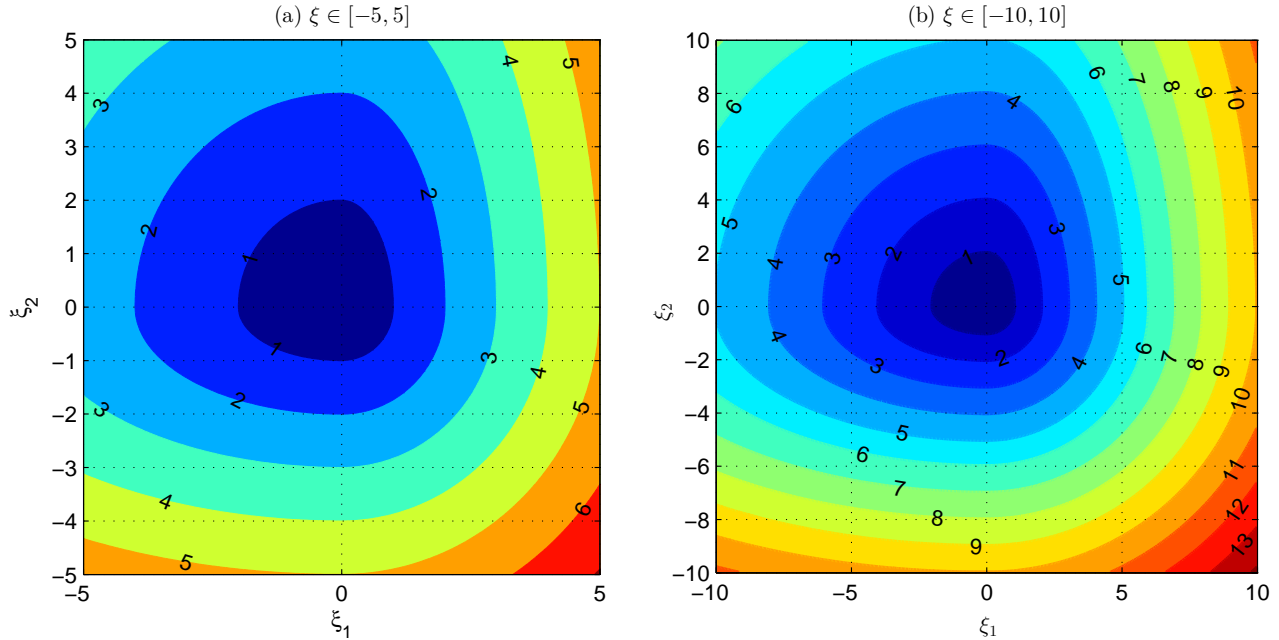

Figure 1: Two 2-dimension examples for uncertainty set $\mathcal{F}_\Omega$ with $\xi_i \in [-5, 5]$ and $\xi_i \in [-10, 10]$ for $i = 1, 2$.

counterpart of stochastic constraint $G(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \geq 0$ under uncertainty set $\mathcal{F}_\Omega$ is,

$$G(\boldsymbol{x}, \boldsymbol{\xi}) \geq 0, \quad \forall \, \boldsymbol{\xi} \in \mathcal{F}_\Omega. \tag{3.10}$$

The following two results are crucial and helpful for our subsequent analysis, and their proofs can be found in Chen, Sim and Sun.(2007).

Theorem 3.1 *If the distribution support $\mathbb{S}$ is finite, the robust counterpart (3.10) is equivalent*

11

to the following result,

$$\exists\, \boldsymbol{\zeta} \in \mathbb{R}^m, \boldsymbol{s}, \boldsymbol{v} \in \mathbb{R}^m_+,$$
$$\begin{cases} h_0(\boldsymbol{x}) \geq \Omega\|\boldsymbol{\zeta}\|_2 + \boldsymbol{s}'\boldsymbol{u} + \boldsymbol{v}'\boldsymbol{l}, \\ \zeta_i \geq -p_i(h_i(\boldsymbol{x}) + s_i - v_i), \quad \zeta_i \geq q_i(h_i(\boldsymbol{x}) + s_i - v_i), i = 1, \cdots, m. \end{cases} \tag{3.11}$$

Theorem 3.2 *Let the random vector $\widetilde{\boldsymbol{\xi}}$ satisfy standard conditions (3.6) and let $\boldsymbol{x}$ be feasible in the robust counterparts (3.11). Then we have that*

$$\mathbb{P}\left\{G(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \geq 0\right\} \geq 1 - \exp(-\Omega^2/2). \tag{3.12}$$

The following results give estimations of $p(\widetilde{\xi})$ and $q(\widetilde{\xi})$ if we have distribution information or historical data for $\widetilde{\xi}$.

Theorem 3.3 [32] *If either the distribution or a set of historical data of random variable $\widetilde{\xi}$ is known, then we can estimate $p(\widetilde{\xi})$ and $q(\widetilde{\xi})$ by:*

$$p(\widetilde{\xi}) = \sup_{\pi>0}\left\{\sqrt{2\frac{\ln\left(\mathbb{E}\left(\exp(\pi\widetilde{\xi})\right)\right)}{\pi^2}}\right\}, \quad q(\widetilde{\xi}) = \sup_{\pi>0}\left\{\sqrt{2\frac{\ln\left(\mathbb{E}\left(\exp(-\pi\widetilde{\xi})\right)\right)}{\pi^2}}\right\}.$$

4 Robust counterparts of multi-period portfolio problems

In this section, we will discuss the robust counterpart of the multi-period portfolio problem (2.14) under asymmetric uncertainty set $\mathcal{F}_\Omega$. To this end, we assume that the uncertainty of $\widetilde{r}^t_i$ (for all $t = 1, \cdots, T$) of risky asset i is completely driven by m random factors, $\widetilde{\boldsymbol{\xi}}^t = (\xi^t_1, \cdots, \xi^t_m)' \in \mathbb{R}^m$, and that the returns vector $\widetilde{\boldsymbol{r}}^t$ can be expressed linearly by these factors. Then, from (2.7), noting that $\boldsymbol{r}^0$ is known, we can assume that the scaled returns vector $\widetilde{\boldsymbol{R}}^{t+1}$ (that in fact corresponds to $\widetilde{\boldsymbol{r}}^t$) has the expression

$$\widetilde{\boldsymbol{R}}^{t+1} = \boldsymbol{\mu}^{t+1} + \boldsymbol{A}^{t+1}\widetilde{\boldsymbol{\xi}}^t, \quad t = 1, \cdots, T, \tag{4.1}$$

where $\boldsymbol{\mu}^{t+1}$ is a constant vector, and $\boldsymbol{A}^{t+1} \in \mathbb{R}^{n\times m}$ is the factor loading matrix. For simplicity, assume that $\widetilde{\boldsymbol{\xi}}^t$ is serially uncorrelated [4], and for each t satisfies standard conditions (3.6) and

---

[4]The serially uncorrelated assumption is a technical convenience and means that the uncertainty in the $(t+1)^{th}$ period is independent of the uncertainty in the tth period. The assumption can be made when we estimate the parameters of an AR(1) process, see Section 5 for details.

has a support set $\mathbb{S}^t = [-\boldsymbol{l}^t, \boldsymbol{u}^t]$. Then clearly,

$$\mathbb{E}\left[\widetilde{\boldsymbol{R}}^{t+1}\right] = \boldsymbol{\mu}^{t+1}, \quad t = 1, \cdots, T. \tag{4.2}$$

We extend the asymmetric uncertainty set $\mathcal{F}_\Omega$ to the multi-period case:

$$\mathcal{F}_\Omega^t = \left\{\boldsymbol{\xi}^t : \exists \underline{\boldsymbol{u}}^t, \overline{\boldsymbol{u}}^t \in \mathbb{R}_+^m, \; \boldsymbol{\xi}^t = \overline{\boldsymbol{u}}^t - \underline{\boldsymbol{u}}^t, \; \|(\boldsymbol{P}^t)^{-1}\overline{\boldsymbol{u}} + (\boldsymbol{Q}^t)^{-1}\underline{\boldsymbol{u}}^t\|_2 \leq \Omega^t, \boldsymbol{\xi}^t \in [-\boldsymbol{l}^t, \boldsymbol{u}^t].\right\}, \tag{4.3}$$

where $\boldsymbol{P}^t = \mathrm{diag}(p_1^t, \cdots, p_m^t)$ and $\boldsymbol{Q}^t = \mathrm{diag}(q_1^t, \cdots, q_m^t)$, and the parameters $p_i^t = p(\widetilde{\xi}_i^t) > 0$ and $q_i^t = q(\widetilde{\xi}_i^t) > 0$ are the 'forward' and 'backward' deviations of random variable $\xi_i^t, i = 1, \cdots, m$ and $t = 1, \cdots, T$, respectively. In the following, what we are interested in is the robust counterpart of (2.14) when the random factors $\widetilde{\boldsymbol{\xi}}^t$ vary in the uncertainty set $\mathcal{F}_\Omega^t$. To this end, we introduce some notation. Let

$$h_0^t(\boldsymbol{y}^t, \boldsymbol{z}^{t+1}) = w_0^t - w_0^{t+1} + \left(\frac{1-\theta}{R_0^{t+1}}\boldsymbol{z}^{s,t+1} - \frac{1+\theta}{R_0^{t+1}}\boldsymbol{z}^{b,t+1}\right)\boldsymbol{\mu}^{t+1}, \quad t = 1, \cdots, T-1, \tag{4.4}$$

and

$$h_j^t(\boldsymbol{y}^t, \boldsymbol{z}^{t+1}) = \boldsymbol{e}_j{}'\left(\boldsymbol{A}^{t+1}\right)'\left(\frac{1-\theta}{R_0^{t+1}}\boldsymbol{z}^{s,t+1} - \frac{1+\theta}{R_0^{t+1}}\boldsymbol{z}^{b,t+1}\right), \quad j = 1, \cdots, m, \quad t = 1, \cdots, T-1, \tag{4.5}$$

where $\boldsymbol{y}^t = (w_0^t; \boldsymbol{w}^t) \in \mathbb{R}^{n+1}$, $\boldsymbol{z}^{t+1} = (\boldsymbol{z}^{b,t+1}; \boldsymbol{z}^{s,t+1}) \in \mathbb{R}^{2n}$, and $\boldsymbol{e}_j$ is an m dimensional vector with one at the jth position and zeros otherwise. We rewrite the stochastic inequality constraint (2.13) as follows

$$w_0^t - w_0^{t+1} + \frac{(1-\theta)}{R_0^{t+1}}\left(\widetilde{\boldsymbol{R}}^{t+1}\right)'\boldsymbol{z}^{s,t+1} - \frac{(1+\theta)}{R_0^{t+1}}\left(\widetilde{\boldsymbol{R}}^{t+1}\right)'\boldsymbol{z}^{b,t+1} \geq 0, \; t = 1, \cdots, T-1. \tag{4.6}$$

Substitute (4.1) into the stochastic constraint (4.6). Then, from (4.1)–(4.5), the stochastic constraint (4.6) is equivalent to the following bilinear function in $(\boldsymbol{y}^t; \boldsymbol{z}^{t+1})$ and $\widetilde{\boldsymbol{\xi}}^t$,

$$G^t(\boldsymbol{y}^t, \boldsymbol{z}^{t+1}; \widetilde{\boldsymbol{\xi}}^t) = h_0^t(\boldsymbol{y}^t, \boldsymbol{z}^{t+1}) + \sum_{j=1}^m \widetilde{\xi}_j^t h_j^t(\boldsymbol{y}^t, \boldsymbol{z}^{t+1}) \geq 0, \quad t = 1, \cdots, T-1. \tag{4.7}$$

Suppose further

$$g(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}, \widetilde{\boldsymbol{\xi}}^t) = w_0^T R_0^{T+1} + (\widetilde{\boldsymbol{R}}^{T+1})'\boldsymbol{w}^T, \tag{4.8}$$

$$h_0^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) = w_0^T R_0^{T+1} + (\boldsymbol{\mu}^{T+1})'\boldsymbol{w}^T, \tag{4.9}$$

and

$$h_j^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) = \boldsymbol{e}_j'\left(\boldsymbol{A}^{T+1}\right)'\boldsymbol{w}^T, \; j = 1, \cdots, m. \tag{4.10}$$

13

Then, from (4.1),

$$
\begin{aligned}
g(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}, \widetilde{\boldsymbol{\xi}}) &= w_0^T R_0^{T+1} + (\widetilde{\boldsymbol{R}}^{T+1})' \boldsymbol{w}^T \\
&= \left(w_0^T R_0^{T+1} + (\boldsymbol{\mu}^{T+1})' \boldsymbol{w}^T \right) + \sum_{i=1}^{m} \widetilde{\xi}_j \left(\boldsymbol{e}_j' \left(\boldsymbol{A}^{T+1} \right)' \boldsymbol{w}^T \right) \\
&= h_0^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) + \sum_{i=1}^{m} \widetilde{\xi}_i h_j^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}).
\end{aligned}
\tag{4.11}
$$

Under the notation above, we can rewrite the original problem (2.14) as

$$
\text{(MLPM)}: \quad \min_{(\boldsymbol{y}^t, \boldsymbol{z}^{t+1}) \in \mathcal{X}} \quad \rho(a, \lambda; g(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}, \widetilde{\boldsymbol{\xi}}^t)) = -\mathbb{E}[g(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}, \widetilde{\boldsymbol{\xi}}^t)] + \lambda \mathbb{E}\left[\left(a - g(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}, \widetilde{\boldsymbol{\xi}}^t) \right)_+ \right]
$$

$$
\text{s.t.} \quad G^t(\boldsymbol{y}^t, \boldsymbol{z}^{t+1}; \widetilde{\boldsymbol{\xi}}^t) = h_0^t(\boldsymbol{y}^t, \boldsymbol{z}^{t+1}) + \sum_{j=1}^{m} \widetilde{\xi}_j^t h_j^t(\boldsymbol{y}^t, \boldsymbol{z}^{t+1}) \geq 0, \quad t = 1, \cdots, T-1,
$$

where $\boldsymbol{y}^t = (w_0^t; \boldsymbol{w}^t) \in \mathbb{R}^{n+1}$, $\boldsymbol{z}^{t+1} = (\boldsymbol{z}^{b,t+1}; \boldsymbol{z}^{s,t+1}) \in \mathbb{R}^{2n}$ and

$$
\mathcal{X} = \left\{ (\boldsymbol{y}^t, \boldsymbol{z}^{t+1}) : \begin{array}{l}
R_0^1 w_0^1 + \sum_{i=1}^{n} R_i^1 w_i^1 = 1, \quad i = 1, \cdots, n, \\
w_i^{t+1} = w_i^t + z_i^{b,t+1} - z_i^{s,t+1} \quad i = 1, \cdots, n, \quad t = 1, 2, \cdots, T, \\
w_i^t \geq 0, \quad i = 0, 1, \cdots, n, t = 1, \cdots, T, \\
z_i^{b,t} \geq 0, z_i^{s,t} \geq 0, i = 1, \cdots, n, t = 1, \cdots, T.
\end{array} \right\}
$$

is the feasible set formed by the deterministic constraints.

Theorem 4.1 *The objective of (2.14) or problem* MLPM *has a tight upper bound*

$$
\begin{aligned}
\rho(a, \lambda; g(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}, \widetilde{\boldsymbol{\xi}}^T)) &= -\mathbb{E}[g(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}, \widetilde{\boldsymbol{\xi}}^T)] + \lambda \mathbb{E}\left[\left(a - g(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}, \widetilde{\boldsymbol{\xi}}^T) \right)_+ \right] \\
&\leq -(1+\lambda) h_0^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) + \lambda a - \lambda(\gamma - 1)\widehat{\beta} + \lambda \left(-\widehat{\beta} \right)_+ \\
&\quad + \lambda \gamma \left(\left(h_0^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) - a \right) + \max_{\boldsymbol{\xi}^T \in \mathcal{F}_\Omega} \sum_{j=1}^{m} h_j^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) \xi_j^T \right),
\end{aligned}
\tag{4.12}
$$

where $h_0^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) = w_0^T R_0^{T+1} + (\boldsymbol{\mu}^T)' \boldsymbol{w}^T$, $h_j^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) = \boldsymbol{e}_j' \left(\boldsymbol{A}^{T+1} \right)' \boldsymbol{w}^T (j = 1, \cdots, m)$, $\gamma = \exp(-\Omega^2/2) > 0$ *is a constant, and*

$$
\widehat{\beta} = h_0^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) - a + \frac{\Omega^2 - 1}{\Omega} \|\boldsymbol{d}^T\|_2,
$$

where $\boldsymbol{d}^T = (d_1^T, \cdots, d_m^T)'$ *with* $d_j^T = \max \left\{ p_j^T h_j^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}), -q_j^T h_j^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) \right\}, j = 1, \cdots, m.$

Proof. See the support information in Online Appendix. □

Denote the right hand side of (4.12) by

$$
\begin{aligned}
\psi(a, \lambda; g(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}, \boldsymbol{\xi}^T)) = \ & -(1+\lambda)h_0^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) + \lambda a - \lambda(\gamma - 1)\widehat{\beta} + \lambda \left(-\widehat{\beta}\right)_+ \\
& + \lambda\gamma\left((h_0^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) - a) + \max_{\boldsymbol{\xi}^T \in \mathcal{F}_\Omega} \sum_{j=1}^m h_j^T(\boldsymbol{y}^T, \boldsymbol{z}^{t+1})\xi_j^T\right).
\end{aligned}
$$

Then $\psi(a, \lambda; g(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}, \boldsymbol{\xi}^T))$ is a tight upper bound of the original objective $\rho(a, \lambda; g(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}, \widetilde{\boldsymbol{\xi}}^T))$. We regard ψ as an approximation of ρ and replace ρ with ψ in (MLPM) in the rest of this paper. From the proof of Theorem 4.1 (see, e.g. the online support information). We have that

$$
\max_{\boldsymbol{\xi}^T \in \mathcal{F}_\Omega} \sum_{j=1}^m h_j^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1})\xi_j^T = \min_{\boldsymbol{d}^T}\{\Omega\|\boldsymbol{d}^T\| : \quad \boldsymbol{P}^{-1}\boldsymbol{d}^T \geq \boldsymbol{h}^T, \boldsymbol{Q}^{-1}\boldsymbol{d}^h \geq -\boldsymbol{h}^T\}.
$$

Note that $(x)_+ = x + (-x)_+$ again and $\widehat{\beta}$ is given in Theorem 4.1, then the approximated objective function is

$$
\begin{aligned}
\psi(a, \lambda; g(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}, \boldsymbol{\xi}^T)) \ &= (\lambda\gamma - (1+\lambda))h_0^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) + \lambda a(1-\gamma) - \lambda\gamma\widehat{\beta} + \lambda\left(\widehat{\beta}\right)_+ + \lambda\gamma\Omega\|\boldsymbol{d}^T\| \\
&= -(\gamma+1)h_0^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) + \lambda(\widehat{\beta})_+ + \frac{\lambda\gamma}{\Omega}\|\boldsymbol{d}^T\| + \lambda a.
\end{aligned}
$$

Next we return to the stochastic inequality constraints of problem (MLPM). With the uncertainty set $\mathcal{F}_\Omega$ and (3.10), the robust counterpart of the stochastic inequality constraint is the determinate inequality,

$$
G^t(\boldsymbol{y}^t, \boldsymbol{z}^{t+1}; \boldsymbol{\xi}^t) \geq 0, \quad \boldsymbol{\xi}^t \in \mathcal{F}_\Omega^t, \ \text{ for } \text{ each } \ t = 1, \cdots, T-1. \tag{4.13}
$$

For any $\epsilon_t > 0$ ($t = 1, \cdots, T-1$), from Theorem 3.2, the feasible solution $(\boldsymbol{y}^t, \boldsymbol{z}^{t+1})$ satisfying inequality (4.7) has a probability guarantee $u_{\Omega_t} = 1 - \epsilon_t = 1 - \exp(-\Omega_t^2/2)$, where $\Omega_t = \sqrt{-2\ln(\epsilon_t)}$ ($t = 1, \cdots, T-1$). The robust counterpart (4.13) is equivalent with the following cone inequality constraints which are denoted by the feasible set $\mathcal{Z}$ of $(w_0^t, \boldsymbol{s}^t, \boldsymbol{v}^t, \boldsymbol{\zeta}^t, \boldsymbol{z}^t)$,

$$
\left\{
\begin{aligned}
& w_0^t - w_0^{t+1} + \left(\frac{1-\theta}{R_0^{t+1}}\boldsymbol{z}^{s,t+1} - \frac{1+\theta}{R_0^{t+1}}\boldsymbol{z}^{b,t+1}\right)\boldsymbol{\mu}^{t+1} \geq \Omega^t\|\boldsymbol{\zeta}^t\|_2 + (\boldsymbol{s}^t)'\boldsymbol{u}^t + (\boldsymbol{v}^t)'\boldsymbol{l}^t, \ t = 1, \cdots, T-1, \\
& \zeta_j^t \geq -p_j^t\left(\boldsymbol{e}_j'(\boldsymbol{A}^{t+1})'\left(\frac{1-\theta}{R_0^{t+1}}\boldsymbol{z}^{s,t+1} - \frac{1+\theta}{R_0^{t+1}}\boldsymbol{z}^{b,t+1}\right) + s_j^t - v_j^t\right), j = 1, \cdots, m, \ t = 1, \cdots, T-1, \\
& \zeta_j^t \geq q_j^t\left(\boldsymbol{e}_j'(\boldsymbol{A}^{t+1})'\left(\frac{1-\theta}{R_0^{t+1}}\boldsymbol{z}^{s,t+1} - \frac{1+\theta}{R_0^{t+1}}\boldsymbol{z}^{b,t+1}\right) + s_j^t - v_j^t\right), j = 1, \cdots, m, \ t = 1, \cdots, T-1, \\
& \boldsymbol{s}^t = (s_1^t, \cdots, s_m^t)', \boldsymbol{v}^t = (v_1^t, \cdots, v_m^t)' \geq \boldsymbol{0}, \boldsymbol{\zeta}^t = (\zeta_1^t, \cdots, \zeta_m^t)' \in \mathbb{R}^m, t = 1, \cdots, T.
\end{aligned}
\right\}
$$

In summary, we have the robust counterpart of original problem, denoted by RLPM:

$$(\text{RLPM}): \quad \min_{\substack{\boldsymbol{y}^t, \boldsymbol{z}^t, \boldsymbol{d}^T, \tau \\ \tau^h, \boldsymbol{s}^t, \boldsymbol{v}^t, \boldsymbol{\zeta}^t}} \quad -(\gamma+1)h_0^T(\boldsymbol{y}^T, \boldsymbol{z}^{T+1}) + \lambda\tau + \frac{\lambda\gamma}{\Omega}\tau^h$$

$$\text{s.t.} \quad \boldsymbol{P}^{-1}\boldsymbol{d}^T \geq \boldsymbol{h}^T, \quad \boldsymbol{Q}^{-1}\boldsymbol{d}^h \geq -\boldsymbol{h}^T, \quad \|\boldsymbol{d}^T\| \leq \tau^h, \tau \geq \widehat{\beta} \qquad (4.14)$$

$$w_0^T R_0^{T+1} + (\boldsymbol{\mu}^{T+1})'\boldsymbol{w}^T + \frac{\Omega^2-1}{\Omega}\|\boldsymbol{d}^T\|_2 \leq \tau + a,$$

$$\tau \geq 0, \tau^h \geq 0, (\boldsymbol{y}^t, \boldsymbol{s}^t, \boldsymbol{v}^t, \boldsymbol{\zeta}^t, \boldsymbol{z}^t) \in \mathcal{X} \bigcap \mathcal{Z}, t = 1, \cdots, T-1,$$

where we drop the several constant terms in objective. The inequality constraints $\widehat{\beta} \leq \tau \geq 0$ are from the transformations of $(\widehat{\beta})_+ = \tau$. The following probability guarantee can be obtained.

Theorem 4.2 *For any $\lambda \geq 1$, $\epsilon > 0$ with $\epsilon_t = \epsilon/(T-1)$, $\Omega = \Omega_t = \sqrt{-2\ln(\epsilon_t)}(t = 1, \cdots, T$, and $\gamma = \exp(-\Omega^2/2)$, then for any feasible solution $(\boldsymbol{y}^t, \boldsymbol{z}^{t+1})$ of problem (RLPM), the stochastic constraint of the original problem (2.14) satisfies*

$$\mathbb{P}\left\{G^t(\boldsymbol{y}^t, \boldsymbol{z}^{t+1}; \widetilde{\boldsymbol{\xi}}^t) \geq 0, \right\} \geq 1 - \epsilon_t, \ t = 1, \cdots, T-1, \qquad (4.15)$$

$$\mathbb{P}\left\{G^t(\boldsymbol{y}^t, \boldsymbol{z}^{t+1}; \widetilde{\boldsymbol{\xi}}^t) \geq 0, t = 1, \cdots, T-1\right\} \geq 1 - \epsilon, \qquad (4.16)$$

where $G^t(\boldsymbol{y}^t, \boldsymbol{z}^{t+1}; \widetilde{\boldsymbol{\xi}}^t)$ is given by (4.7).

The parameter Ω must satisfy $\Omega \geq 1$, to ensure the convex objective and the convexity of the constraint $\widehat{\beta} - a \leq \tau$ of problem (RLPM). This condition can be attainted easily if $\epsilon_t \leq \exp(-1/2)$ for $t = 1, \cdots, T-1$.

The impact of Theorem 4.2 is double. On the one hand, one can always take a desired confidence level ϵ, such as $\epsilon = 0.05$, such that the probability guarantee (4.16) holds. However, on the other hand, this will lead to a very conservative individual constraint, in other words, if $\epsilon = 0.05$, then we must take $\epsilon_t \leq \epsilon/(T-1) = 0.0005$ and $\mathbb{P}\left\{G^t(\boldsymbol{y}^t, \boldsymbol{z}^{t+1}; \widetilde{\boldsymbol{\xi}}^t) \geq 0\right\} \geq 0.9995$ for $T = 101$. Indeed, this is too conservative and is unpractical. In Section 5, we will provide a simple approach to decrease T, but keep the whole investment period unchanged.

4.1   The Relationship Between RLPM and CVaR

Let $f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})$ be a loss (of wealth) function associated with the portfolio $\boldsymbol{x} \in \mathcal{X} \subseteq \mathbb{R}^{n+1}$ and the random vector $\widetilde{\boldsymbol{\xi}} \in \mathbb{R}^m$.

VaR is defined as the minimal level β, such that the probability that the portfolio loss $f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})$ exceeds β is below α, where $\alpha \in (0, 1)$ is the probability level specified by the user

16

(typically, $\alpha = 1\%$ *or* 5%):

$$\mathrm{VaR}_{1-\alpha}(f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})) = \min \left\{ \beta \mid \mathbb{P}\{f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \geq \beta\} \leq \alpha \right\},$$

CVaR measures the conditional expected loss of portfolio subject to $f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \geq \mathrm{VaR}_{1-\alpha}$. Mathematically, the $\mathrm{CVaR}_{1-\alpha}(f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}))$ measure is defined as (Rockafellar and Uryasev 2000)

$$\mathrm{CVaR}_{1-\alpha}(f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})) = \mathbb{E}\left[f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \mid f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \geq \mathrm{VaR}_{1-\alpha} \right].$$

For given α, if we take the target a of LPM as $a = 1 - \mathrm{VaR}_{1-\alpha}$, and $g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) = 1 - f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})$. Then,

$$\begin{aligned} \mathrm{LPM}(a; g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})) &= \mathbb{E}\left[a - g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \right]_+ = \mathbb{E}\left\{ a - g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \mid a - g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \geq 0 \right\} \\ &= \mathbb{E}\left\{ f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) - \mathrm{VaR}_{1-\alpha}(f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})) \mid f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) \geq \mathrm{VaR}_{1-\alpha} \right\} \end{aligned}$$

If α is taken such that the maximum loss $\mathrm{VaR}_{1-\alpha}$ is nonnegative. Then, clearly,

$$\mathrm{LPM}(a; g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})) \leq \mathrm{CVaR}_{1-\alpha}(f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})).$$

Under the conditions that α and the distribution of $\widetilde{\boldsymbol{\xi}}$ are given, such as $\mathrm{VaR}_{1-\alpha}$ is determinate, then we further have that

$$\mathrm{LPM}(a; g(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})) = \mathrm{CVaR}_{1-\alpha}(f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})) - \mathrm{VaR}_{1-\alpha}(f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})).$$

In this case, the minimization of CVaR is equivalent to the minimization of the sum of LPM and VaR. Thus, the differences between mean-CVaR and mean-LPM are clear. First, in the minimization of CVaR problem, we can understand that the target $\mathrm{VaR}_{1-\alpha}(f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}))$ is in fact endogenous. But in LPM, the target a is exogenous.

Second, the advantage of the exogenous a can avoid the effect of uncertainty set in robust portfolio choice which can reduce a chance constraint in multi-period case. To see this, applying the coherent property of CVaR, we have

$$\min_{\boldsymbol{x}} \mathrm{CVaR}_{1-\alpha}(f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}})) = \min_{\boldsymbol{x}, \gamma_0} \left\{ \gamma_0 \mid \mathrm{CVaR}_{1-\alpha}(f(\boldsymbol{x}, \widetilde{\boldsymbol{\xi}}) + \gamma_0) \leq 0 \right\}. \tag{4.17}$$

The right hand side of problem (4.17) includes a CVaR constraint which is still difficult to be expressed explicitly when the distribution information $\widetilde{\boldsymbol{\xi}}$ is only given by (3.7). However, we find

that the CVaR constraint can be viewed as an approximation of the chance constraint, namely,

$$\text{CVaR}_{1-\epsilon}(f(\boldsymbol{x},\widetilde{\boldsymbol{\xi}})+\gamma_0) \leq 0 \Rightarrow \mathbb{P}\left\{f(\boldsymbol{x},\widetilde{\boldsymbol{\xi}})+\gamma_0 \geq 0\right\} \geq 1-\epsilon, \tag{4.18}$$

which is not the sufficient condition. As stated in Natarajan, Pachamanova and Sim (2008), to solve the minimization of CVaR under the uncertainty set $\mathcal{F}_\Omega$, it needs to add a new chance constraint (the right hand side of (4.18)), which will lead to a more conservative solution from Theorem 4.2 than LPM.

Third, when the exogenous target a is taken not too extreme value, it can help LPM reduce the variance of portfolio. For example, if a is close to the mean return of portfolio, then LPM in fact controls the down-side semi-variance which clearly is important to limit the variance. Our numerical tests in Section 5 validate this statement.

Interestingly, we may give the reference value of exogenous a from VaR. In view of this, let the returns of a portfolio $\widetilde{r}_P(\widetilde{\boldsymbol{\xi}})$ be normal distribution. Using $a = 1 - \text{VaR}$, then we have that[5]

$$a = 1 - \mathcal{Z}_{1-\alpha}\sigma_{\widetilde{r}_P}. \tag{4.19}$$

where $\mathcal{Z}_{1-\alpha}$ is the quantile with the confidence level $1-\alpha$. For example, under $\sigma_{\widetilde{r}_P} = 10\%$, and $\mu_{\widetilde{r}_P} = 0$, the target $a = 0.767$ for $\alpha = 0.01$ with $\mathcal{Z}_{1-\alpha} = 2.33$, $a = 0.804$ for $\alpha = 0.05$ with $\mathcal{Z}_{1-\alpha} = 1.96$, and $a = 0.835$ for $\alpha = 0.10$ with $\mathcal{Z}_{1-\alpha} = 1.65$. In the numerical tests, we will take exogenously the value a referring to (4.19) for the volatility of samples.

5 Numerical results

5.1 Parameters estimation approach

Let $\widetilde{r}_i^t = (\widetilde{S}_i^t - \widetilde{S}_i^{t-1})/\widetilde{S}_i^{t-1}$ be the average asset return of risky asset i at time t. Denote its logarithmic return by

$$\widetilde{\boldsymbol{Y}}^t = \ln(1+\widetilde{\boldsymbol{r}}_t), \quad with \quad \widetilde{Y}_i^t = \ln\frac{\widetilde{S}_i^t}{\widetilde{S}_i^{t-1}} = \ln(1+\widetilde{r}_t^i), \quad \widetilde{\boldsymbol{Y}}^t = (\widetilde{Y}_1^t,\cdots,\widetilde{Y}_n^t)'. \tag{5.1}$$

---

[5]VaR can also be calculated by under normal distribution,

$$\text{VaR} = \mathcal{Z}_{1-\alpha}\sigma_{\widetilde{r}_P} + \mu_{\widetilde{r}_P}.$$

In our numerical tests, we set target a using the calculation of Hull (2015) who ignores the expected return because the expected return is too small relative to volatility. See Hull, John, Options, futures and other derivatives. 9th ed. Upper Saddle River, NJ: Prentice Hall, Chapter 22: Value at risk for detail.

In order to predict the returns, we assume that $\{\widetilde{\boldsymbol{Y}}^t\}$ is a VAR(1) process:

$$\widetilde{\boldsymbol{Y}}^t = \boldsymbol{\nu} + \boldsymbol{\Phi}\boldsymbol{Y}^{t-1} + \widetilde{\boldsymbol{\eta}}^t, \quad \widetilde{\boldsymbol{\eta}}^t \sim \text{MVN}\,[\boldsymbol{0}, \boldsymbol{\Sigma}]. \tag{5.2}$$

Assume that the covariance matrix $\boldsymbol{\Sigma}$ of $\boldsymbol{\eta}^t$ is fixed for all t, and $\widetilde{\boldsymbol{\eta}}^t = \widetilde{\boldsymbol{\eta}} \sim \text{MVN}\,(\boldsymbol{0}, \boldsymbol{\Sigma})$. Take $\widetilde{\boldsymbol{\xi}} = \boldsymbol{\Sigma}^{-1/2}\widetilde{\boldsymbol{\eta}}$, then $\widetilde{\boldsymbol{\xi}}$ satisfies the standard conditions (3.6) and (5.2), and can be rewritten as

$$\widetilde{\boldsymbol{Y}}^t = \boldsymbol{\nu} + \boldsymbol{\Phi}\boldsymbol{Y}^{t-1} + \boldsymbol{\Sigma}^{1/2}\widetilde{\boldsymbol{\xi}}. \tag{5.3}$$

Let the current time be $t = 0$, and $\widehat{\boldsymbol{\nu}}, \widehat{\boldsymbol{\Phi}}$ and $\widehat{\boldsymbol{\Sigma}}$ be the estimation of $\boldsymbol{\nu}, \boldsymbol{\Phi}$ and $\boldsymbol{\Sigma}$ for an observed sample series $\{\boldsymbol{Y}^t\}_{t \geq -T_0}^0$. Then we have that

$$\widehat{\boldsymbol{\mu}}_Y^t = \mathbb{E}\left[\widetilde{\boldsymbol{Y}}^t | \boldsymbol{Y}^{t-1}\right] = \widehat{\boldsymbol{\nu}} + \widehat{\boldsymbol{\Phi}}\boldsymbol{Y}^{t-1}.$$

Let $\{\widehat{\boldsymbol{\xi}}^t\}_{t \geq -T_0}^0$ be the residual error series with the same length as series $\{\boldsymbol{Y}^t\}_{t \geq -T_0}$. Then, we can view $\{\widehat{\boldsymbol{\xi}}^t\}_{t \geq T_0}^0$ as the T_0 observations of $\widetilde{\boldsymbol{\xi}}$. So, using Theorem 3.3, we can obtain the estimation of p_i and q_i $(i = 1, \cdots, m)$. For simplicity, we assume that p_i and q_i $(i = 1, \cdots, m)$ are fixed for all periods.[6] In multi-period portfolio optimization, we need to estimate $\boldsymbol{\mu}^t$ in advance. Thus, we need to estimate the possible realizations of $\{\boldsymbol{Y}^t\}$ in advance using the following equation:

$$\widetilde{\boldsymbol{Y}}^t = \widehat{\boldsymbol{\nu}} + \widehat{\boldsymbol{\Phi}}\boldsymbol{Y}^{t-1} + \widehat{\boldsymbol{\Sigma}}^{1/2}\widetilde{\boldsymbol{\xi}}, \quad t > 0.$$

The key step to estimate $\{\boldsymbol{Y}^t\}$ in advance is to estimate $\widetilde{\boldsymbol{\xi}}$. Denote the support set of $\widetilde{\boldsymbol{\xi}} = (\widetilde{\xi}_1, \cdots, \widetilde{\xi}_n)'$ by $\mathbb{S} = [-\boldsymbol{l}, \boldsymbol{u}](-\infty \leq -l_i \leq u_i \leq \infty)^7$, and assume that $\widetilde{\xi}_i$ instead of the normal distribution, follows the following distribution, $\mathrm{P}\{\widetilde{\xi}_i = \xi_i^\ell\} = \alpha_\ell^i, \quad \ell = 1, \cdots, J$, where J is an even number and

$$\xi_i^\ell = \frac{(-1)^\ell \sqrt{\prod\limits_{\ell=1}^J \alpha_\ell^i}}{\alpha_\ell^i}. \tag{5.4}$$

For each $i = 1, \cdots, n$, α_ℓ^i is computed by

$$\alpha_\ell^i = \frac{\ell}{J}\left(\frac{1}{\ell} + \frac{2i}{J(J-1)(n+1)}\right), \ell = 1, \cdots, J-1, \quad \text{and} \quad \alpha_J^i = 1 - \sum_{\ell=1}^{J-1} \alpha_\ell^i.$$

---

[6]We also can estimate p_i and q_i by simulating the realizations of $\widetilde{\boldsymbol{\xi}}$, see the following GSP.

[7]In this section and next section, we fix support set $\mathbb{S}$ with $l_i = u_i = 10, i = 1, \cdots, m$.

In particular, when $J = 2$,

$$\alpha_1^i = \frac{1}{2}\left(1 + \frac{i}{n+1}\right), \quad \text{and} \quad \alpha_2^i = 1 - \alpha_1^i.$$

This has been considered by Natarajan, Pachamanova and Sim (2008) and Gulpinar, Canakoglu and Pachamanova (2014). Thus, the assumption of distribution of the random variables $\widetilde{\xi}_i$ $(i = 1, \cdots, n)$ is a generalization of Natarajan, Pachamanova and Sim (2008). Note that $\widetilde{\xi}_i$ is a discrete random variable and J gives the number of possible states of the random variable $\widetilde{\xi}_i$. The shock of returns $\widetilde{\boldsymbol{Y}}^t$ is driven by the random variable $\widetilde{\xi}_i$ and therefore, in general, the number of states of random variable $\widetilde{\xi}_i$ is large in real markets. Thus, a large J is more appropriate than a small J to fit the stochastic shocks of real markets. Several related numerical experiments are designed by Bertsimas and Pachamanova (2008), Gulpinar and Pachamanova (2013) to simulate data series, see also Gulpinar, Pachamanova and Canakoglu (2016) for details. We consider a similar simulation procedure by taking $J = 4$ for computational purposes. Formally, we state the following general simulation procedure (GSP) to obtain a data sequence of $\{\boldsymbol{Y}^t\}_{t=0}^T$

General Simulation Procedure (GSP):

1. Input an initial $\widehat{\boldsymbol{Y}}^0$, two integers K and T. Set $k = 1, t = 0$.

2. Generate $\boldsymbol{\xi}$ independently using probability distribution (5.4) for every k, and compute the corresponding $\overline{\boldsymbol{Y}}_k^{t+1}$ by

$$\overline{\boldsymbol{Y}}_k^{t+1} = \widehat{\boldsymbol{\nu}} + \widehat{\boldsymbol{\Phi}}\,\widehat{\boldsymbol{Y}}^t + \widehat{\boldsymbol{\Sigma}}^{1/2}\boldsymbol{\xi}, \quad k = 1, \cdots, K \tag{5.5}$$

3. Put $\widehat{\boldsymbol{Y}}^{t+1} = (1/K)\Sigma_{k=1}^K \overline{\boldsymbol{Y}}_k^{t+1}$, and then set $t = t + 1$.

4. If $t > T$, output the sequence of $\{\widehat{\boldsymbol{Y}}^t\}_{t=0}^T$, and stop. Otherwise, go to Step 2.

K is the times of simulations between the t-th and the $(t+1)$-th period. The simulated $\boldsymbol{\xi}$ can be viewed as the realizations of $\widetilde{\boldsymbol{\xi}}$ and therefore can be used to estimate p_i, q_i from Theorem 3.3.[8] The output series $\{\boldsymbol{Y}^t\}_{t=0}^T$ generated by (GSP) can be used to estimate model parameters $\boldsymbol{\mu}^t, \boldsymbol{A}^t$.

$$\widetilde{\boldsymbol{R}}^{t+1} - 1 \approx \ln\left(1 + (\widetilde{\boldsymbol{R}}^{t+1} - 1)\right) = \ln\left(\widetilde{\boldsymbol{R}}^{t+1}\right) = \sum_{\ell=0}^t \ln\left(1 + \widetilde{\boldsymbol{r}}^\ell\right) = \sum_{\ell=0}^t \widetilde{\boldsymbol{Y}}^\ell.$$

---

[8]We actually generate $T * K$ samples of random vectors $\widetilde{\boldsymbol{\xi}}$ independently when we generate sequence $\{\boldsymbol{Y}^t\}$ using (GSP). These random vector samples $\{\widetilde{\boldsymbol{\xi}}_{t=1}^{t=T*K}\}$ are viewed as the realizations of $\widetilde{\boldsymbol{\xi}}$ and therefore can be used to estimate p_i and q_i which are the same for all t.

Thus,

$$\widetilde{\boldsymbol{R}}^{t+1} = \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \sum_{\ell=0}^{t} \widehat{\boldsymbol{\mu}}_Y^\ell}_{\widehat{\boldsymbol{\mu}}^{t+1}} + \underbrace{\sqrt{t}\widehat{\Sigma}^{1/2}}_{\widehat{\boldsymbol{A}}^{t+1}} \widetilde{\boldsymbol{\xi}} = \widehat{\boldsymbol{\mu}}^{t+1} + \widehat{\boldsymbol{A}}^{t+1}\widetilde{\boldsymbol{\xi}}. \tag{5.6}$$

where $\widehat{\boldsymbol{\mu}}^{t+1}$, $\widehat{\boldsymbol{A}}^{t+1}$ are the estimates of $\boldsymbol{\mu}^{t+1}$, $\boldsymbol{A}^{t+1}$ in (4.1).

We close this subsection by giving some remarks on how to mitigate the probability guarantee between the joint stochastic constraints and the individual stochastic constraint. As described in Theorem 4.2, in order to get a given probability guarantee, the optimal strategy will become more conservative when T is large. The problem can be mitigated by changing the adjustment frequency when the out-of-sample term is fixed.[9] For example, for a fixed 3-years out-of-sample investment term, assume that the probability guarantee of all joint stochastic constraints is set as $1 - \epsilon = 0.95$. Then $T = 36$ for monthly adjustment, corresponding to $\Omega = 3.62$ and the probability guarantee for each stochastic constraint is 0.9986. If we consider quarterly adjustment, then T is reduced to 12 and $\Omega = 3.29$, and the probability guarantee for each stochastic constraint is relaxed to 0.9955. Further, if we consider yearly adjustment, then T is reduced to 3 and $\Omega = 2.72$, and the probability guarantee for each stochastic constraint is relaxed to 0.9750, which is acceptable.

When the investment horizon is long, we can use the following M-stage rolling procedure (M-SRP) to obtain the appropriate probability guarantee. Let the investment horizon T' be long and the adjustment frequency be Δt years. Consider a T-period portfolio model with $T' = M * T$. The multi-stage rolling procedure for a long investment horizon, requires the T-period portfolio mode RLPM to be implemented M times. Specifically, at the kth ($k = 1, \cdots, M$) stage, we run the RLPM to get the optimal portfolio sequence $\{\mathbf{w_t^k}\}_{\mathbf{t=1:T}}$. For simplicity, we consider an example. Suppose that $T' = 30$ years, $\Delta t = 1$ year (annual adjustment), and $T = 10$. Then, we can take $M = 3$. We run the RLPM model to obtain the first 10-year portfolio sequence $\{\mathbf{w_t^1}\}_{\mathbf{t=1:10}}$ in which the optimal portfolio $\{\mathbf{w_t}\}$ is adjusted annually. Using the new historical data, we run the RLPM model again to get the second 10-year portfolio sequence $\{\mathbf{w_t^2}\}_{\mathbf{t=1:10}}$. Finally, the third 10-year portfolio sequence $\{\mathbf{w_t^3}\}_{\mathbf{t=1:10}}$ can be also obtained by running RLPM model at the third time. The whole 30-year portfolio sequence $\{\mathbf{w_t^1}\}_{\mathbf{t=1:10}}$, $\{\mathbf{w_t^2}\}_{\mathbf{t=1:10}}$, and $\{\mathbf{w_t^3}\}_{\mathbf{t=1:10}}$ with each sequence adjusted annually, can be obtained by running RLPM with a 3-stage rolling procedure.

In order to know the differences of performance for different models, we will make some comparisons among the proposed model, equally-weighted model ($1/N$), mean-WVaR, mean-

---

[9]There is much literature that discusses the approximately deterministic expression of joint stochastic constraints, such as Chen et al. (2010). The discussion of algorithms is clearly another important topic that is beyond the scope of this paper.

CVaR and mean-WCVaR [10]. Models including WVaR, CVaR, WCVaR and RLPM are certain moment minimization problems and are comparable. We will compare the expected return, volatility and Sharpe ratio with respect to risk averse parameter λ among these models, where three performance indices are calculated by

$$\bar{\mu} = \frac{1}{\Delta t}\left[\frac{1}{T}\sum_{t=1}^{T}((1 + \mathbf{x}^T\mathbf{r}_t)(1 - \theta\sum_{i=1}^{n}|\mathbf{x}_{i,t+1} - \mathbf{x}_{i,t}|) - 1)\right],$$

$$\sigma = \frac{1}{\Delta t}\sqrt{\left[\frac{1}{T-1}\sum_{t=1}^{T}((1 + \mathbf{x}^T\mathbf{r}_t)(1 - \theta\sum_{i=1}^{n}|\mathbf{x}_{i,t+1} - \mathbf{x}_{i,t}|) - 1 - \bar{\mu})^2\right]},$$

$$SR = (\bar{\mu} - r_f)/\sigma.$$

Here $\mathbf{r}_t$ is the realized return of $\widetilde{\mathbf{r}}_t$ at time t, T is the number of periods and Δt is the adjustment frequency of the strategy.

5.2 Numerical comparisons with the real market data

As the first example, we consider a portfolio consisting of five assets, which is 5Ind portfolios from the French data library [11]. The whole sample is from Jan. 2nd 1946 to Dec. 31st 2016 with Jan. 2nd 1946—Dec. 31st 1986 as in-sample data and Jan. 2nd 1987—Dec. 31st 2016 as out-of-sample data. The investment starts from Jan. 2nd 1987 with different investment horizons. Using the in-sample data, we can estimate the parameters $\boldsymbol{\nu}, \boldsymbol{\Phi}$ and $\boldsymbol{\Sigma}$ by a VAR(1) process (5.1) to obtain

$$\widehat{\boldsymbol{\nu}} = [-0.0013, 0.0010, -0.0040, -0.0007, 0.0037]',$$

---

[10] The compared models are described as follows.

- $\mathrm{Mean - WVaR}$(WVaR for short) : $\min_{\boldsymbol{x}}\left\{-\mathbb{E}[\boldsymbol{x}'\widetilde{\boldsymbol{r}}] + \lambda\mathrm{WVaR}_{1-\alpha} : \quad \boldsymbol{x} \geq 0, \sum_{i=1}^{n}x_i = 1\right\}$ with a normal distribution assumption. The α is the worst-case probability that the loss of portfolio is greater than worst-case VaR (WVaR), see El Ghaoui et al.(2003).

- $\mathrm{Mean - CVaR}$(CVaR for short) : $\min_{\boldsymbol{x}}\left\{-\mathbb{E}[\boldsymbol{x}'\widetilde{\boldsymbol{r}}] + \lambda\mathrm{CVaR}_{1-\alpha} : \quad \boldsymbol{x} \geq 0, \sum_{i=1}^{n}x_i = 1\right\}$ is solved by linear programming under the discrete distribution, see Rockafellar and Uryasev (2000) for details, where different α are taken.

- $\mathrm{Mean - WCVaR}$(WCVaR for short) : $\min_{\boldsymbol{x}}\left\{-\mathbb{E}[\boldsymbol{x}'\widetilde{\boldsymbol{r}}] + \lambda\mathrm{WCVaR}_{1-\alpha} : \quad \boldsymbol{x} \geq 0, \sum_{i=1}^{n}x_i = 1\right\}$ is solved by SOCP under the ellipsoidal uncertainty set $\mathcal{U}_{\Omega}$, see Zhu, Fukushima (2009) for details, where different α are taken.

[11]French-Data Library: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

$$\widehat{\boldsymbol{\Phi}} = \begin{bmatrix} 0.1909 & -0.1498 & 0.1461 & -0.0137 & 0.2181 \\ -0.1086 & -0.0250 & 0.2295 & -0.1858 & 0.2800 \\ 0.0986 & 0.2505 & 0.0081 & -0.1029 & -0.1215 \\ -0.0116 & 0.0501 & 0.0840 & -0.1741 & 0.0961 \\ 0.0209 & 0.0490 & 0.1070 & 0.1420 & -0.1301 \end{bmatrix},$$

$$\widehat{\boldsymbol{\Sigma}}^s = \begin{bmatrix} 0.0027 & 0.0026 & 0.0014 & 0.0019 & 0.0017 \\ 0.0026 & 0.0045 & 0.0022 & 0.0024 & 0.0023 \\ 0.0014 & 0.0022 & 0.0033 & 0.0015 & 0.0013 \\ 0.0019 & 0.0024 & 0.0015 & 0.0019 & 0.0016 \\ 0.0017 & 0.0023 & 0.0013 & 0.0016 & 0.0019 \end{bmatrix},$$

where $\widehat{\boldsymbol{\Sigma}}^s$ is the covariance matrix estimated from real market data. Then, the return series $\boldsymbol{Y}^t$ can be predicted as follows.

$$\widetilde{\boldsymbol{Y}}^t = \widehat{\boldsymbol{\nu}} + \widehat{\boldsymbol{\Phi}} \boldsymbol{Y}^{t-1} + (\widehat{\boldsymbol{\Sigma}}^s)^{1/2} \widetilde{\boldsymbol{\xi}}, \tag{5.7}$$

Assume that a new simulated sequence of $\widehat{\boldsymbol{Y}}^t$ using (GSP), for which K is set to be equal to the trading dates between the t-th and $(t+1)$-th period, for example, if adjustment frequency is annual, we set $K = 252$. Then we can obtain the new estimates of $\boldsymbol{\mu}^t$ and $\boldsymbol{\Sigma}$ by the simulated sequence of $\widehat{\boldsymbol{Y}}^t$. To combine the new simulated sequence of $\widehat{\boldsymbol{Y}}^t$ into the estimation of covariance $\boldsymbol{\Sigma}$, we estimate $\boldsymbol{\Sigma}$ by the simulated time series $\widehat{\boldsymbol{Y}}^t$ once again to get

$$\widehat{\boldsymbol{\Sigma}}^{simulate} = \mathrm{Cov}(\widehat{\boldsymbol{Y}}^t | \mathcal{F}_{t-1}).$$

Update $\boldsymbol{\Sigma}$ by combining $\boldsymbol{\Sigma}$ and $\widehat{\boldsymbol{\Sigma}}^{simulate}$ as

$$\widehat{\boldsymbol{\Sigma}}^{new} = \delta \widehat{\boldsymbol{\Sigma}}^s + (1 - \delta) \widehat{\boldsymbol{\Sigma}}^{simulate}, \tag{5.8}$$

where $\delta \in [0, 1]$. Clearly, $\widehat{\boldsymbol{\Sigma}}^{new}$ includes both historical and simulating information of $\boldsymbol{Y}^t$. We set $\delta = 0.7$ in the numerical tests.

Figure 2 gives comparative statics on RLPM, with a 10-year investment horizon, i.e., from Jan. 2nd, 1987 to Dec. 31st, 1996, and thus we take $T = 10$ years and annual adjustment. The change in expected annual returns $\overline{\mu}$, annual volatility σ and Sharpe ratios SR of RLPM for different parameter values are described in this figure.

Figure 2 indicates that the expected return, volatility and Sharpe ratio of the optimal portfolio of RLPM are decreasing with respect to the risk-aversion coefficient λ for a fixed target a. This is easily understood since the large risk aversion parameter λ corresponds to a relatively large degree of risk aversion or to a relatively conservative case which generally gives
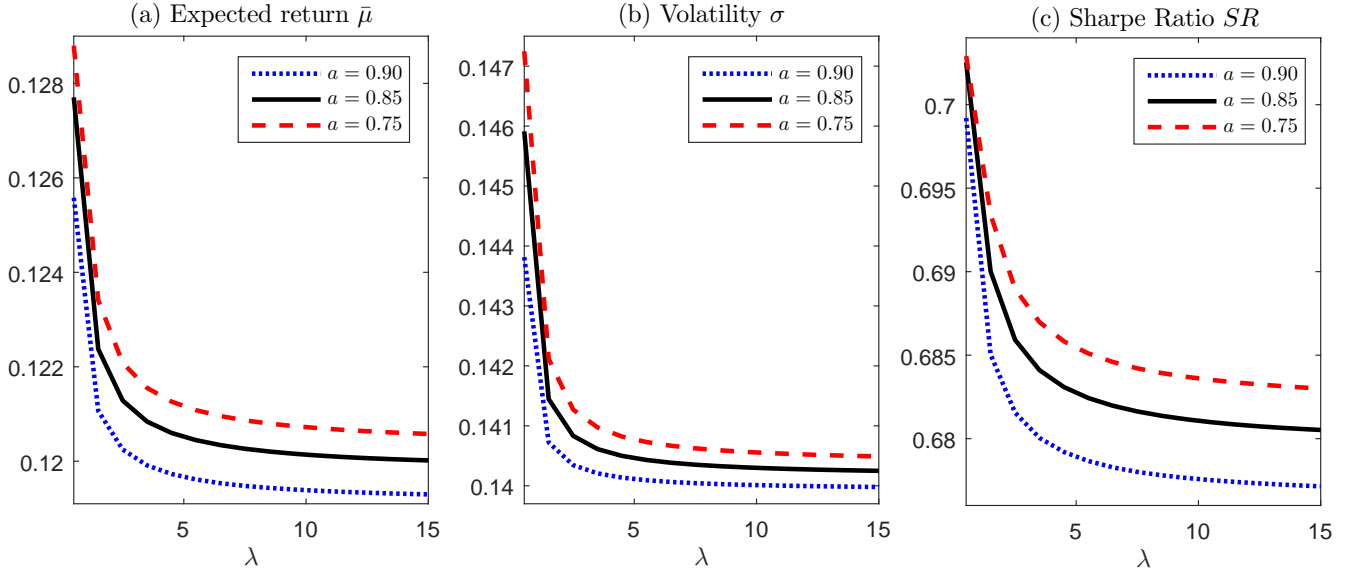
Figure 2: The comparative statics of the optimal average expected returns, average volatility and average Sharpe ratio with respect to risk averse coefficient λ under the three target values $a = 0.90, 0.85, 0.75$, where we take $T = 10$, and $\epsilon = 0.05$, $\epsilon_t = \epsilon/(T-1) = 0.0056$, $\Omega = \Omega_t = \sqrt{-2\ln\epsilon_t} = 3.22(t = 1, \cdots, 10)$.

a small expected return. The expected return will remain constant when λ reaches a relative large value. For example, when $\lambda > 15$, the expected return will keep nearly constant. This is because for larger λ, the lower partial moment is smaller. In particular,

$$\lambda \cdot \mathbb{E}\left[\left(a - w_0^T R_0^{T+1} - (\widetilde{\boldsymbol{R}}^{T+1})'\boldsymbol{w}^T\right)_+\right] \to 0$$

as $\lambda \to \infty$. Thus, the impact of LPM on the optimal portfolio will disappear if λ is large. This case happens for volatility and Sharpe ratio with the same reason.

In theory, the target a is exogenous and can be set freely by investors. Here, we set a by (4.19) which is an approximate reflection VaR with the given volatility and confidence level. For simplicity, we take $\sigma = 12.6\%$ in (4.19), which is calculated by the average volatility of five Industries assets from Jan. 2nd 1946 to Dec. 31st 1986, then $a = 0.75, 0.85, 0.90$ in Figure 2 are approximately corresponding to a single asset VaR with volatility 12.6% per year and confidence level 0.975, 0.90 and 0.80. Note that a is the investment target at the terminal period and $1 - a$ reflects the acceptable loss for the investor. Thus, the smaller the target a is, the larger the acceptable loss is, which in return indicates the investor would like to find a big bet with possible large loss or return. In other words, the investor who sets a relative small target a expects to get

24

a relative high expected return at the terminal period. This fact is showed by Figure 2, where the expected return, volatility and Sharpe ratio decrease as the targe a increases.

We aggregate the sum of weights of all risky assets at each investment period and then take the average sum for all $T = 10$ investment periods. Figure 3 gives the effect of risk averse coefficient λ on the total weight of risky assets. We find from Figure 3 that the total weight of risky assets is decreasing as risk averse coefficient λ and the target a. As stated above, a relative large λ means the investor is risk aversion more, and therefore he(she) will hold the risky assets fewer. Especially, when λ is close to zero, the risk seeking dominates the risk aversion and the total weight of risky assets is close one. In turn, when λ is large, the risk aversion dominates the risk seeking and the total weight risky assets is close to zero. When λ is fixed, Figure 3 also shows that the investor who sets a relative larger target a will hold fewer risky assets. Because a relative larger target a shows that the investor can only accept the less loss, which can lead to reduce the weights of risky assets.
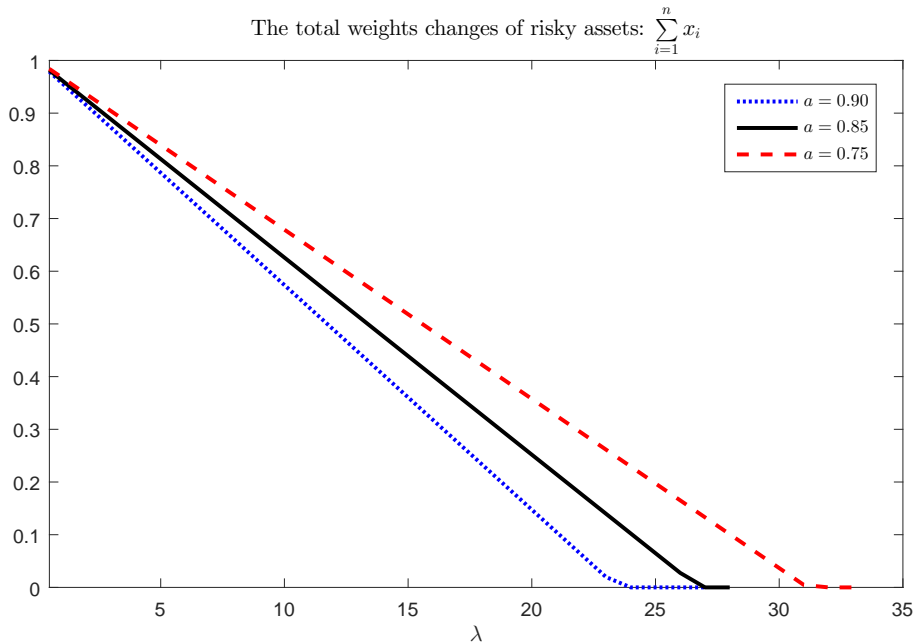

Figure 3: Comparative statics on the weights of risky assets with respect to risk averse coefficient λ under three target $a = 0.90, 0.85, 0.75$, where we take $T = 10$, and $\epsilon = 0.05$, $\epsilon_t = \epsilon/(T-1) = 0.0056$, $\Omega = \Omega_t = \sqrt{-2\ln \epsilon_t} = 3.22 (t = 1, \cdots, 10)$.

Figure 4 gives the comparisons of expected return, volatility and Sharpe ratio for WVaR, CVaR, WCVaR, equally-weighted strategy ($1/N$) and RLPM. One can find from Figure 4 that the expected return and volatility are all decreasing functions of λ for WVaR, CVaR, WCVaR and RLPM, but $1/N$ strategy is an exception. This reason is similar to Figure 2. Second, when λ is small (for example is less than 1.75), the expected return of RLPM is the greatest among

four models, but its expected return is decreasing sharply and lower than those of WVaR and CVaR, WCVaR, see Figure 4(a). This shows that the expected return of RLPM is more sensitive to the risk averse coefficient λ than other models .

However, third, RLPM can get the smallest volatility among all four models for any λ, which leads to RLPM has the best Sharpe ratio among four models. As we have seen, the WVaR model controls only the probability of extreme loss, and even though the CVaR and WCVaR model can limit the expectation of left-tail losses, it also only controls the expectation under the condition of extreme loss. Extreme loss in general is a low-probability event whose contribution to controlling the whole variance and improving the Sharpe ratio is possible small. RLPM can also control the lower parity loss, but the left-tail event is not extreme and a target a can be set by investors. Thus, the lower parity loss limited by RLPM has a greater probability than the extreme event of WVaR, WCVaR and CVaR models. This is the reason that the controlled lower parity loss in RLPM can more efficiently reduce the variance and improve the Sharpe ratio than the controlled extreme loss in the WVaR, WCVaR and CVaR models.

Fourth, the asymmetry uncertainty set is significantly important to reduce the volatility of portfolio relative to the symmetry ellipsoidal uncertainty set $\mathcal{U}_\Omega$. See Figure 5, in where, we set $\alpha = 20\%$ for WVaR, WCVaR and CVaR models and take $a = 0.90$. In other words, when the nearly same and not extreme tail is considered for four models, the proposed RLPM still can obtain the smallest volatility. There are two possible reasons. One is that, as stated in Introduction, the asymmetry attitude of investors on the gain and loss can form the asymmetry distribution feature. The asymmetry uncertainty set model is clearly more appropriate than the symmetry ellipsoidal uncertainty set model. The another is that, the asymmetry uncertainty set $\mathcal{F}_\Omega$, as stated in Section 3, is the generalization of the symmetry ellipsoidal uncertainty set $\mathcal{U}_\Omega$ and can lead to the relative more conservative case than the ellipsoidal uncertainty set $\mathcal{U}_\Omega$, which can lead to the smaller variance of portfolio. Figure 4 and Figure 5 show that the performance of RLPM model is robust for different target a and comparable to WVaR, WCVaR and CVaR models with the different loss level α.

We finally give some cumulative wealth comparisons on these 5Industry assets for models RLPM, WCVaR, CVaR, WVaR and equally-weighted strategy $(1/N)$ with different adjustment frequencies. Figure 6 reports these comparisons with three adjustment frequencies, quarterly, semiannual and annual adjustment with an investment horizon: from Jan. 2nd 1987 to Dec. 30th 2016. We fix the $T = 10$ to avoid the individual stochastic constraint too conservatively. Thus, in 30 years investment horizon case, we use M-SRP with different M for different adjustment frequencies, see Subsection 5.1 for MSRP. We fix $\lambda = 1.75$ in Figure 6 since the four models, CVaR, WCVaR, WVaR and RLPM have nearly equivalent expected return when $\lambda = 1.75$, see Figure 4(a).
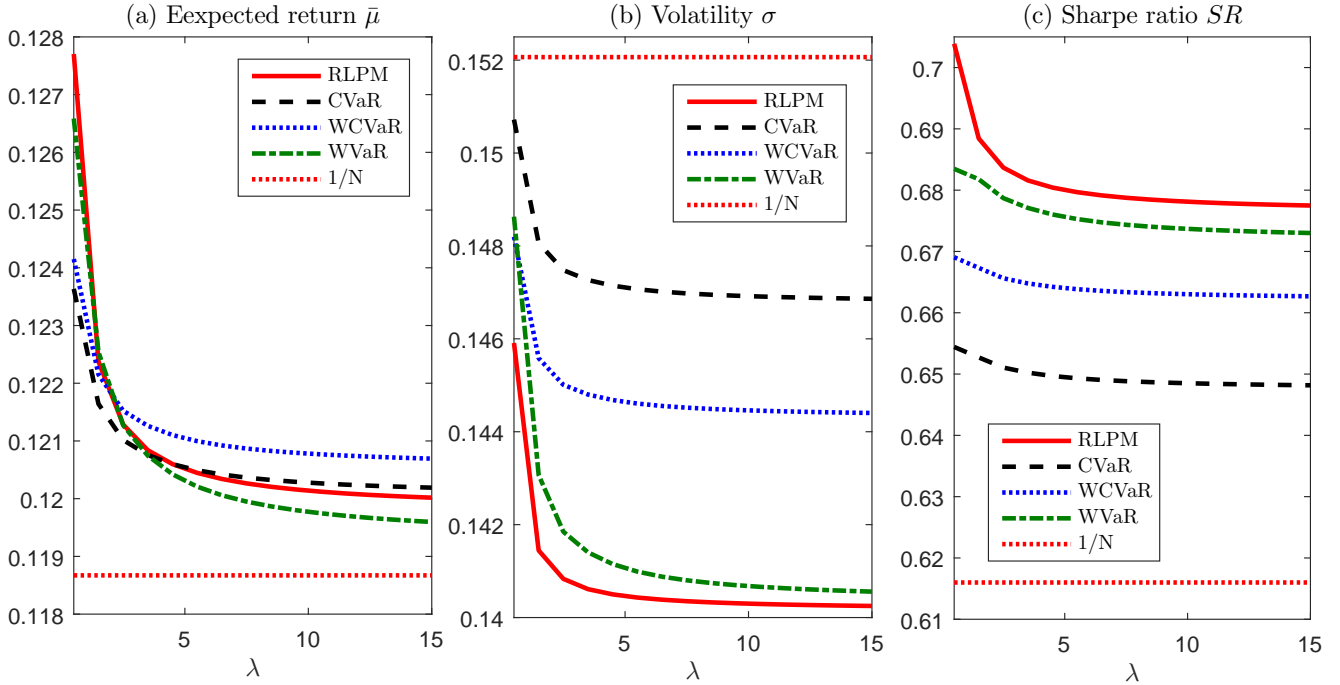
Figure 4: Average performances comparisons of RLPM, CVaR, WCVaR, WVaR and equally-weighted strategy $(1/N)$ for different risk aversion coefficient λ, where $a = 0.85, \epsilon = 0.05$, and $T = 10$ are fixed. $\epsilon_t = \epsilon/(T-1) = 0.0056$, $\Omega = \Omega_t = \sqrt{-2\ln\epsilon_t} = 3.22(t = 1, \cdots, 10)$. The loss level $\alpha = 5\%$ is considered in this examples for CVaR, WCVaR, WVaR. The investment horizon of portfolios performance: Jan. 2nd 1987 to Dec. 31st 1996.

Some interesting results are shown in Figure 6 for the long term investment horizon. In the trading dates between the 1000th and 2000th, the cumulative wealth of the RLPM model is weakly better than other models but after the 2000th trading date, the cumulative wealth of the RLPM model is clearly better than other models, especially, for quarterly adjustment. When the adjustment frequency is high, such as quarterly adjustment, the terminal wealth of the WVaR model is the lowest among the four models because the transaction cost is high for WVaR model[12]. WCVaR has the weaker performance than CVaR in this case. This is because that WCVaR has the weaker diversification than CVaR, which can lead to higher transection cost when the adjustment frequency is high. The performance of WCVaR becomes better than CVaR as the adjustment frequency decreases, see Figure 6 (b) and (c). The WCVaR model

---

[12]Transaction cost is relating to the turnover ratio which is calculated by

$$TR = \frac{1}{\Delta t}\left[\frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n}|\boldsymbol{x}_{i,t+1} - \boldsymbol{x}_{i,t}|\right]$$

for all models. See Table 1 for a comparison of average TR.
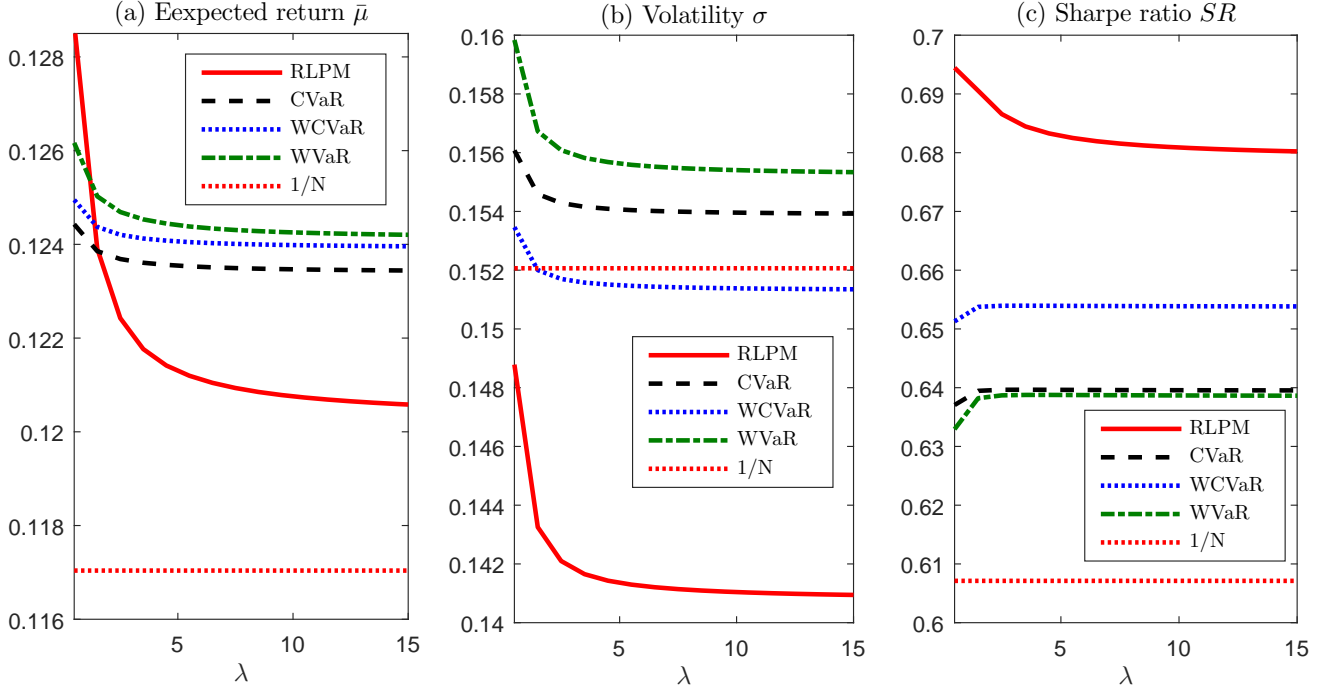
Figure 5: Average performances comparisons of RLPM, CVaR, VaR and equally-weighted strategy $(1/N)$ for different risk aversion coefficient λ, where $a = 0.90, \epsilon = 0.05$, and $T = 10$ are fixed. $\epsilon_t = \epsilon/(T-1) = 0.0056$, $\Omega = \Omega_t = \sqrt{-2\ln\epsilon_t} = 3.22(t = 1, \cdots, 10)$. The loss level $\alpha = 20\%$ is considered in this examples for CVaR, WCVaR, WVaR. The investment horizon of portfolios performance: Jan. 2nd 1987 to Dec. 31st 1996.

has the second best terminal wealth among the five models under the semiannual and annual adjustment for the long term horizon. The transaction cost is controlled endogenously by the RLPM model, which is a possible reason, allowing the RLPM model to yield the highest terminal wealth among the five models whatever the adjustment frequency is quarterly, semiannual or annual.

In order to show that the Sharpe ratio of RLPM is better than other models not only at the average level, but also at the statistical level, Table 1 reports a performance test of the Sharpe ratio using a significance test proposed by Jobson and Korkie (1981) and Memmel(2003) for the 5Industry assets. As we see in Figure 4 and Figure 5, four models, CVaR, WCVaR, WVaR and RLPM, have the nearly close expected returns when at $\lambda = 1.75$, but the terminal wealth of RLPM is the largest among the four models for different adjustment frequencies in Figure 6. Panel A in Table 1 shows further this feature of RLPM when $\lambda = 1.75$. In order to show that the property of RLPM is robust with respect to λ, we also consider the case of $\lambda = 6.50$ where RLPM has the smallest expected return among four models, CVaR, WCVaR, WVaR and RLPM. Panel B in Table 1 shows the statistical results of SR of RLPM with respect to other

three models. It is clearly found that the Sharpe ratio of RLPM relative to that of other models has the significant p-value at the 10% level for all cases under two λ. Most of regressions are significant at the 5% level, and some results are significant at the 1% level, see Panel A, WVaR column with Annually Adjustment.

[**Insert Table 1 about here**]

5.3   Further numerical comparisons.

In the rest of this section, we compare RLPM model with the recent robust growth-optimal portfolio (RGOP) model of Rujeerapaiboon, Kuhn and Wiesemann (2016) that can outperform equally weighted strategy. RGOP model is essentially the worst-case VaR portfolio model and can expressed as follows.

$$\begin{aligned}
\max_{\mathbf{w}} \mathrm{WVaR}_\alpha(\widetilde{\gamma}_T(\mathbf{w})) \quad &= \max_{\mathbf{w}} \min_{\mathbb{P} \in \mathcal{P}} \mathbb{P} - \mathrm{VaR}_\alpha(\widetilde{\gamma}_T(\mathbf{w})) \\
&= \max_{\mathbf{w}, \gamma \in \mathbb{R}} \left\{ \gamma : \ \mathbb{P}\left(\tfrac{1}{T} \sum_{t=1}^{T} \log(1 + \mathbf{w}'\widetilde{\mathbf{r}}_t) \geq \gamma \right) \geq 1 - \alpha, \ \ \forall \mathbb{P} \in \mathcal{P} \right\}
\end{aligned}$$

where the n-dimension vector $\mathbf{w}$ denotes the portfolio weight and is adjusted at the beginning of each period, such that $\mathbf{w}$ is the same for all periods $t = 1, \cdots, T$. The uncertainty set

$$\mathcal{P} = \left\{ \mathbb{P} \in \mathcal{P}_0^{nT} : \ \mathbb{E}_{\mathbb{P}}(\widetilde{\boldsymbol{r}}_t) = \boldsymbol{\mu}, \ \mathbb{E}_{\mathbb{P}}(\widetilde{\boldsymbol{r}}_t \widetilde{\boldsymbol{r}}_s') = \delta_{st}\boldsymbol{\Sigma} + \boldsymbol{\mu}\boldsymbol{\mu}', \quad \forall s, t : 1 \leq s \leq t \leq T \right\}$$

is a symmetry uncertainty set and similar to El Ghaoui et al. (2003). Here, δ_{st} is the coefficient of scaling the covariance matrix $\boldsymbol{\Sigma}$. The model can be solved by positive semidefinite programming (SDP). See Rujeerapaiboon, Kuhn and Wiesemann (2016) for more details.

We use the same four data sets, 10Ind, 12Ind, iShare and DJIA, as in Rujeerapaiboon, Kuhn and Wiesemann (2016), with all data from the French data library, Yahoo finance and the Wind database. The optimal portfolio of RLPM is adjusted annually for all data sets. Table 2 reports these comparisons.

[**Insert Table 2 about here**]

We observe from Table 2 that the RLPM model still attains better returns and Sharpe ratios than two RGOP models. Moreover, the standard deviations of RLPM is the lowest for all data . This is also not surprising, because RGOP is only a probability minimization problem

29

which is a zero-order moment optimization problem and is similar to VaR that does not consider the first-order moment. But, in RLPM, the first-order downside moment is considered as the minimization objective, which leads clearly to the lower standard deviation. Finally, the turnover ratios of RLPM dominates completely the RGOP model for all data sets. The reasons can be understood clearly because the transaction costs in our robust model are not exogenous but are integrated with the decision process, and are therefore a part of the optimal decision, while the transaction costs are completely exogenous in RGOP.

We close this section by discussing some characteristics of the RLPM model. First, unlike the multi-period scenarios generating model of Shen and Zhang (2008), where they only consider a two-period case due to the dilemma of scenarios. Our multi-period model can avoid the dilemma of the scenarios. Second, our model can be too conservative when the number of periods T is large. This is because a high probability guarantee for joint stochastic constraints requires a higher probability guarantee for each individual stochastic constraint. But we can basically overcome this problem using a multi-stage rolling procedure even though the investment horizon is long. Third, the exogenous target a can help to reduce the volatility of optimal solution due to avoiding the extreme threshold in CVaR and VaR models. Finally, we need to mention that our model still has many input parameters, such as, target a, risk aversion coefficient λ, $\boldsymbol{\mu}^t$ and $\boldsymbol{A}^t$ which is dependent on covariance matrix $\boldsymbol{\Sigma}$. Some parameters, for example, λ, a and $\boldsymbol{\Sigma}$, are sensitive for the performance of model. So how to mitigate the sensitivity of model on these parameters will be important problems to be further studied.

6   Conclusions

We consider a multi-period robust portfolio selection problem with an asymmetric uncertainty distribution. In this problem, the worst-case downside risk under an asymmetric uncertainty set is defined and its upper bound is explored. We establish the robust multi-period mean-LPM portfolio selection model with transaction costs under an asymmetric uncertainty set. We explore the second-order cone representation of the robust model by the first-order moment upper bound of the random variable. The theoretical results show that the optimal solution of RLPM can be obtained with a given probability for individual and joint stochastic constraints. We also analyze the effects of the target a and risk averse λ on the optimal performance using a comparative statics method. In order to achieve an efficient implementation, using VAR(1) process, we provide a multi-period simulation approach to estimate the parameters. Our numerical results and comparisons with real market data indicate that the RLPM model achieves better excess returns and Sharpe ratios, and has the lowest standard deviation among WVaR, CVaR WCVaR and equally-weighted strategy. Our multi-period model can deal

with the portfolio problem with the long investment horizon by M-SRP approach. Moreover, we also suggest that an appropriate rebalancing frequency should be chosen, depending on the investment horizon.

References

[1] Ang, A., Chen, J. and Xing, Y., (2006). Downside risk. Review of Financial Studies 19: 1191-239.

[2] Artzner, P., Delbaen, F., Eber, J. and Heath, D., (1999). Coherent measures of risk. Mathematical Finance, 9: 203-228.

[3] Bali, T., Cakici, N. and Whitelaw, R., (2014). Hybrid Tail Risk and Expected Stock Returns: When Does the Tail Wag the Dog? Reviews Asset Pricing Studies, 4 (2): 206-246.

[4] Bawa, V. , Lindenberg, E., (1977). Capital market equilibrium in a mean-lower partial moment framework. Journal of Financial Economics 5, 189-200.

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

[6] Ben-Tal, A., Margelit, T., Nemirovski, A.. Robust modeling of multi-stage portfolio problems. In: Frenk H, Roos K, Terlaky T, Zhang S, editors. High-performance optimization. Dordrecht: Kluwer Academic Publishers; 2000, 303-8.

[7] Bertsimas, D. and Pachamanova, D., (2008). Robust multiperiod portfolio management in the presence of transaction costs. Computers & Operations Research, 35, 3-17.

[8] Bodnar, T., Parolya, N. and Schmid, W., (2015). A Closed-Form Solution of the Multi-Period Portfolio Choice Problem for a Quadratic Utility Function. Annals of Operations Research 229(1): 121-158.

[9] Chen, L. He, S. and Zhang, S., (2011). Tight Bounds for Some Risk Measures, with Applications to Robust Portfolio Selection. Operations Research 59(4), 847-865.

[10] Chen, W., Sim, M., Sun, J. and Teo, C.-P., 2010. From CVaR to Uncertainty Set: Implications in Joint Chance-Constrained Optimization Oper. Res., 58(2), 470-485.

[11] Chen, X., M. Sim and P. Sun., (2007). A robust optimization perspective on stochastic programming. Operations Research, 55(6), 1058-1107.

[12] Cui, X., Li, D. and Li, X., (2017). Mean-Variance Policy for Discrete-Time Cone-Constrained Markets: Time Consistency in Efficiency and the Minimum-Variance Signed Supermartingale Measure. Mathematical Finance, 27(2): 471-504

[13] Dokuchaev, N., (2010). Optimality of myopic strategies for multi-stock discrete time market with management costs. European Journal of Operational Research, 200(2), 551-556.

[14] Ghaoui, L. El, Oks, M. and Oustry, F., (2003). Worst-case value-at-risk and robust portfolio optimization: a conic programming approach. Operations Research 51, 543-556.

[15] Gârleanu, N. and Pedersen, L., (2013). Dynamic trading with predictable returns and transaction costs. Journal of Finance, 68(6): 2309-2340.

[16] Goldfarb, D. and Iyengar, G., (2003). Robust Portfolio selection problems. Mathematics of Operations Research 28, 1-38.

[17] Gulpinar, N., Canakoglu, E. and Pachamanova, D., (2014). Robust Investment Decisions under Disruption in Petroleum Markets. Computers and Operations Research, 44(4): 75-91

[18] Gulpinar, N. and Pachamanova, D., (2013). A Robust Optimization Approach to Asset Liability Management under Time-Varying Investment Opportunities. Journal of Banking and Finance, 37(6): 2031-2041

[19] Gulpinar, N., Pachamanova, D. and Canakoglu, E., (2016). A Robust Asset-Liability Management Framework for Investment Products with Guarantees. OR Spectrum, 38(4): 1007-1041

[20] Huang, D., Zhu, S., Fabozzi, F. and Fukushima, M., (2008). Portfolio selection with uncertain exit time: A robust CVaR approach. Journal of Economic Dynamics and Control 32(2), 594-623.

[21] Huang, D., Zhu, S., Fabozzi, F. and Fukushima, M., (2010). Portfolio Selection under Distributional Uncertainty: A Relative Robust CVaR Approach. European Journal of Operational Research, 203(1), 185-194.

[22] Hull, John, (2015). *Options, futures and other derivatives.* 9th ed. Upper Saddle River, NJ: Prentice Hall,

[23] Jobson, J., and Korkie, B., (1981). Performance hypothesis testing with the Sharpe and Treynor measures. Journal of Finance 36, 4: 889-908.

[24] Kahneman, D., and Tversky, A., (1979). Prospect Theory: An Analysis of Decision under Risk, Econometrica, XVLII: 263-291

[25] Kelly, B. and Jiang, H., (2014).Tail risk and asset prices. Review of Financial Studies, 27(10): 2841-2871.

[26] Li, D. and Ng, W., (2000). Optimal Dynamic Portfolio Selection: Multiperiod Mean-Variance Formulation, Mathematical Finance, Vol. 10, No. 3, 387-406.

[27] Ling, A., Sun J. and Yang, X., (2014). Robust tracking error portfolio selection with worst-case downside risk measures. Journal of Economic Dynamics & Control 39, 178-207.

[28] Ling, A., Sun J., Xiu, N. and Yang, X., (2017). Robust two-stage stochastic linear optimization with risk aversion. European Journal of Operational Research 256, 215-229.

[29] Liu, J. and Chen, Z., (2018). Time consistent multi-period robust risk measures and portfolio selection models with regime-switching. 268 (1): 373-385.

[30] Mei, X., DeMiguel, V. and Nogales, F., (2016). Multiperiod portfolio optimization with multiple risky assets and general transaction costs. Journal of Banking & Finance 69: 108-120.

[31] Memmel, C., (2003). Performance hypothesis testing with the Sharpe ratio. Finance Letters, 1, 21-23.

[32] Natarajan, K., Pachamanova, D. and Sim, M., (2008). Incorporating Asymmetric Distributional Information in Robust Value-at-Risk Optimization, Management Science, 54(3), 573-585.

[33] Natarajan, D. Pachamanova and M. Sim., (2009). Constructing Risk Measures from Uncertainty Sets. Operations Research, 57(5): 1129-1141.

[34] Rujeerapaiboon, N., Kuhn, D. and Wiesemann, W., (2016). Robust Growth-Optimal Portfolios. Management Science, 62 (7): 2090-2109.

[35] Ruszczyński, A. and Shapiro, A., (2006). Optimization of convex risk functions. Math. Oper. Res., 31, 433-452.

[36] Rockafellar, R. and Uryasev, S., (2000). Optimization of conditional Value-at-Risk. Journal of Risk 2, 21-41.

[37] Shen, R. and Zhang, S., (2008). Robust Portfolio Selection Based on a Multi-stage Scenario Tree. European Journal of Operational Research 191, 864-887.

[38] Zhang, W., Liu,Y. and Xu, W., (2012). A possibilistic mean-semivariance-entropy model for multi-period portfolio selection with transaction costs. European Journal of Operational Research 222(2), 341-49.

[39] Zhu, S., Fan, M. and Li, D., (2016). Portfolio management with robustness in both prediction and decision: A mixture model based learning approach. Journal of Economic Dynamics and Control 48, 1-25

[40] Zhu, S., Fukushima, M., (2009). Worst-case conditional Value-at-Risk with application to robust portfolio management, Operations Research, 57(5), 1155-1168.

[41] Zhu, S., Li, D., and Wang, S., (2009). Robust Portfolio Selection under Downside Risk Measures. Quantitative Finance 7, 869-885.

[42] Zymler, S., Kuhn, D. and Rustem, B., (2013). Worst-Case Value-at-Risk of Non-Linear Portfolios. Managment Science, 59(1), 172-188.

[43] Zymler, S., Rustem, B. and Kuhn, D., (2011). Robust Portfolio Optimization with Derivative Insurance Guarantees. European Journal of Operational Research 210(2), 410-424.

Appendix: Variables, parameters and symbols

Table B.1: This table summarizes the symbols for the key variables used in the model and the parameter values. Bold lowercase, e.g. $\boldsymbol{a}, \boldsymbol{\mu}, \cdots$, denote a vector. $\widetilde{\xi}$ and $\widetilde{\boldsymbol{\xi}}$ denote a random variable and vector, and their realizations are denoted by ξ and $\boldsymbol{\xi}$, respectively. Bold uppercase letters, e.g. $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\Sigma}, \cdots$, will generally denote a matrix. The 'transpose' is denoted by '′'.

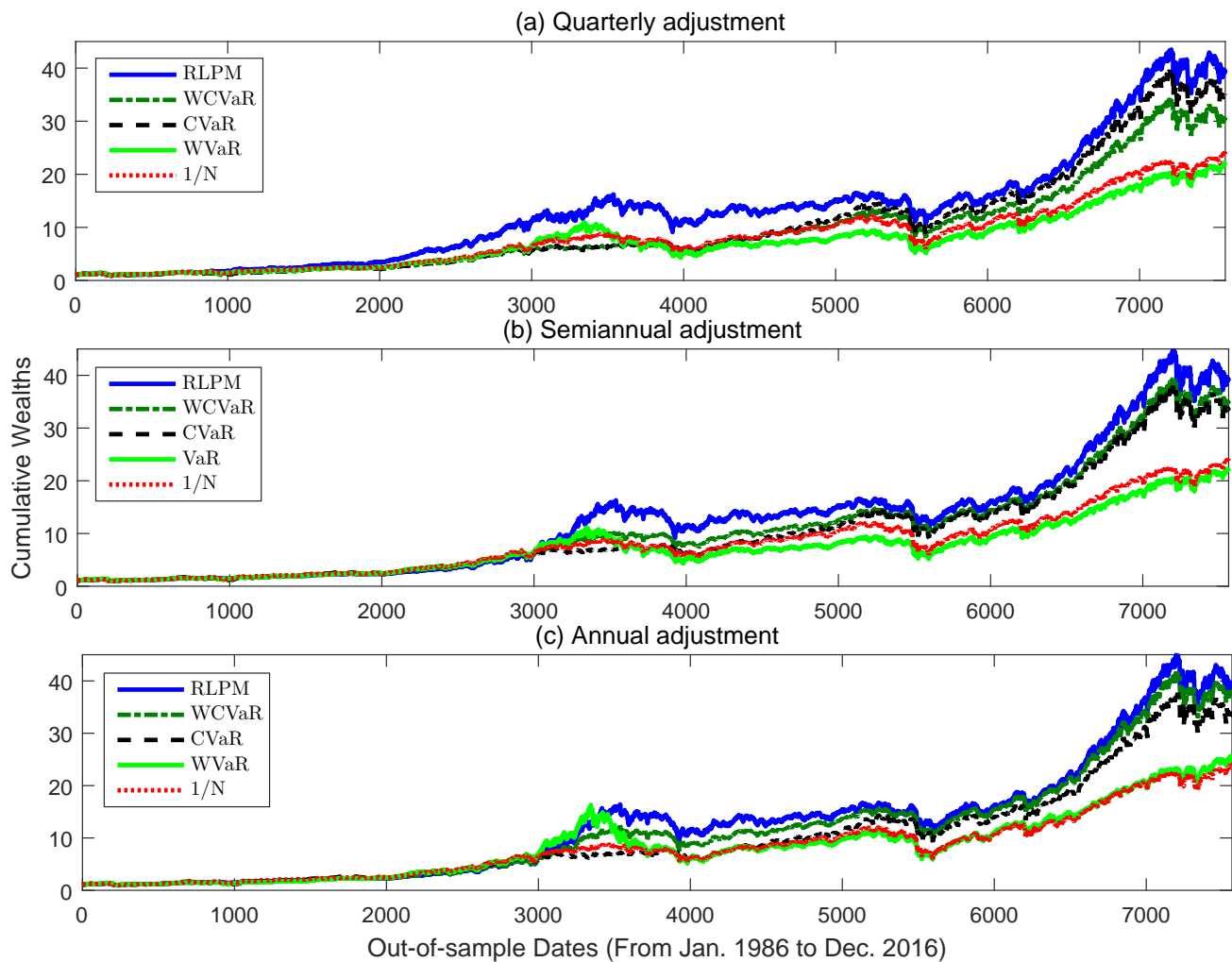| Variables	Symbols	
Downside risk measure	$\rho(\cdot)$	
Uncertainty shock at the tth period	$\widetilde{\boldsymbol{\xi}}^t,\ \widetilde{\xi}$	$t = 1, \cdots, T$
Realization of $\widetilde{\boldsymbol{\xi}}^t,\ \widetilde{\xi}$	$\boldsymbol{\xi}^t,\ \xi$	$t = 1, \cdots, T$
Returns vector of risky assets at time t	$\widetilde{\boldsymbol{r}}^t$	$t = 1, \cdots, T$
Cumulative returns vector of risky assets until to time t	$\widetilde{\boldsymbol{R}}^{t+1},$	$t = 1, \cdots, T$
Cumulative returns of riskless assets until to time t	$R_0^{t+1},$	$t = 1, \cdots, T$
Uncertainty set	$\mathcal{F}_\Omega$	
Portfolio vector with dollar amount at the tth period	$\boldsymbol{x}^t,$	$t = 1, \cdots, T$
Cash amount at the tth period	$x_0^t,$	$t = 1, \cdots, T$
Dollar amount of buying risky assets at the beginning of tth period	$\boldsymbol{\Delta}^{b,t},$	$t = 2, \cdots, T$
Dollar amount of selling risky assets at the beginning of tth period	$\boldsymbol{\Delta}^{s,t},$	$t = 2, \cdots, T$
Terminal wealth of portfolio	W_P^T	
Scaled decision variables	$\boldsymbol{w}^t,$	$t = 1, \cdots, T$
Scaled dollar amount of buying risky assets	$\boldsymbol{z}^{b,t},$	$t = 2, \cdots, T$
Scaled dollar amount of selling risky assets	$\boldsymbol{z}^{s,t},$	$t = 2, \cdots, T$
Other auxiliary decision variables	$\boldsymbol{d}^T, \tau, \tau^h, \boldsymbol{s}^t, \boldsymbol{v}^t, \boldsymbol{\zeta}^t,$	$t = 1, \cdots, T$
Parameters	**Symbols**	**Estimations (values)**
Expected cumulative returns vector until to time t	$\boldsymbol{\mu}^{t+1}$	$\widehat{\boldsymbol{\mu}}^{t+1},$ for all t
Factor loading matrix until to time t	$\boldsymbol{A}^{t+1}$	$\widehat{\boldsymbol{A}}^{t+1},$ for all t
Covariance matrix of VAR(1) process at time t	$\boldsymbol{\Sigma}^t$	$\widehat{\boldsymbol{\Sigma}},$ for all t
Coefficients of VAR(1) process	$\boldsymbol{\nu}, \Phi$	$\widehat{\boldsymbol{\nu}}, \widehat{\boldsymbol{\Phi}}$
Forward and backward deviations	p_i^t, q_i^t	$\widehat{p}_i, \widehat{q}_i,$ for all t
Return of risk-free asset at time t	$r_0^t = r_f$	1.5% for all t
Risk averse coefficient	λ	1.75, 2.15, 6.50
Investment target	a	0.75/0.85/0.90
Transaction cost	θ	0.2%
Updating coefficient of covariance matrix $\widehat{\boldsymbol{\Sigma}}^s$	δ	0.7
Size of uncertainty set	Ω, Ω_t	$\sqrt{-2\ln\epsilon_t}$
Confident level	ϵ, ϵ_t	$\epsilon = 0.05, \epsilon_t = \epsilon/(T-1)$
Bound of support set of $\widetilde{\boldsymbol{\xi}}^t$	$\boldsymbol{l}^t, \boldsymbol{u}^t$	$\boldsymbol{l} = \boldsymbol{u} = \boldsymbol{10}$ for all t
Confident level of VaR and CVaR	α	5%, 20%

34

Figure 6: The cumulative wealth comparisons of RLPM, WCVaR, CVaR, WVaR and equally-weighted strategy $(1/N)$ with different adjustment frequencies, where $a = 0.85$, $\lambda = 1.75$, $T = 10$ and $\epsilon = 0.05$, $\epsilon_t = \epsilon/(T-1) = 0.0056$, $\Omega = \Omega_t = \sqrt{-2\ln\epsilon_t} = 3.22(t = 1, \cdots, 10)$. $\alpha = 5\%$ for WCVaR, CVaR, WVaR. The investment horizon from Jan. 2nd 1987 to Dec. 30th 2016 is fixed for all models and M-SRP is used to ensure that $T = 10$ for different adjustment frequencies.

Table 1. Sharpe ratio (SR) of the RLPM test is reported by using a significance test proposed by Jobson and Korkie (1981) and Memmel (2003), where $\lambda=1.75$, target $a=0.85$, T=10 and $\epsilon=0.05$, $\epsilon_t = \epsilon/(T-1) = 0.0056$, $\Omega=\Omega_t=\sqrt{-2\ln \epsilon}=3.22$ (t = 1,···,10). $\alpha = 5\%$ for CVaR, WCVaR and WVaR. The investment horizons are 10 years (left column) and 20 years (right column). Turnover ratio (TR) is the average annual turnover ratio. The bracket is the p-value, and ***, **, * mean that p-values are statistically significant at the 1%, 5% and 10% levels.

Models	Semi-annual Adjustment (2-stage portfolio procedure)					Annually Adjustment (2-stage portfolio procedure)				
	$WVaR$	$WCVaR$	$CVaR$	1/N	RLPM	$WVaR$	$WCVaR$	$CVaR$	1/N	RLPM
Panel A: λ=1.75										
SR	0.782**	0.365**	0.621**	0.632**	0.691	0.645***	0.416**	0.577**	0.667**	0.694
	(0.015)	(0.032)	(0.025)	(0.013)		(0.007)	(0.034)	(0.041)	(0.031)	
TR	0.452	0.462	0.412	0.426	0.387	0.406	0.402	0.382	0.408	0.375
Panel B: λ=6.50										
SR	0.422**	0.237**	0.467**	0.528**	0.661	0.471**	0.257**	0.415*	0.544**	0.667
	(0.042)	(0.031)	(0.033)	(0.027)		(0.016)	(0.036)	(0.062)	(0.035)	
TR	0.442	0.438	0.415	0.426	0.390	0.413	0.412	0.383	0.408	0.363

Table 2. Out-of-Sample numerical comparisons with Robust growth-optimal portfolio (RGOP) of Rujeerapaiboon, Kuhn, Wiesemann(2016). Here, $\lambda=2.15$, target $a=0.85$, T=6 and $\epsilon=0.05$, $\epsilon_t = \epsilon/(T-1) = 0.01$, $\Omega=\Omega_t=\sqrt{-2\ln \epsilon_t}=2.54$ (t = 1,···,8), and 2-SRP is used. The first column specifies the data set used in the numerical comparisons from [34]. The results of RGOP (RGOP[+]) are from [48], in which they report the monthly performance, and we transfer their monthly performance into annual results. That is, the $\bar{\mu}$ in this table are obtained by multiplying 12 to the monthly mean value in [34], σ are obtained by multiplying $\sqrt{12}$ to the monthly volatility in [34]. The *SR* in this table are annual and calculated by $(\bar{\mu} - \mu_f)/\sigma$. The average annual *TR* for RGOP (RGOP[+]) is obtained by multiplying 12 to the monthly average *TR* in [34].

Data	Models	$\bar{\mu} - \mu_f$	σ	*SR*	*TR*
10Ind (01/2000-12/2012)	RLPM	0.0780	0.1246	0.6260	0.4743
	RGOP	0.0744	0.1247	0.5966	0.5244
	RGOP[+]	0.0768	0.1251	0.6139	0.5196
12Ind (01/2000-12/2012)	RLPM	0.0781	0.1235	0.6324	0.5036
	RGOP	0.0756	0.1244	0.6077	0.5340
	RGOP[+]	0.0780	0.1251	0.6235	0.5328
iShares (04/2006-07/2013)	RLPM	0.0410	0.1916	0.2140	0.4241
	RGOP	0.0396	0.1985	0.1995	0.4656
	RGOP[+]	0.0396	0.1985	0.1995	0.4656
DJIA (04/2000-07/2013)	RLPM	0.0688	0.1319	0.5216	0.6138
	RGOP	0.0588	0.1320	0.4455	0.8016
	RGOP[+]	0.0684	0.1320	0.5182	0.7812