*Quantitative Finance*, 2020
https://doi.org/10.1080/14697688.2020.1781237



# A cost-effective approach to portfolio construction with range-based risk measures

CHI SENG PUN * and LEI WANG

School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, Singapore

(*Received 10 February 2020; accepted 5 June 2020; published online 29 July 2020*)

In this paper, we introduce a new class of risk measures and the corresponding risk minimizing portfolio optimization problem. Instead of measuring the expected deviation of a daily return from a single target value, we propose to measure its deviation from a range of values centered on the single target value. By relaxing the definition of deviation, the proposed risk measure is robust to the variation of data input and thus the resulting risk-minimizing portfolio has a lower turnover rate and is resilient to outliers. To construct a practical portfolio, we propose to impose an $\ell_2$-norm constraint on the portfolio weights to stabilize the portfolio's out-of-sample performance. We show that for some cases of our proposed range-based risk measures, the corresponding portfolio optimization can be recast as a support vector regression problem. This allows us to tap into the machine learning literature on support vector regression and effectively solve the portfolio optimization problem even in high dimensions. Moreover, we present theoretical results on the robustness of our range-based risk minimizing portfolios. Simulation and empirical studies are conducted to examine the out-of-sample performance of the proposed portfolios.

*Keywords*: Portfolio optimization; Transaction costs; Risk measures; Statistical learning theory; $\ell_2$-regularized portfolios; Support vector regression; Robustness

*JEL Classification*: C31 – Multiple Variables - Quantile Regressions, C55 – EconometricModeling - Large Data Sets: Modeling and Analysis, D81 – Information, Knowledge, andUncertainty - Criteria for Decision-Making under Risk and Uncertainty, G11 – GeneralFinancial Markets - Portfolio Choice/Investment Decisions

## 1. Introduction

Markowitz (1952) shows that with relevant beliefs about future performances, investors should pick a portfolio from the efficient frontier. This groundbreaking work on mean-variance analysis has had tremendous influence on investment practices. However, two critical factors: estimation error and transaction costs, undermine the successful implementation of the optimal mean-variance portfolios.

Estimation error, which perverts the empirical mean and covariance matrices, leads to poor out-of-sample performance of the resulting portfolios as shown in Michaud (1989), Chopra and Ziemba (2013). To reduce the impact of estimation error, much contemporary research considers only the minimum variance portfolios, for example, Clarke *et al.* (2011), Kempf and Memmel (2006). Minimum variance portfolios are less prone to estimation error as the estimation of the expected returns is no longer required. Another attempt to reduce the estimation error is by imposing constraints on portfolio norms. Jagannathan and Ma (2003) show the equivalence between imposing non-negativity constraints on the portfolio-weight vector and shrinking the extreme elements in the covariance matrix. DeMiguel *et al.* (2009) introduce the constraint that the norm of portfolio weights should be below a threshold and provide a unified framework for portfolio optimization with norm constraints. Norm-constrained portfolios can stablize portfolio performance and generally outperform unconstrained portfolios out-of-sample; see Brodie *et al.* (2009). Fan *et al.* (2012) interpret an $\ell_1$-norm constraint on portfolio weights as a gross exposure constraint, giving a theoretical insight for the conclusions of Jagannathan and Ma (2003). Hence, though the constraint may exclude the theoretically optimal portfolio, it decreases the estimation error substantially and leads to decent out-of-sample performance.

Another shortcoming of traditional portfolios is that transaction cost is ignored by minimum-variance portfolios. Due to the sensitivity of the inverse covariance matrix to the data input, the estimated minimum-variance and most other model-based portfolios are unstable in the sense of high turnover rate, as evidenced in DeMiguel *et al.* (2009). Unstable portfolio

*Corresponding author. Email: cspun@ntu.edu.sg

weight vectors result in large transaction costs, which can easily erode away the profit obtained from such portfolios. Though the imposition of a norm constraint helps control the gross exposure and indirectly stabilizes the portfolio-weight vector, the norm constrained strategy still produces portfolios with high transaction cost. An analytical approach to portfolio selection with transaction costs was initiated by Davis and Norman (1990) and has attracted a lot of attention; see Muthuraman (2007), Mei and Nogales (2018), Pun and Ye (2019). However, the analytical approach is restricted to a small number of assets being invested and thus a more practical approach is desired in this era of big data. In this paper, we directly address the transaction cost problem from an empirical point of view. In addition, our approach does not suffer from a tractability issue and can be easily applied to high dimensional portfolio selection problems.

The two problems above (estimation error and transaction cost) are more pronounced when we select a large-scale portfolio, such as an institutional portfolio, which consists of hundreds or thousands of assets. Due to non-stationarity limits, financial analysts usually only consider the recent 1-year or 3-year data. The ratio between the number of available stocks and the number of observations is large, leading to a high-dimensional statistical problem and potential instability in portfolios; see Still and Kondor (2010), Chiu *et al.* (2017) for the big data challenges in portfolio construction. The existing approaches to high-dimensional portfolio selection are mainly regularization; see Brodie *et al.* (2009), Fan *et al.* (2012), Pun and Wong (2016), Chiu *et al.* (2017), Pun and Wong (2019). However, their resulting portfolios still have a high turnover rate.

We attack the two aforementioned problems by proposing a novel risk measure and adapting the new risk measure to the portfolio risk minimization problem. We made a simple yet important observation that the conventional deviation risk measure considers any deviation from the central tendency. Thus, when we implement the traditional risk minimizing portfolio for a series of time points with rolling windows of data, we have to constantly adjust the portfolio because it is quite impossible that the new data is the same as the remaining data. Hence, we propose to relax the definition of the deviation risk by measuring the risk only for those deviations from a target range instead of a single target. As long as the remaining data and the new data have the same characteristic of being within the target range, the resulting risk minimizing portfolio remains unchanged. As a result, the turnover rate is reduced. This initiative is analogous to the proposal of $\epsilon$-insensitive objectives in support vector machines (Cortes and Vapnik 1995) in the machine learning literature and the proposal of the set-valued risk measure in Jouini *et al.* (2004) in the financial mathematics literature. However, the distinct features of our proposed range-based risk minimizing portfolios are their robustness to data variation and their applicability to large-scale asset allocation subject to transaction costs.

Not only does the introduction of the range-based risk measures control transaction cost and estimation error, it also aligns with human economic intuition. Fund managers do not adjust their portfolio weights often until there is a big change in the return rates or the market condition. They rather accept a range of portfolio returns and keep the portfolio weights the same if the portfolio return falls within the acceptable range, as it is normal for the market to fluctuate in the short term. Such a trading behavior provides a natural financial interpretation for our proposed range-based risk minimizing portfolios.

The contribution of this paper is threefold. First, we introduce a new class of risk measures, which are robust to the observations. We show desirable properties for the new risk measures and convert some common deviation risk measures into our proposed robust risk measures while preserving their interpretation. Second, we provide a general framework for constructing cost-effective portfolios by minimizing range-based risk measures. We address the implementation issue by connecting the range-based risk minimization with support vector regressions. Moreover, we theoretically prove that our proposed portfolios possess the aforementioned desired robustness property; see our main results in Propositions 2 and 3 below. They pave a new path for studying the large-scale portfolio selection problem subject to transaction costs from an empirical perspective. Third, we conduct extensive simulation and empirical studies to compare a number of risk minimizing portfolios and examine our proposed range-based improvement scheme. Our results supplement the empirical observations in DeMiguel *et al.* (2009), which considered only global minimum variance (GMV) portfolios in the class of risk minimizing portfolios and concluded that the GMV portfolio is the best among the model-based portfolios. Both our simulation and empirical results suggest that our proposed range-based risk minimizing portfolios have lower turnover rate and typically higher Sharpe ratio when the transaction cost is taken into account.

The remainder of the paper is organized as follows. Section 2 provides the motivation and a general definition of the proposed range-based risk measures. Section 3 introduces risk-minimizing portfolio problems with range-based risk measures and relates the problems to support vector regression problems. Section 4 provides the theoretical results on the robustness of the proposed portfolios. Simulation and empirical results are presented in Sections 5 and 6 respectively. Section 7 concludes the paper and proposes a few future research directions.

Throughout this paper, we adopt the following notation. For a vector or a matrix $A$, the transpose of $A$ is denoted by $A'$. We denote by $I$ and $\mathbf{1}$ the identity matrix and the vector of ones with appropriate dimensions, respectively. For $a = (a_1, \ldots, a_p)' \in \mathbb{R}^p$, we denote by $\ell_1$-norm $\|a\|_1 = \sum_{j=1}^p |a_j|$, by $\ell_2$-norm $\|a\|_2 = \sqrt{\sum_{j=1}^p a_j^2}$, and by $\mathrm{diag}(a)$ a diagonal matrix with diagonal elements of $a$. We study the problem of our interest in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. $\mathcal{I}_E := \mathcal{I}_E(\omega)$ denotes the indicator function of the event $E \in \mathcal{F}$.

## 2. Range-based risk measures

In this section, we introduce a new type of range-based risk measure for portfolio diversification. The new measure extends the deviation risk measure in the way that the corresponding risk minimizing portfolio will yield less turnover rate at the minimal cost of under-diversification. We first provide motivation of our initiative and then introduce the new risk measure.

### 2.1. Motivation of using alternative risk measures

To illustrate our motivation, we introduce the estimation aspect of our empirical studies in advance, which follows (DeMiguel *et al.* 2009) and usual market practice. In order to construct a portfolio for the investment period $[t, t+1]$, we use the *latest* $M$ (daily or monthly) data from $t-M+1$ to $t$ to estimate the parameters required in the portfolio construction, where $t$ denotes the current time. Moving forward from $t$ to $t+1$, the estimation window rolls to include the newest data at $t+1$ and exclude the oldest data at $t-M+1$. Even though most ($M-2$ out of $M$) data are overlapped, the computed portfolio weights at times $t$ and $t+1$, denoted by vectors $\hat{w}_t$ and $\hat{w}_{t+1}$, respectively, could be very different due to the portfolio's sensitivity to data. As a result, *when the conventional risk measure is used, there will be high turnover rate or equivalently, the investor needs to suffer a significant transaction cost*. We illustrate the portfolio's (or the underlying risk measure's) sensitivity to data from the following empirical and mathematical points of view.

The extensive empirical studies in DeMiguel *et al.* (2009), Pun and Wong (2019) reveal that equally weighted (EW) portfolio, which maintain the equal weights on all the assets in the market at all times, yields the least turnover rate, compared to the model-based theoretically optimal portfolios. The results in DeMiguel *et al.* (2009) imply that if $\hat{w}_t = \hat{w}_{t+1}$, then the portfolio turnover is relatively low because in this case, the turnover is only due to the variation of the asset prices over the investment period. On the other hand, taking minimum variance portfolio as an example, as long as the data for estimating the covariance matrix are different (even for two consecutive rolling estimation windows or for a small perturbation), the resulting portfolios could be very different and thus the turnover rate is high. It is ascribed to the non-robustness of variance as a risk measure in risk minimization. To illustrate this ascription, it is convenient for us to first revisit the deviation risk measures (DRMs) as follows:

$$DRM(Z) := \mathbb{E}[D(Z - \mathcal{T})], \tag{1}$$

where $Z$ is the portfolio's return, $\mathcal{T} \in \mathbb{R}$ is a statistic of $Z$ interpreted as a deterministic target return, and $D : \mathbb{R} \mapsto \mathbb{R}^+$ is some nonnegative convex function measuring the deviation of $Z$ from its target $\mathcal{T}$ with $D(0) = 0$. We further assume that there exists $k \geq 1$ such that $D(\lambda x) = \lambda^k D(x)$ for any $\lambda, x \in \mathbb{R}$. Moreover, we impose the following assumptions on $\mathcal{T}$:

Assumption 2.1 The target $\mathcal{T} : \mathcal{L}^2 \mapsto \mathbb{R}$, where $\mathcal{L}^2$ is the $L^2$ space of random portfolio returns, satisfies the followings:

(i) Normalization: $\mathcal{T}(0) = 0$;
(ii) Positive homogeneity: $\mathcal{T}(\lambda Z) = \lambda \mathcal{T}(Z)$ for any $Z \in \mathcal{L}^2$ and $\lambda \in \mathbb{R}^+$;
(iii) Constant additivity: $\mathcal{T}(Z + c) = \mathcal{T}(Z) + c$ for $Z \in \mathcal{L}^2$ and any $c \in \mathbb{R}$.

Under Assumption 2.1, it is easy to show that $(DRM(Z))^{1/k}$ is normalized, positive, positively homogeneous, and shift-invariant ($DRM(Z + c) = DRM(Z)$ for any $c \in \mathbb{R}$). The target is typically chosen as $\mathbb{E}[Z]$, while in practice, we may adopt an endogenous form as follows:

$$\mathcal{T}_D := \text{argmin}_{a \in \mathbb{R}} \mathbb{E}[D(Z - a)] \tag{2}$$

that depends on the choice of the deviation function. Note that when $D(x) = x^2$, we have $\mathcal{T}_D = \mathbb{E}[Z]$ and (1) becomes the variance of $Z$. It is easy to see that (2) satisfies Assumption 2.1.

Heuristically speaking, the measure (1) considers any deviation of the resulting portfolio's return from its target as risk and thus amendment on portfolio weights ($\hat{w}_{t+1} - \hat{w}_t \neq 0$) becomes inevitable at every time of rebalancing. It motivates us to extend the DRMs in (1) to *range-based risk measures (RRMs)*, where we re-define the deviation as the distance between $Z$ and a target range instead of a single target. In other words, we expect that if the future $Z$ (at $t+1$) falls within the target range, we will have $\hat{w}_{t+1} = \hat{w}_t$ such that the portfolio turnover is minimal. Note that we only propose to alter the risk measure used in the portfolio construction procedure. For the evaluation of portfolios' performance, we will still use the common measures like variance, Sharpe ratio, and portfolio turnover.

Our initiative to introduce a new type of risk measure is analogous to introducing proportional transaction cost constraints into the risk minimization problem. However, dealing with proportional transaction cost constraint poses analytical and implementation challenges, so most related studies are limited to a small universe of the assets; see Davis and Norman (1990), Mei and Nogales (2018), Pun and Ye (2019). Moreover, the distinct feature of our framework is that the portfolio minimizing the proposed RRM will be robust to the variation of data input and be resilient to outliers. In next section, we will illustrate the tractability of our framework.

### 2.2. Generalization from deviation risk measures to range-based risk measures

Instead of setting a single target, we establish a novel range-based risk measure (RRM) that is based on an interval of target values: $(\mathcal{T} - \epsilon, \mathcal{T} + \epsilon)$, which is characterized by the target center $\mathcal{T}$ and the acceptance radius $\epsilon \geq 0$. The RRM considers the event that $Z$ falls within the target range contributes no risk. Therefore, the target range can be interpreted as an acceptance range of $Z$. This feature facilitates the robustness of the empirical RRM with respect to the input data of $Z$. Mathematically, the RRM takes a general form of

$$RRM(Z) := \mathbb{E}[D((|\Delta| - \epsilon)_+ \text{sgn}(\Delta))], \tag{3}$$

where $\Delta := Z - \mathcal{T}$, $x_+ = \max(x, 0)$ is the positive part function, $\text{sgn}(x) = \mathcal{I}_{\{x \geq 0\}} - \mathcal{I}_{\{x<0\}}$ is the sign function, and $D$ and $\mathcal{T}$ are as introduced in (1). It is easy to see that when $\epsilon = 0$, the RRM (3) is reduced to DRM (1). Generally speaking, $\epsilon$ is also a statistic of $Z$ as with $\mathcal{T}$. In RRM, we assume that the target center $\mathcal{T}$ satisfies Assumption 2.1 and we impose the following assumptions on the $\epsilon$:

Table 1. Deviation risk measures and the corresponding range-based risk measures. $Q_\eta(Z)$ is the $\eta$-quantile of $Z$.

| DRM in (1) | $D(x)$ | $\mathcal{T}_D$ in (2) | RRM in (3) |
|---|---|---|---|
| Var: $\mathbb{E}[\Delta^2]$ | $x^2$ | $\mathbb{E}[Z]$ | R-Var: $\mathbb{E}[(\lvert\Delta\rvert-\epsilon)_+^2]$ |
| MAD: $\mathbb{E}[\lvert\Delta\rvert]$ | $\lvert x\rvert$ | $Q_{0.5}(Z)$ | R-MAD: $\mathbb{E}[(\lvert\Delta\rvert-\epsilon)_+]$ |
| QL: $\mathbb{E}[\Delta(\eta-\mathcal{I}_{\{\Delta<0\}})]$ | $\lvert x\rvert\lvert\eta-\mathcal{I}_{\{x<0\}}\rvert$ | $Q_\eta(Z)$ | R-QL: $\mathbb{E}[(\lvert\Delta\rvert-\epsilon)_+\lvert\eta-\mathcal{I}_{\{\Delta<0\}}\rvert]$ |

Assumption 2.2 The acceptance radius $\epsilon:\mathcal{L}^2\mapsto\mathbb{R}^+$ satisfies the followings:

(i) Normalization: $\epsilon(c)=0$ for any $c\in\mathbb{R}$ and $\epsilon(0)=0$ in particular;
(ii) Positive homogeneity: $\epsilon(\lambda Z)=\lambda\epsilon(Z)$ for any $Z\in\mathcal{L}^2$ and $\lambda\in\mathbb{R}^+$;
(iii) Translation invariance: $\epsilon(Z+c)=\epsilon(Z)$ for any $Z\in\mathcal{L}^2$ and $c\in\mathbb{R}$.

Proposition 1 *Under Assumptions* 2.1 *and* 2.2*, we have that* $(RRM(Z))^{1/k}$*, where $k$ is the constant such that* $D(\lambda x)=\lambda^kD(x)$*, satisfies the followings:*

(i) *normalization:* $(RRM(0))^{1/k}=0$*;*
(ii) *Non-negativity:* $(RRM(Z))^{1/k}\geq 0$ *for all non-constant* $Z\in\mathcal{L}^2$ *and* $(RRM(Z))^{1/k}=0$ *for any constant* $Z\in\mathbb{R}$.
(iii) *positively homogeneous:* $(RRM(\lambda Z))^{1/k}=\lambda(RRM(Z))^{1/k}$ *for any* $Z\in\mathcal{L}^2$ *and* $\lambda\in\mathbb{R}^+$*;*
(iv) *and shift-invariant:* $(RRM(Z+c))^{1/k}=(RRM(Z))^{1/k}$ *for any* $Z\in\mathcal{L}^2$ *and* $c\in\mathbb{R}$.

*Proof* By the normalization of $\mathcal{T}$ and $\epsilon$, we have $RRM(0)=0$. The non-negativity of $RRM(Z)$ is guaranteed by the non-negativity of $D(\cdot)$. If $Z$ is a constant, then by Assumptions 2.1 and 2.2, we have $\mathcal{T}(Z)=Z$ and $\epsilon(Z)=0$ and thus $\Delta=0$ in (3). Hence $RRM(Z)=\mathbb{E}[D(0)]=0$. The first two properties are proved.

For any $Z\in\mathcal{L}^2$ and $\lambda\in\mathbb{R}^+$, we have

$$\begin{aligned}(RRM(\lambda Z))^{1/k}&=(\mathbb{E}[D(\lambda(|\Delta|-\epsilon)_+\mathrm{sgn}(\Delta))])^{1/k}\\&=\lambda(\mathbb{E}[D((|\Delta|-\epsilon)_+\mathrm{sgn}(\Delta))])^{1/k}\\&=\lambda(RRM(Z))^{1/k}.\end{aligned}$$

Hence, the third property is proved

For any $Z\in\mathcal{L}^2$ and $c\in\mathbb{R}$, we note that $\mathcal{T}(Z+c)=\mathcal{T}(Z)$ and $\epsilon(Z+c)=\epsilon(Z)$. Hence, $\Delta(Z+c)=\Delta(Z)$. Therefore, it is easy to see from (3) that $RRM(Z+c)=RRM(Z)$. Then the results follow. ■

The advantage of using (3) in risk minimization is its robustness which leads to lower portfolio turnover during portfolio rebalancing; see Propositions 2 and 3 below. In table 1, we provide the RRM counterparts of the DRMs that commonly used in portfolio construction, including variance (Var), mean absolute deviation (MAD), and $\eta$-quantile loss (QL) with $\eta\in(0,1)$; see Markowitz (1952), Konno and Yamazaki (1991), Bonaccolto *et al.* (2017). Hereafter, we use a prefix 'R-' to indicate the RRM generalization of the corresponding DRM, such as R-Var, R-MAD, and R-QL.

Analogous to DRM, the target center and $\epsilon$ could be user-defined, as long as they satisfy Assumptions 2.1 and 2.2. However, to facilitate the tractability of the proposed framework, we propose to use the following target center that extends (2):

$$\mathcal{T}_{D,\epsilon}=\mathrm{argmin}_{a\in\mathbb{R}}\ \mathbb{E}[D((|Z-a|-\epsilon)_+\mathrm{sgn}(Z-a))].\quad (4)$$

Note that $\mathcal{T}_{D,0}=\mathcal{T}_D$. It is easy to show that (4) satisfies Assumption 2.1 given that $\epsilon$ satisfies Assumption 2.2. The choice of $\epsilon$ will be detailed in the Section 3.3.3.

Remark 1 Theoretically, the target $\mathcal{T}$ should satisfy Assumption 2.1 to embrace the desirable properties of the RRM. In practice, the target $\mathcal{T}$ may be set as a fixed value. Although our theoretical results would not nest this case, our numerical scheme for range-based risk minimization problems can accommodate with a fixed target, which is illustrated in the next section.

## 3. Portfolio construction with range-based risk measures

In this section, we discuss how to construct portfolios that minimize the range-based risk measures proposed in Section 2. Following the framework of Fan *et al.* (2012), suppose that there are $p$ risky assets in the market, whose simple return vector for the period $[t,t+1]$ is denoted by $R_t=(R_{1t},\ldots,R_{pt})'\in\mathbb{R}^p$. We assume that the asset returns are correlated across the $p$ assets but they are independent in time $t$. Let $w_t=(w_{1t},\ldots,w_{pt})'\in\mathbb{R}^p$ be the portfolio weight vector with $w_t'\mathbf{1}=1$. Then the random portfolio return is given by $w_t'R_t$ for the period $[t,t+1]$.

The risk minimization problem is at the center of modern portfolio theory which was initiated by Markowitz (1952). It is widely adopted in portfolio tracking and asset allocation. In this section, we introduce the linkage between the risk minimization problem with the newly introduced range-based risk measure. We also explain how the problem can be efficiently solved by support vector regression. The problems presented in this section are indexed by $t$ since we are interested in the portfolio turnover between two days.

### 3.1. Connection with regression problem

With the specification of the center target $\mathcal{T}_{D,\epsilon}$ in (4), the range-based risk minimization problem can be recast as follows:

$$\begin{aligned}&\min_{w_t\in\{x\in\mathbb{R}^p|x'\mathbf{1}=1\},\ a_t\in\mathbb{R}}\ \mathbb{E}_t[D((|w_t'R_t-a_t|-\epsilon)_+\mathrm{sgn}(w_t'R_t-a_t))]\\&=\min_{\widetilde{w}_t\in\mathbb{R}^{p-1},\ a_t\in\mathbb{R}}\ \mathbb{E}_t[D((|Y_t-a_t\\&\quad-X_t'\widetilde{w}_t|-\epsilon)_+\mathrm{sgn}(Y_t-a_t-X_t'\widetilde{w}_t))],\end{aligned}\quad (5)$$

where $\widetilde{w}_t = (w_{1t}, \ldots, w_{p-1,t})$, the equality is due to the substitution of $w_{pt} = 1 - (\widetilde{w}_t)'\mathbf{1}$, $Y_t = R_{pt}$ and $X_t = (X_{1t}, \ldots, X_{p-1,t})'$ with $X_{jt} = R_{pt} - R_{jt}$ for $j = 1, \ldots, p-1$. $Y_t$ plays a role as risk-free asset as in many portfolio selection problems; see Pun (2018). The risk minimization problem in (5) is independent of the choice of $Y_t$ and finding the optimal weights $w_t$ is equivalent to finding the optimal $\widetilde{w}_t$ in (5). In practice, the expectation term in (5) is replaced by its empirical counterpart with the historical data of the random asset return vector.

However, from the practical point of view, the risk minimizing portfolios might suffer from accumulation of estimation errors. As shown in DeMiguel *et al.* (2009), the problem of estimation error accumulation can be mitigated by imposing a norm constraint. Such constraint imposition is equivalent to adding a regularization/penalty term of the portfolio weight vector as seen in Brodie *et al.* (2009), Fan *et al.* (2012). From the statistical perspective, the regularization approach can avoid overfitting and yield better our-of-sample performance. Moreover, with the regularization, the portfolio can be implemented even when the number of samples is less than the number of assets. Since using more recent data to estimate preserves stationarity more effectively; see Broadie (1993) but raises the concern of overfitting and high-dimensional statistical problem, it becomes popular to introduce regularization term in portfolio construction; see Pun and Wong (2016), Chiu *et al.* (2017), Pun and Wong (2019). For our application, we reformulate the problem as a regularized risk minimization problem:

$$\min_{w_t \in \{x \in \mathbb{R}^p | x'\mathbf{1}=1\},\ a_t \in \mathbb{R}} \mathbb{E}_t[D((|w_t'R_t - a_t| - \epsilon)_+ \mathrm{sgn}(w_t'R_t - a_t)] + \mathcal{R}_\lambda(w_t) \tag{6}$$

where $\mathcal{R}_\lambda(w_t)$ is a regularization term of $w_t$ and $\lambda \in \mathbb{R}^+$ controls the tradeoff between the efficiency and the model complexity. As discussed in Fan *et al.* (2012), the problem (6) is similar to a transformed regularized risk minimization problem based on (5):

$$\min_{\widetilde{w}_t \in \mathbb{R}^{p-1},\ a_t \in \mathbb{R}} \mathbb{E}_t[D((|Y_t - a_t - X_t'\widetilde{w}_t| - \epsilon)_+ \mathrm{sgn}(Y_t - a_t - X_t'\widetilde{w}_t))] + \widetilde{\mathcal{R}}_\lambda(\widetilde{w}_t), \tag{7}$$

where $\widetilde{\mathcal{R}}_\lambda$ is the same as $\mathcal{R}_\lambda$ except that the former's argument is $(p-1)$-dimensional vectors while the latter's argument is $p$-dimensional vectors. The aforementioned regularized problems are not equivalent as (7) depends on the choice of $Y_t$, while (6) does not. The optimal solution to (7) gives an suboptimal solution to (6). However, as suggested by the numerical and empirical evidences in Fan *et al.* (2012), when $Y_t$ is properly chosen, the solution to is nearly optimal. We will discuss the choices of $Y_t$ and $\epsilon$ in the subsequent sections. For the implementation of our range-based risk minimizing portfolios, we adopt the form (7). We denote the optimal weights by $\hat{w}_t \in \mathbb{R}^{p-1}$ that minimizes (7). The corresponding optimal $\hat{a}$ is interpreted as the central tendency of $\hat{w}_t'R_t$.

Two common choices of $\widetilde{\mathcal{R}}_\lambda$ are $\ell_1$-norm and $\ell_2$-norm, i.e. $\widetilde{\mathcal{R}}_\lambda(\widetilde{w}_t) = \lambda\|\widetilde{w}_t\|_1$ and $\widetilde{\mathcal{R}}_\lambda(\widetilde{w}_t) = \frac{\lambda}{2}\|\widetilde{w}_t\|_2^2$, respectively. In this paper, we consider the latter over the former for the following two reasons. First, Olivares-Nadal and DeMiguel (2018) empirically shows that data-driven portfolio with calibrated quadratic transaction cost which is equivalent to an $\ell_2$-regularized regression problem outperforms its counteraprt with proportional transaction cost which is equivalent to an $\ell_1$-regularized regression problem. Second, van den Doel *et al.* (2012) found that not only is $\ell_2$-norm-regularization computationally easier, it can occasionally outperform $\ell_1$-norm regularization when subjected to noisy data.

### 3.2. Choice of *Y*

If (5) is formulated for portfolio tracking and the $p$th asset with return $Y_t = R_{pt}$ is the portfolio to be tracked, then problem (5) can be interpreted as finding the optimal weights on available $p-1$ assets to minimize a ranged-based tracking error. In this case, choosing the $p$th asset as $Y_t$ in (7) makes perfect sense.

If (5) is formulated for allocation on an universe with $p-1$ assets (without loss of generality as $p$ can vary), then following the framework of Fan *et al.* (2012), $Y_t$ could be a fixed portfolio on $p-1$ assets (to be tracked). Denote by $w_y \in \mathbb{R}^{p-1}$ the portfolio weight vector of $Y_t$ with $w_y'\mathbf{1} = 1$. Once we obtain the optimal weights from solving (7), we have the allocation vector on $p-1$ assets given by $\hat{w}_t + (1 - \hat{w}_t'\mathbf{1})w_y$. We would like to remark here that when the regularization parameter $\lambda$ is very large and dominates the risk measure, (7) will suggest $\hat{w}_t \equiv 0$, which implies that the best allocation is $w_y$.

In Fan *et al.* (2012), the authors suggested the followings for the fixed portfolio $Y$: a tradable index, an electronically traded fund (ETF), or a non-short-sale efficient portfolio, and finally adopted the last one. However, we propose to use the EW portfolio as $Y$, i.e. $w_y = 1/(p-1)\mathbf{1}$. We also experimented using the non-short-sale efficient portfolio but it slightly underperforms than using the EW portfolio, which may be attributed to the estimation error incurred in the efficient portfolio. Moreover, EW portfolio has been empirically proven to yield higher Sharpe ratio than many portfolios, including the value- and price-weighted portfolios; see DeMiguel *et al.* (2009), Plyakha *et al.* (2012). As stated in Section 2.1, the EW yields the least turnover rate that is consistent with the objective of our range-based risk minimizing portfolios. Based on the reasons above, we advocate using the EW portfolio as $Y$.

### 3.3. Relation with support vector regression

With some specification of the RRM or the deviation function $D$, we can link the risk minimization problem (7) to a machine learning technique: support vector regression (SVR); see Drucker *et al.* (1997), Smola and Schlkopf (2004) for more details on SVR. Suppose that we collected the historical data, $\{R_s\}_{s=t-M}^{t-1}$, up to the current time $t$, where we assume the intervals of the historical return periods are the same as $[t, t+1]$. Then the empirical version of (7) is given by

$$\min_{\widetilde{w}_t \in \mathbb{R}^{p-1},\ a_t \in \mathbb{R}} \frac{1}{M} \sum_{s=t-M}^{t-1} [D((|Y_s - a_t - X_s'\widetilde{w}_t| - \epsilon)_+ \mathrm{sgn}(Y_s - a_t - X_s'\widetilde{w}_t))] + \frac{\lambda}{2}\|\widetilde{w}_t\|_2^2, \tag{8}$$

where $Y_s = R_{ps}$ and $X_s = (X_{1s}, \ldots, X_{p-1,s})'$ with $X_{js} = R_{ps} - R_{js}$ for $j = 1, \ldots, p-1$ and $s = t-M, \ldots, t-1$.

In this paper, we consider the range-based risk minimizing portfolios with the RRMs listed in table 1, namely minimum R-Var, R-MAD, and R-QL portfolios.

**3.3.1. Range-based mean absolute deviation (R-MAD) and range-based variance (R-Var).** The MAD and Var admit an unified form of the deviation function $D(x) = |x|^m$, where $m = 1$ refers to the MAD and $m = 2$ refers to the Var. Hence, it is convenient for us to discuss them together here. With such a specification of $D$, (8) becomes

$$\min_{\widetilde{w}_t \in \mathbb{R}^{p-1},\, a_t \in \mathbb{R}} \frac{1}{M} \sum_{s=t-M}^{t-1} (|Y_s - a_t - X_s'\widetilde{w}_t| - \epsilon)_+^m + \frac{\lambda}{2}\widetilde{w}_t'\widetilde{w}_t,$$

which can be rewritten as

$$\min_{\widetilde{w}_t \in \mathbb{R}^{p-1},\, a_t \in \mathbb{R},\, \{\xi_s, \xi_s^*\}_{s=t-M}^{t-1}} \frac{1}{2}\widetilde{w}_t'\widetilde{w}_t + C \sum_{s=t-M}^{t-1} \times [\xi_s^m + (\xi_s^*)^m] \quad \text{subject to } (\mathcal{C}) \tag{9}$$

where $C = 1/(M\lambda)$ is interpreted as a cost parameter to be tuned and the constraint set $(\mathcal{C})$ is defined as

$(\mathcal{C})$ For $s \in \{t-M, \ldots, t-1\}$,

$$\begin{aligned} &Y_s - a_t - X_s'\widetilde{w}_t - \epsilon \le \xi_s, \\ &Y_s - a_t - X_s'\widetilde{w}_t + \epsilon \ge -\xi_s^*, \quad \xi_s, \xi_s^* \ge 0. \end{aligned} \tag{10}$$

Problem (9) is a standard L1-loss SVR problem for $m = 1$ and a standard L2-loss SVR problem for $m = 2$, which can be solved even for a large scale ($M$ is large) with an efficient algorithm developed in Ho and Lin (2012). In our implementation, we solve for (9) with R package 'LiblineaR'.

Remark 2 When we use a fixed target, i.e. $a_t \equiv \mathcal{T}$, we can simply substitute $Y_s$ by $Y_s - \mathcal{T}$ and still leverage R package 'LiblineaR' with the setting of no intercept to solve for (9).

In our framework, one can also consider incorporating linear portfolio constraints, which nest no-shorting and leverage constraints. Note that $\tilde{w}$ is a linear transformation of $w$. Hence, linear constraints on $w$ can be converted to another set of linear constraints on $\tilde{w}$. The corresponding risk minimization problem can be formulated as a support vector regression with linear constraints, for which one can refer to Klopfenstein and Vaiter (2019) for the sequential minimal optimization algorithm to solve. However, the addition of the constraints is not the focus of this paper and cannot bring additional insights to this paper. We focus on the risk minimization problem (8).

**3.3.2. Range-based quantile loss (R-QL).** For the quantile loss, we take $D(x) = |x||\eta - \mathcal{I}_{\{x<0\}}|$, where $\eta \in (0,1)$ and typically takes value of 0.25 or 0.75. Note that when $\eta = 0.5$, the QL is equivalent to MAD. Then, (8) can be rewritten as

$$\min_{\widetilde{w}_t \in \mathbb{R}^{p-1},\, a_t \in \mathbb{R}} \frac{1}{M} \sum_{s=t-M}^{t-1} \left[(|Y_s - a_t - X_s'\widetilde{w}_t| - \epsilon)_+ |\eta - \mathcal{I}(Y_s - a_t - X_s'\widetilde{w}_t < 0)|\right] + \frac{\lambda}{2}\widetilde{w}_t'\widetilde{w}_t,$$

The problem above can be recast as a support vector quantile regression form:

$$\min_{\widetilde{w}_t \in \mathbb{R}^{p-1},\, a_t \in \mathbb{R},\, \{\xi_s, \xi_s^*\}_{s=t-M}^{t-1}} \frac{1}{2}\widetilde{w}_t'\widetilde{w}_t + C \sum_{s=t-M}^{t-1} \times [\eta\xi_s + (1-\eta)\xi_s^*] \quad \text{subject to } (\mathcal{C}), \tag{11}$$

where $C$ is defined and interpreted as in (9) and $(\mathcal{C})$ is defined in (10). Following Kecman (2005), Seok *et al.* (2010), the dual problem of (11) is

$$\begin{aligned} \min_{\{\alpha_{t,s}, \alpha_{t,s}^*\}_{s=t-M}^{t-1}} &\frac{1}{2} \sum_{s_1, s_2 = t-M}^{t-1} (\alpha_{t,s_1} - \alpha_{t,s_1}^*)(\alpha_{t,s_2} - \alpha_{t,s_2}^*) X_{s_1}' X_{s_2} \\ &+ \sum_{s=t-M}^{t-1} (\epsilon - Y_s)\alpha_{t,s} + \sum_{s=t-M}^{t-1} (Y_s + \epsilon)\alpha_{t,s}^* \end{aligned} \tag{12}$$

subject to constraints: for $s \in \{t-M, \ldots, t-1\}$,

$$\sum_{s=t-M}^{t-1} (\alpha_{t,s} - \alpha_{t,s}^*) = 0,$$

$$0 \le \alpha_{t,s} \le \eta C, \quad 0 \le \alpha_{t,s}^* \le (1-\eta)C.$$

The primal-dual relationship indicates that the optimal solution to the dual problem (12), denoted by $\{(\hat{\alpha}_{t,s}^{QL}, \hat{\alpha}_{t,s}^{*QL})\}_{s=t-M}^{t-1}$, and the optimal solution to the primal problem (11), $\hat{w}_t^{QL}$, satisfy

$$\hat{w}_t^{QL} = \sum_{s=t-M}^{t-1} (\hat{\alpha}_{t,s}^{QL} - \hat{\alpha}_{t,s}^{*QL}) X_s. \tag{13}$$

Moreover, the optimal target $\hat{a}_t^{QL}$ is obtained via Kuhn-Tucker conditions (Kuhn and Tucker 1951) as

$$\hat{a}_t^{QL} = \frac{1}{N_t + N_t^*}\left[\sum_{s \in I_t}\left(Y_s - X_s'\hat{w}_t^{QL} - \epsilon\right) + \sum_{s \in I_t^*}\left(Y_s - X_s'\hat{w}_t^{QL} + \epsilon\right)\right], \tag{14}$$

where $N_t$ is the size of the set $I_t = \{s \in \{t-M, \ldots, t-1\} | \; 0 < \hat{\alpha}_{t,s}^{QL} < \eta C\}$ and $N_t^*$ is the size of the set $I_t^* = \{s \in \{t-M, \ldots, t-1\} | \; 0 < \hat{\alpha}_{t,s}^{*QL} < (1-\eta)C\}$. Problem (12) can be solved as a constrained quadratic programming problem. In our implementation, we solve for (11) with R package 'quadprog'.

Remark 3 When we use a fixed target, i.e. $a_t \equiv \mathcal{T}$, we can derive a similar dual problem as in (12), which is detailed in Appendix. The optimal solution is still given in the form of (13).

**3.3.3. Choice of $\epsilon$.** After we have converted the risk minimization problem into support vector regression problem, we can borrow the ideas from the machine learning engineers to determine the choice of $\epsilon$ for the non-expert users. A practical selection suggested in Cherkassky and Ma (2004) is that the optimal choice of $\epsilon$ should be proportional to the input noise level. Specifically, based on the empirical tuning, the $\epsilon$ should take the form of

$$\epsilon = 3\sigma\sqrt{\frac{\log M}{M}}, \tag{15}$$

where $\sigma^2$ is the second moment of $Y_t - a_t - X_t'\widetilde{w}_t$, which is equivalent to the variance of $w_t'R_t$. Since standard deviation is a deviation risk measure, it is easy to see that the $\epsilon$ in (15) satisfies Assumption 2.2. Note that the $\sigma$ requires prior knowledge of the true model (the oracle $w_t$), which is not available. In Cherkassky and Ma (2004), the authors used the $K$-nearest-neighbours regression method to approximate the noise level $\sigma$. In our implementation, we apply this method to the first training dataset and obtain an estimated $\epsilon$ value, which is used throughout the test datasets.

The value of $\epsilon$ is critical to the computation of portfolio as it determines which daily return data are support vectors. A change of $\epsilon$ value can lead to an enormous change in the portfolio-weight vector, potentially incurring huge transaction cost. In practice, many investment companies make major change in their portfolios on a monthly basis while modifying their portfolio-weights slightly on a daily basis. In order to apply our method to real financial investment, investors can update $\epsilon$ value monthly and use a constant $\epsilon$ throughout a month. Such practice ensures that enough updated information is captured in portfolio selection and the transaction cost is controlled at an acceptable level.

## 4. Theoretical results: robustness of minimum RRM portfolios

One of the major advantages of the range-based risk minimizing portfolios over the conventional risk minimizing portfolios is their robustness to the data input, which generally leads to a lower turnover rate and transaction cost. In this section, we echo our motivation of introducing range-based risk measures by showing that $\hat{w}_\tau = \hat{w}_{\tau+1}$ for some scenarios and some time $\tau$, where $\hat{w}_t$ is the optimal weights of $\widetilde{w}_t$ that minimize the problem (8) indexed by $t$. We also denote by $\hat{a}_t$ the optimal target of $a_t$ that minimizes the time $t$-problem (8). Since $\tau$ could be arbitrary, the analysis below is general for the whole test period.

By investigating the change from time $\tau$-problem of (8) to time $(\tau+1)$-problem of (8), we recognize the only difference is that the latter replaces the oldest data $(X_{\tau-M}, Y_{\tau-M})$ by the newest data $(X_\tau, Y_\tau)$ in the former. Hence, our investigation is also highly related to the robustness of data input. In order to investigate the change in the corresponding resulting portfolios, we examine the similarity between $(X_{\tau-M}, Y_{\tau-M})$ and $(X_\tau, Y_\tau)$. We introduce the following concept of acceptance set.

Definition 1 *Given a pair $(\hat{w}_\tau, \hat{a}_\tau)$, we define the acceptance set as*

$$\mathcal{A}^\epsilon(\hat{w}_\tau, \hat{a}_\tau) = \left\{(X, Y) \in \mathbb{R}^{p-1} \times \mathbb{R} \,\middle|\, |Y - \hat{a}_\tau - X'\hat{w}_\tau| < \epsilon\right\}.$$

Here, $Y = R_p$ and $X = (X_1, \ldots, X_{p-1})'$ with $X_j = R_p - R_j$ for $j = 1, \ldots, p-1$ are the transformed data of return vector $R = (R_1, \ldots, R_p)$ as in (5). If we set $\hat{w}_\tau^* = (\hat{w}_\tau', 1 - \mathbf{1}'\hat{w}_\tau)'$, then we have $|Y - \hat{a}_\tau - X'\hat{w}_\tau| = |R'\hat{w}_\tau^* - \hat{a}_\tau|$. Hence, the acceptance set represents the outcomes, characterized by $R'\hat{w}_\tau^*$, that are within an acceptable interval $(\hat{a}_\tau - \epsilon, \hat{a}_\tau + \epsilon)$ with the target $\hat{a}_\tau$ determined at time $\tau$. It is noteworthy that when $\epsilon = 0$, $\mathcal{A}^0(\hat{w}_\tau, \hat{a}_\tau) \equiv \emptyset$ for any pair of $(\hat{w}_\tau, \hat{a}_\tau)$.

### 4.1. Robustness of minimum R-MAD and R-QL portfolios

In this subsection, we show the robustness of minimum R-QL portfolios that are given by (11). Note that the minimum R-MAD problem is a special case of (11) with $\eta = 0.5$ and a scaled cost parameter: $C_{QL} = 2C_{MAD}$. Hence, our results in this section hold true for the minimum R-MAD and R-QL portfolios with any $\eta \in (0, 1)$. To distinguish a different class of optimal portfolios from next subsection, we denote the minimum R-QL portfolio weight by $\hat{w}^{QL}$ together with the optimal target $\hat{a}^{QL}$.

Proposition 2 *Given that $(\hat{w}_\tau^{QL}, \hat{a}_\tau^{QL})$ minimizes the R-QL problem (11) with $t = \tau$. If $(X_{\tau-M}, Y_{\tau-M}) \in \mathcal{A}^\epsilon(\hat{w}_\tau^{QL}, \hat{a}_\tau^{QL})$ and $(X_\tau, Y_\tau) \in \mathcal{A}^\epsilon(\hat{w}_\tau^{QL}, \hat{a}_\tau^{QL})$, then we have $\hat{w}_{\tau+1}^{QL} = \hat{w}_\tau^{QL}$ and $\hat{a}_{\tau+1}^{QL} = \hat{a}_\tau^{QL}$.*

*Proof* The primal-dual formulation of the minimum R-QL problem in (11) sheds light on the robustness of the methodology. The idea of the proof is to first write down the optimality conditions for time $(\tau+1)$-problem of (11) and then verify that $(\hat{w}_\tau^{QL}, \hat{a}_\tau^{QL})$ is the minimum point. To this end, we also need to derive the optimality conditions that are satisfied by $(\hat{w}_\tau^{QL}, \hat{a}_\tau^{QL})$. Hence, we first present the dual problem and the Karush–Kuhn–Tucker (KKT) conditions for the problem (11) indexed by $t$ for $t \in \{\tau, \tau+1\}$.

Based on (13) and (14), in order to show that $\hat{w}_{\tau+1}^{QL} = \hat{w}_\tau^{QL}$ and $\hat{a}_{\tau+1}^{QL} = \hat{a}_\tau^{QL}$, we ought to show that

$$\{(\hat{\alpha}_{\tau,s}^{QL}, \hat{\alpha}_{\tau,s}^{*QL})\}_{s=\tau-M}^{\tau-1} = \{(\hat{\alpha}_{\tau+1,s}^{QL}, \hat{\alpha}_{\tau+1,s}^{*QL})\}_{s=\tau-M+1}^{\tau}. \tag{16}$$

Note that at the optimal solutions, (13) and (14), the following KKT complementarity conditions must be fulfilled: for $s \in \{t-M, \ldots, t-1\}$,

$$\begin{aligned} \hat{\alpha}_{t,s}^{QL}[Y_s - \hat{a}_t^{QL} - X_s'\hat{w}_t^{QL} - \epsilon - \hat{\xi}_s] &= 0, \\ \hat{\alpha}_{t,s}^{*QL}[Y_s - \hat{a}_t^{QL} - X_s'\hat{w}_t^{QL} + \epsilon + \hat{\xi}_s^*] &= 0, \end{aligned} \tag{17}$$

where $\hat{\xi} \geq 0$ and $\hat{\xi}^* \geq 0$ are the optimal $\xi$ and $\xi^*$, respectively, for the problem (11).

Since $(X_{\tau-M}, Y_{\tau-M}) \in \mathcal{A}^\epsilon(\hat{w}_\tau^{QL}, \hat{a}_\tau^{QL})$, (17) implies that $\hat{\alpha}_{\tau,\tau-M}^{QL} = \hat{\alpha}_{\tau,\tau-M}^{*QL} = 0$. Hence, by (13), we have $\hat{w}_\tau^{QL} = \sum_{s=\tau-M+1}^{\tau-1}(\hat{\alpha}_{\tau,s}^{QL} - \hat{\alpha}_{\tau,s}^{*QL})X_s$. In order to show that $\hat{w}_\tau^{QL}$ is the solution to $(\tau+1)$-problem (11), it remains to show that

$$(\hat{\alpha}^{QL}_{\tau+1,s}, \hat{\alpha}^{*QL}_{\tau+1,s}) = (\hat{\alpha}^{QL}_{\tau,s}, \hat{\alpha}^{*QL}_{\tau,s}) \text{ for } s = \tau - M + 1, \ldots, \tau - 1 \tag{18}$$

and $\hat{\alpha}^{QL}_{\tau+1,\tau} = \hat{\alpha}^{*QL}_{\tau+1,\tau} = 0$, because in this case, by (13),

$$\begin{aligned}\hat{w}^{QL}_{\tau+1} &= \sum_{s=\tau-M+1}^{\tau} (\hat{\alpha}^{QL}_{\tau+1,s} - \hat{\alpha}^{*QL}_{\tau+1,s})X_s \\ &= \sum_{s=\tau-M+1}^{\tau-1} (\hat{\alpha}^{QL}_{\tau,s} - \hat{\alpha}^{*QL}_{\tau,s})X_s = \hat{w}^{QL}_{\tau}.\end{aligned}$$

By the optimality of $\hat{w}^{QL}_{\tau}$ and $\{\hat{\alpha}^{QL}_{\tau,s}, \hat{\alpha}^{*QL}_{\tau,s}\}_{s=\tau-M+1}^{\tau-1}$, the associated KKT conditions for $(\tau+1)$-primal and dual problems, which are the same as that of $\tau$-problems, are satisfied with (18). Note that $(X_\tau, Y_\tau) \in \mathcal{A}^\epsilon(\hat{w}^{QL}_\tau, \hat{a}^{QL}_\tau)$. Similarly from (17) with $t = \tau+1$ and $s = \tau$, we have $\hat{\alpha}^{QL}_{\tau+1,\tau} = \hat{\alpha}^{*QL}_{\tau+1,\tau} = 0$. Then (16) is proved and the results follow. ■

Proposition 2 implies that if both exiting data $(X_{\tau-M}, Y_{\tau-M})$ and entering data $(X_\tau, Y_\tau)$ are in the acceptable range, then the optimal weights are unchanged. The robustness of minimum R-MAD and R-QL portfolios help them yield lower portfolio turnover than minimum MAD and QL portfolios. However, if the exiting data and entering data are not in the acceptable range, there will be a need to rebalance the optimal weights. Financially, we should consider that the distribution of the return data has undergone a significant change. In contrast, conventional risk measures set $\epsilon = 0$ and their corresponding portfolios' acceptance sets are empty. As a result, conventional risk measures consider the distribution of financial data to be ever-changing no matter how small the variation in the training data is. In this sense, RRM aligns with people's financial intuition better than DRM.

### *4.2. Robustness of minimum R-Var portfolios*

In this subsection, we show the robustness of minimum R-Var portfolios that are given by (9) with $m = 2$. We denote the minimum R-Var portfolio weight by $\hat{w}^{Var}$ and the optimal target by $\hat{a}^{Var}$. We obtain the similar results as with Proposition 2.

PROPOSITION 3 *Given that $(\hat{w}^{Var}_\tau, \hat{a}^{Var}_\tau)$ minimizes the R-Var problem (9) with $m = 2$ and $t = \tau$. If $(X_{\tau-M}, Y_{\tau-M}) \in \mathcal{A}^\epsilon(\hat{w}^{Var}_\tau, \hat{a}^{Var}_\tau)$ and $(X_\tau, Y_\tau) \in \mathcal{A}^\epsilon(\hat{w}^{Var}_\tau, \hat{a}^{Var}_\tau)$, then we have $\hat{w}^{Var}_{\tau+1} = \hat{w}^{Var}_\tau$ and $\hat{a}^{Var}_{\tau+1} = \hat{a}^{Var}_\tau$.*

*Proof* The proof is similar to that of Proposition 2. They are only different in the forms of the dual problem and the KKT conditions.

It is well-known (Vapnik 2000, Ho and Lin 2012) that the dual problem of (9) with $m = 2$ is

$$\begin{aligned}\min_{\{\alpha_{t,s},\alpha^*_{t,s}\}_{s=t-M}^{t-1}} \; & \frac{1}{2}\sum_{s_1,s_2=t-M}^{t-1} (\alpha_{t,s_1} - \alpha^*_{t,s_1})(\alpha_{t,s_2} - \alpha^*_{t,s_2})X'_{s_1}X_{s_2} \\ & + \sum_{s=t-M}^{t-1} (\epsilon - Y_s)\alpha_{t,s} \\ & + \sum_{s=t-M}^{t-1} (Y_s + \epsilon)\alpha^*_{t,s} + \frac{1}{4C}\sum_{s=t-M}^{t-1} (\alpha^2_{t,s} + \alpha^{*2}_{t,s})\end{aligned} \tag{19}$$

subject to constraints: for $s \in \{t-M, \ldots, t-1\}$,

$$\sum_{s=t-M}^{t-1} (\alpha_{t,s} - \alpha^*_{t,s}) = 0, \quad \alpha_{t,s} \geq 0, \quad \alpha^*_{t,s} \geq 0.$$

It is noteworthy that (19) can be solved alone. Moreover, the optimal solution to the dual problem (19), denoted by $\{(\hat{\alpha}^{Var}_{t,s}, \hat{\alpha}^{*Var}_{t,s})\}_{s=t-M}^{t-1}$, and the optimal solution to the primal problem (9) with $m = 2$, $\hat{w}^{Var}_t$, satisfy

$$\hat{w}^{Var}_t = \sum_{s=t-M}^{t-1} (\hat{\alpha}^{Var}_{t,s} - \hat{\alpha}^{*Var}_{t,s})X_s$$

and the optimal target $\hat{a}^{Var}_t$ is given by

$$\begin{aligned}\hat{a}^{Var}_t = & \frac{1}{N_t + N^*_t} \\ & \times \left[\sum_{s\in I_t}\left(Y_s - X'_s\hat{w}^{Var}_t - \epsilon\right) + \sum_{s\in I^*_t}\left(Y_s - X'_s\hat{w}^{Var}_t + \epsilon\right)\right],\end{aligned} \tag{20}$$

where with slightly abuse of notations, $N_t$ is the size of the set $I_t = \{s \in \{t-M, \ldots, t-1\} | \hat{\alpha}^{Var}_{t,s} > 0\}$ and $N^*_t$ is the size of the set $I^*_t = \{s \in \{t-M, \ldots, t-1\} | \hat{\alpha}^{*Var}_{t,s} > 0\}$. The KKT complementarity conditions of (9) also imply that the optimal solutions, $\{(\hat{\alpha}^{Var}_{t,s}, \hat{\alpha}^{*Var}_{t,s})\}_{s=t-M}^{t-1}$ and $(\hat{w}^{Var}_t, \hat{a}^{Var}_t)$ satisfy that for $s \in \{t-M, \ldots, t-1\}$,

$$\begin{aligned}\hat{\alpha}^{Var}_{t,s}[Y_s - \hat{a}^{Var}_t - X'_s\hat{w}^{Var}_t - \epsilon - \hat{\xi}_s] = 0, \\ \hat{\alpha}^{*Var}_{t,s}[Y_s - \hat{a}^{Var}_t - X'_s\hat{w}^{Var}_t + \epsilon + \hat{\xi}^*_s] = 0.\end{aligned}$$

By the same arguments as in the proof of Proposition 2, we can show that

$$\hat{\alpha}^{Var}_{\tau,\tau-M} = \hat{\alpha}^{*Var}_{\tau,\tau-M} = \hat{\alpha}^{Var}_{\tau+1,\tau} = \hat{\alpha}^{*Var}_{\tau+1,\tau} = 0.$$

Thus, the results follow. ■

The minimum Var portfolio is an interesting case as it was widely recognized as one of the most stable model-based portfolios in practice and in the literature on empirical finance (DeMiguel *et al.* 2009). The regularized minimum variance portfolios are recently studied by Fan *et al.* (2012), Pun and Wong (2019) (with $\ell_1$-norm) and Olivares-Nadal and DeMiguel (2018) (with $\ell_1$- and $\ell_2$-norms of change in portfolio-weight vector). Our proposed minimum R-Var portfolio improves over the existing portfolios by introducing data robustness, evidenced by Proposition 3. In our simulation and empirical studies, we particularly compare between minimum R-Var portfolio and $\ell_2$-regularized minimum Var portfolio.

## 5. Simulation study

In this section, we numerically examine the performance of 13 portfolios: $\ell_2$-regularized minimum R-MAD (Min R-MAD), $\ell_2$-regularized minimum R-Var (Min R-Var), $\ell_1$-regularized minimum Var (LASSO Min Var), $\ell_2$-regularized minimum Var (Ridge Min Var), $\ell_2$-regularized minimum R-QL (Min R-QL $\eta$) with $\eta = .25$ and .75, $\ell_2$-regularized minimum QL (Min QL $\eta$) with $\eta = .25$ and .75, equally weighted (EW), no-short sale (NS), global minimum variance (GMV) with sample covariance matrix, GMV with shrinkage covariance matrix estimate (Shrink), and oracle GMV (Oracle) portfolios. Except for the last five portfolios, other portfolios will have tuning parameters, where LASSO Min Var and Ridge Min Var have tuning parameter $\lambda$ with the interpretation as in (7) and the rest has tuning parameter $C = 1/(M\lambda)$.

### *5.1. Simulation setup and evaluation methodology*

We closely follow the data simulation process for a simulated Fama–French three-factor model, described in Section 4.1 of Fan *et al.* (2012) with parameters calibrated in Fan *et al.* (2008). The data simulation process is run $N_{sim} = 100$ times for two cases of different number of assets $p$ (100 or 400), where for the regression-based portfolios, we add an EW portfolio of the $p$ assets as the last asset. The EW portfolio is treated as $Y$ in the regression where the excess return of all the $p$ assets with respect to $Y$ are $X$. For each simulation, we simulate one-year daily returns of length $M = 252$ (days) for initial portfolio construction/training and another subsequent $N_{te} = 5$ days for testing. Then, we totally generated 100 $((M + N_{te}) \times p)$-daily return matrices. The first case with $p = 100$ is non-degenerate ($p < M$) whereas the second case with $p = 400$ is degenerate ($p > M$), which resembles a modern portfolio selection problem.

We use the latest $M$ days of training data to construct the portfolio weight vector. By assuming that the investors hold and rebalance this portfolio for the subsequent $N_{te}$ test days, we record the portfolio returns and the portfolio turnover on these $N_{te}$ days. Such an out-of-sample evaluation procedure is repeated for $N_{sim}$ simulations to obtain the estimates of the statistics of the out-of-sample portfolio returns. The metrics of the portfolio performance are listed in the table 2.

In addition to the performance metrics, we also report the mean and standard deviation (SD) of the endogenous daily targets induced in each portfolio. They are found with formulas (20) for Var-related portfolios and (14) for QL-related (including MAD) portfolios.

### *5.2. Simulation results*

The simulation results are summarized in tables 3 and 4. Overall, our proposed range-based risk minimizing portfolios exhibit lower turnover rate and comparable Sharpe ratio as against traditional risk minimizing portfolios at similar gross exposure level. The simulation results confirm our theoretical results on the cost-effectiveness of the range-based risk minimization. In what follows, we interpret the results from different aspects.

One of the most important observations is that range-based risk minimizing portfolios successfully bring down the daily turnover rate as compared to its conventional counterparts. In some cases listed in table 3 and 4, despite a higher gross exposure, the Min R-Var portfolio demonstrates a lower daily turnover rate than Ridge Min Var portfolio. To see this, we further compare turnover rates of the two portfolios over a range of gross exposure, which are reported in figure 1. In both cases ($p = 100$ and $p = 400$), the turnover rate of Min R-Var is entirely below that of Ridge Min Var. Moreover, for both .25 and .75 quantile, Min R-QL portfolios have lower gross exposure and turnover than the Min QL portfolios with the same tuning parameter $C$. The introduction of range-based risk concept helps reduce gross exposure and transaction cost of Min QL portfolios.

Another key feature of the range-based risk minimizing portfolios is their stability in both the degenerate and non-degenerate cases. Their turnover rate remain relatively stable when the number of stocks increases from 100 to 400. In contrast, a traditional portfolio optimization technique, GMV, has its daily turnover climb up by six times, whereas the Shrink portfolio's turnover doubles from 15.16% to 30.97%.

The range-based risk minimizing portfolios not only stabilize the weight-vector for $p$ stocks, but also improves the Sharpe ratio in some cases. Specifically, from tables 3 and 4, we can see that Min R-Var occasionally have higher Sharpe ratio than the Ridge Min Var at similar level in gross exposure, while Min R-QL performs similarly as with Min QL. These observations verify that the use of range-based risk measure in portfolio selection can bring down the turnover rate while the efficiency or the Sharpe ratio can remain at a similar or even improve to a better level.

Table 2. The metrics of the portfolio performance. To illustrate the formulas of the statistics, we denote by $w^{(k,j)} \in \mathbb{R}^p$ the estimated portfolio weights with any method for the $k$th simulation and $j$th test day's investment. $R^{(j)} \in \mathbb{R}^p$ is the simple asset return vector for the test period $[j-1, j]$ for $j = 1, \ldots, N_{te}$. $w^{(k,j-)} \in \mathbb{R}^p$ is the portfolio weights right before rebalancing on day $j-1$, given by $w_i^{(k,j-)} = \frac{w_i^{(k,j-1)}(1+R_i^{(j-1)})}{\sum_{i=1}^p w_i^{(k,j-1)}(1+R_i^{(j-1)})}$ for $i = 1, \ldots, p$. $\hat{\sigma}(\{Z\})$ is the sample standard deviation based on the dataset $\{Z\}$.

| Annual return | Annual volatility | Sharpe ratio |
|---|---|---|
| $252 \sum_{k=1}^{N_{sim}} \sum_{j=1}^{N_{te}} w^{(k,j)\prime} R^{(j)} / (N_{sim} N_{te})$ | $\sqrt{252}\hat{\sigma}(\{w^{(k,j)\prime} R^{(j)}\}_{k=1,\ldots,N_{sim};j=1,\ldots,N_{te}})$ | $\frac{\text{Annual return}}{\text{Annual volatility}}$ |
| **Gross exposure** | **Daily turnover rate** | |
| $\sum_{k=1}^{N_{sim}} \sum_{j=1}^{N_{te}} \|\|w^{(k,j)}\|\|_1 / (N_{sim} N_{te})$ | $\sum_{k=1}^{N_{sim}} \sum_{j=2}^{N_{te}} \|\|w^{(k,j)} - w^{(k,j-)}\|\|_1 / (N_{sim}(N_{te} - 1))$ | |

Table 3. Results of simulation study with $p = 100$. Note that the annual returns, volatilities, daily turnovers, and the statistics of daily target are reported in percentage.

| Portfolio | Tuning parameter | Gross exposure | Annual return (%) | Annual volatility (%) | Annual Sharpe ratio | Daily turnover (%) | Daily Target Mean (%) | Daily Target SD (%) |
|---|---|---|---|---|---|---|---|---|
| Min R-MAD | $C = 10^{-7}$ | 1 | 0.86 | 2.11 | 0.41 | 0.32 | 0.00 | 0.00 |
| Min R-MAD | $C = 10^{-6}$ | 1.18 | 0.74 | 1.75 | 0.42 | 1.64 | 0.00 | 0.00 |
| Min R-MAD | $C = 10^{-5}$ | 2.35 | 0.68 | 1.15 | 0.59 | 4.58 | 0.00 | 0.00 |
| Min R-Var | $C = 10^{-5}$ | 1 | 0.84 | 2.14 | 0.39 | 0.27 | 0.00 | 0.00 |
| Min R-Var | $C = 10^{-4}$ | 1.01 | 0.69 | 1.96 | 0.35 | 0.74 | 0.00 | 0.00 |
| Min R-Var | $C = 10^{-3}$ | 1.63 | 0.48 | 1.46 | 0.33 | 2.36 | 0.00 | 0.00 |
| LASSO Min Var | $\lambda = 0.002$ | 1.08 | 0.76 | 2.07 | 0.36 | 0.94 | 0.01 | 0.02 |
| LASSO Min Var | $\lambda = 0.001$ | 1.63 | 0.70 | 1.57 | 0.44 | 3.26 | 0.01 | 0.02 |
| LASSO Min Var | $\lambda = 0.0005$ | 2.09 | 0.24 | 1.25 | 0.19 | 5.82 | 0.01 | 0.02 |
| Ridge Min Var | $\lambda = 0.2$ | 1.00 | 0.75 | 1.97 | 0.38 | 0.73 | 0.01 | 0.02 |
| Ridge Min Var | $\lambda = 0.05$ | 1.36 | 0.63 | 1.60 | 0.39 | 2.05 | 0.01 | 0.02 |
| Ridge Min Var | $\lambda = 0.01$ | 2.42 | 0.67 | 1.06 | 0.64 | 5.00 | 0.01 | 0.02 |
| Min R-QL .25 | $C = 5 * 10^{-2}$ | 1.00 | 0.78 | 2.04 | 0.38 | 0.73 | − 0.29 | 0.08 |
| Min R-QL .25 | $C = 2 * 10^{-1}$ | 1.20 | 0.53 | 1.72 | 0.31 | 2.22 | − 0.78 | 0.24 |
| Min R-QL .25 | $C = 5 * 10^{-1}$ | 1.81 | 0.32 | 1.36 | 0.23 | 3.94 | − 3.00 | 0.76 |
| Min QL .25 | $C = 5 * 10^{-2}$ | 1.00 | 0.80 | 2.03 | 0.40 | 0.88 | − 0.25 | 0.04 |
| Min QL .25 | $C = 2 * 10^{-1}$ | 1.27 | 0.41 | 1.66 | 0.25 | 2.84 | − 0.78 | 0.19 |
| Min QL .25 | $C = 5 * 10^{-1}$ | 2.05 | − 0.21 | 1.26 | − 0.17 | 6.06 | − 3.61 | 0.80 |
| Min R-QL .75 | $C = 5 * 10^{-2}$ | 1.00 | 0.84 | 2.04 | 0.41 | 0.59 | 2.51 | 0.07 |
| Min R-QL .75 | $C = 2 * 10^{-1}$ | 1.21 | 0.71 | 1.72 | 0.41 | 1.89 | − 2.12 | 0.23 |
| Min R-QL .75 | $C = 5 * 10^{-1}$ | 1.82 | 0.80 | 1.36 | 0.59 | 3.55 | − 2.69 | 0.80 |
| Min QL .75 | $C = 5 * 10^{-2}$ | 1.00 | 0.85 | 2.02 | 0.42 | 0.74 | 2.00 | 0.03 |
| Min QL .75 | $C = 2 * 10^{-1}$ | 1.29 | 0.81 | 1.66 | 0.49 | 2.66 | − 0.34 | 0.20 |
| Min QL .75 | $C = 5 * 10^{-1}$ | 2.07 | 0.67 | 1.27 | 0.53 | 5.71 | − 3.20 | 0.78 |
| EW | | 1 | 0.87 | 2.17 | 0.38 | 0.26 | | |
| NS | | 1 | 1.06 | 1.44 | 0.74 | 3.48 | | |
| GMV | | 4.31 | 0.52 | 0.99 | 0.52 | 26.26 | | |
| Shrink | | 3.81 | 0.63 | 0.92 | 0.70 | 15.16 | | |
| Oracle | | 3.29 | 1.71 | 0.81 | 2.12 | 1.07 | | |

From the estimated daily targets, we can see that Min R-MAD and Min R-VaR portfolios implicitly adopt almost zero-return targets. The Min R-QL portfolios use more extreme values as targets, which are parallel to the corresponding quantiles. They can be used to reflect the market situations we simulated. We remark that based on the data simulation process for a Fama–French three-factor model, the market situations are random and we cannot conclude on any link between the studies with $p = 100$ and $p = 400$ in terms of the simulated markets. However, the portfolios in each study (table) are studied within the same universe of risky assets such that they are comparable.

## 6. Empirical study

In this section, we examine the performance of the proposed portfolios with two empirical datasets of S&P 500 components' adjusted closing prices, whose list is as of 3 January 2018, for different periods. The first dataset is from 31 December 2015 to 29 December 2017 ($T = 504$), while the second dataset is from 1 December 2006 to 30 June 2009 ($T = 648$). The former is a recent dataset representing a boom period, while the latter is the financial crisis period. The data is provided by Pun (2018)†.

† The data is openly available at https://data.mendeley.com/datasets/ndxfrshm74/3/files/6b7f4cd6-6996-4031-a002-f99701100f54

We used 252 days of return data to construct the portfolio-weight vector of the next day. This process is repeated until the end of the dataset and we obtain $T - 252$ daily portfolio-weight vectors. With these portfolio-weight vectors, we compute the mean gross exposure, annual return, annual volatility, Sharpe ratio, and daily turnover rate as in table 2. We also conduct the equality tests of portfolio's Sharpe ratio and variance following the methodology in DeMiguel *et al.* (2009) that are parallel to Ledoit and Wolf (2008), Ledoit and Wolf (2011), respectively, and report the $p$-values. The results are summarized in tables 5 and 6. We provide our findings in two evaluation measures of the performance: daily turnover and risk reduction. Then we separately discuss the endogenous targets with their time-series plots. Finally, we comment on the choice of some tuning parameters in our proposed models.

### *6.1. Reduction on daily turnover*

The portfolios perform very differently in terms of their daily turnover under the two studies. In 2016-2017 empirical test, the turnover rate of all range-based risk minimizing portfolios uniformly outperform the corresponding risk minimizing portfolios. In some cases, Min R-Var portfolio has a lower turnover than Ridge Min Var even when the former has a larger gross exposure.

In order to illustrate our point clearly, we compute the turnover rates of Min R-Var and Ridge Min Var portfolios with $C$ ranging from $10^{-5}$ to 1 and $\lambda$ ranging from 0.01 to

Table 4. Results of simulation study with $p = 400$. Note that the annual returns, volatilities, daily turnovers, and the statistics of daily target are reported in percentage.

| Portfolio | Tuning parameter | Gross exposure | Annual return (%) | Annual volatility (%) | Annual Sharpe ratio | Daily turnover (%) | Daily Target Mean (%) | Daily Target SD (%) |
|---|---|---|---|---|---|---|---|---|
| Min R-MAD | $C = 10^{-7}$ | 1.02 | 10.75 | 2.01 | 5.36 | 0.96 | 0.00 | 0.00 |
| Min R-MAD | $C = 10^{-6}$ | 2.24 | 6.24 | 1.07 | 5.86 | 4.00 | 0.00 | 0.00 |
| Min R-MAD | $C = 10^{-5}$ | 2.85 | 5.00 | 0.81 | 6.20 | 4.71 | 0.00 | 0.00 |
| Min R-Var | $C = 10^{-5}$ | 1.00 | 11.31 | 2.13 | 5.30 | 0.47 | 0.00 | 0.00 |
| Min R-Var | $C = 10^{-4}$ | 1.45 | 8.76 | 1.56 | 5.60 | 1.82 | 0.00 | 0.00 |
| Min R-Var | $C = 10^{-3}$ | 2.40 | 6.00 | 0.99 | 6.06 | 3.54 | 0.00 | 0.00 |
| LASSO Min Var | $\lambda = 0.002$ | 1.16 | 10.90 | 2.01 | 5.44 | 1.40 | 0.08 | 0.02 |
| LASSO Min Var | $\lambda = 0.001$ | 1.75 | 7.13 | 1.32 | 5.38 | 4.09 | 0.08 | 0.02 |
| LASSO Min Var | $\lambda = 0.0005$ | 2.23 | 4.90 | 0.95 | 5.17 | 6.80 | 0.08 | 0.02 |
| Ridge Min Var | $\lambda = 0.5$ | 1.05 | 10.46 | 1.94 | 5.37 | 1.03 | 0.08 | 0.02 |
| Ridge Min Var | $\lambda = 0.1$ | 1.85 | 7.36 | 1.29 | 5.71 | 2.99 | 0.08 | 0.02 |
| Ridge Min Var | $\lambda = 0.05$ | 2.40 | 5.68 | 0.96 | 5.90 | 4.16 | 0.08 | 0.02 |
| Min R-QL .25 | $C = 2 * 10^{-2}$ | 1.01 | 10.77 | 2.03 | 5.30 | 1.13 | − 0.26 | 0.03 |
| Min R-QL .25 | $C = 5 * 10^{-2}$ | 1.25 | 9.16 | 1.70 | 5.38 | 2.40 | − 0.28 | 0.04 |
| Min R-QL .25 | $C = 10^{-1}$ | 1.83 | 7.09 | 1.29 | 5.47 | 3.86 | − 0.36 | 0.06 |
| Min QL .25 | $C = 2 * 10^{-2}$ | 1.01 | 10.74 | 2.02 | 5.31 | 1.35 | − 0.21 | 0.03 |
| Min QL .25 | $C = 5 * 10^{-2}$ | 1.31 | 9.01 | 1.66 | 5.42 | 3.00 | − 0.25 | 0.03 |
| Min QL .25 | $C = 10^{-1}$ | 2.04 | 6.49 | 1.18 | 5.51 | 5.44 | − 0.35 | 0.06 |
| Min R-QL .75 | $C = 2 * 10^{-2}$ | 1.01 | 10.74 | 2.03 | 5.29 | 1.02 | 0.27 | 0.03 |
| Min R-QL .75 | $C = 5 * 10^{-2}$ | 1.25 | 9.14 | 1.70 | 5.38 | 2.23 | 0.24 | 0.03 |
| Min R-QL .75 | $C = 10^{-1}$ | 1.82 | 6.97 | 1.29 | 5.39 | 3.76 | 0.15 | 0.04 |
| Min QL .75 | $C = 2 * 10^{-2}$ | 1.01 | 10.70 | 2.02 | 5.30 | 1.23 | 0.24 | 0.03 |
| Min QL .75 | $C = 5 * 10^{-2}$ | 1.31 | 8.93 | 1.66 | 5.38 | 2.82 | 0.21 | 0.03 |
| Min QL .75 | $C = 10^{-1}$ | 2.04 | 6.35 | 1.17 | 5.43 | 5.29 | 0.09 | 0.04 |
| EW | | 1 | 11.93 | 2.28 | 5.24 | 0.25 | | |
| NS | | 1 | 5.37 | 1.26 | 4.25 | 3.81 | | |
| GMV | | 7.36 | 2.41 | 0.96 | 2.52 | 157.23 | | |
| Shrink | | 5.17 | 0.38 | 0.69 | 0.55 | 30.97 | | |
| Oracle | | 3.65 | 0.69 | 0.45 | 1.53 | 1.21 | | |

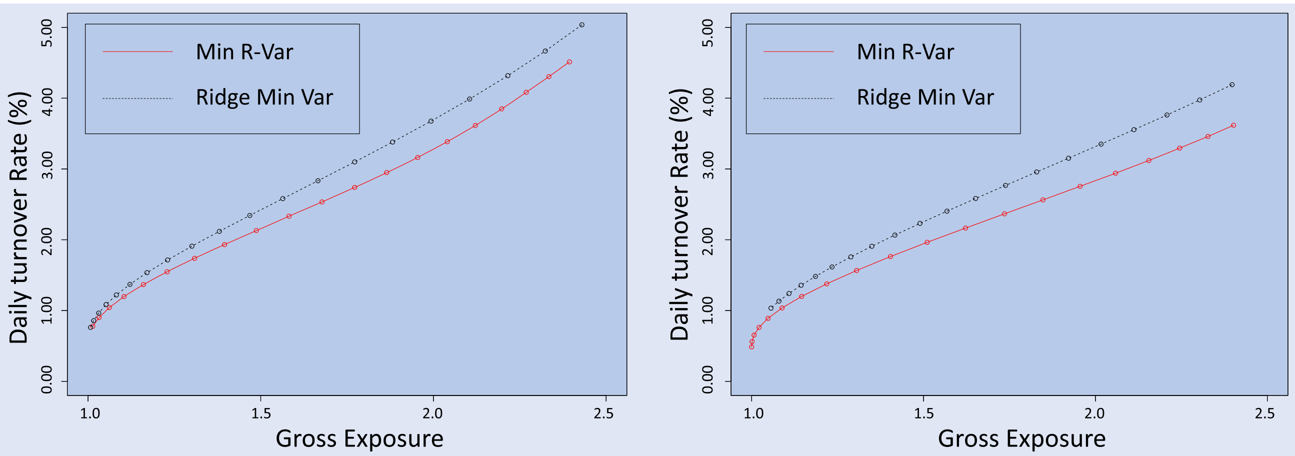


Figure 1. Turnover rate of Min R-Var and Ridge Min Var portfolios for $p = 100$ (left panel) and $p = 400$ (right panel)

1. The interpolation of the turnover rates are plotted against the gross exposure of the portfolios in Figure 2. The left panel shows that Min R-Var portfolio's daily turnover is entirely below Ridge Min Var's. On top of that, both .75 and .25 quantile Min R-QL portfolios exhibit lower gross exposure and daily turnover than their respective Min QL portfolios when $C$ used are the same. At similar gross exposure level, Min R-QL portfolios show lower turnover rate than Min QL portfolios.

On the other hand, for the second empirical test with 2006-2009 data, the comparison between the turnover rates of the two portfolios is inconclusive. The right panel of figure 2 shows that the turnover rates of the Min R-Var and Ridge Min Var are very close to each other across a broad range of gross exposure. Similarly, when we compare the turnover of Min R-QL versus that of Min QL portfolios, neither of them dominate another in terms of daily turnover at similar gross exposure.

The difference can be explained by reconciling the market conditions of the two periods. 2016-2017 has a witnessed resurgent economic growth and blockbuster corporate profit with a sweeping tax cut, hence, it is no surprise that the stock market has a steady expansion. With a stable stock market, it

Table 5. Empirical results with S&P 500 components data in a boom period (2016–2017). In parentheses next to Annual volatility (resp. Sharpe ratio) is the $p$-value of the difference between the variance (resp. Sharpe ratio) of each portfolio from that of the EW portfolio, which is computed using the methodology in Ledoit and Wolf (2011) (resp. Ledoit and Wolf 2008). Note that the annual returns, volatilities, and turnovers are reported in percentage.

| Portfolio | Tuning parameter | Gross exposure | Annual return (%) | Annual volatility (%) | Annual Sharpe ratio | Daily turnover (%) |
|---|---|---|---|---|---|---|
| Min R-MAD | $C = 10^{-7}$ | 1 | 18.12 | 7.07 (0.001) | 2.56 (0.3736) | 0.85 |
| Min R-MAD | $C = 10^{-6}$ | 1.00 | 16.72 | 6.43 (0.001) | 2.60 (0.6913) | 3.04 |
| Min R-MAD | $C = 10^{-5}$ | 1.36 | 14.10 | 5.50 (0.001) | 2.56 (0.9411) | 4.02 |
| Min R-Var | $C = 10^{-5}$ | 1 | 18.03 | 7.03 (0.005) | 2.56 (0.4595) | 0.88 |
| Min R-Var | $C = 10^{-4}$ | 1.01 | 16.68 | 6.37 (0.002) | 2.61 (0.6673) | 1.44 |
| Min R-Var | $C = 10^{-3}$ | 1.27 | 14.31 | 5.68 (0.001) | 2.51 (0.9570) | 3.16 |
| LASSO Min Var | $\lambda = 0.002$ | 1.05 | 15.83 | 6.11 (0.001) | 2.59 (0.8671) | 3.46 |
| LASSO Min Var | $\lambda = 0.001$ | 1.13 | 11.53 | 5.54 (0.001) | 2.08 (0.4955) | 6.21 |
| LASSO Min Var | $\lambda = 0.0005$ | 1.41 | 11.02 | 5.47 (0.001) | 2.01 (0.5325) | 10.55 |
| Ridge Min Var | $\lambda = 0.5$ | 1.00 | 16.78 | 6.17 (0.001) | 2.71 (0.4256) | 1.52 |
| Ridge Min Var | $\lambda = 0.2$ | 1.07 | 15.53 | 5.70 (0.001) | 2.72 (0.5884) | 2.55 |
| Ridge Min Var | $\lambda = 0.1$ | 1.23 | 14.51 | 5.39 (0.001) | 2.68 (0.7403) | 4.02 |
| Min R-QL .25 | $C = 10^{-2}$ | 1.02 | 16.72 | 6.23 (0.001) | 2.68 (0.5015) | 1.75 |
| Min R-QL .25 | $C = 2 * 10^{-2}$ | 1.10 | 14.82 | 5.89 (0.001) | 2.51 (0.9720) | 2.66 |
| Min R-QL .25 | $C = 5 * 10^{-2}$ | 1.33 | 12.40 | 5.63 (0.001) | 2.20 (0.5634) | 4.36 |
| Min QL .25 | $C = 10^{-2}$ | 1.05 | 16.53 | 6.04 (0.001) | 2.73 (0.4795) | 2.67 |
| Min QL .25 | $C = 2 * 10^{-2}$ | 1.20 | 15.51 | 5.72 (0.001) | 2.70 (0.6753) | 4.70 |
| Min QL .25 | $C = 5 * 10^{-2}$ | 1.68 | 14.30 | 5.49 (0.001) | 2.60 (0.9361) | 9.82 |
| Min R-QL .75 | $C = 10^{-2}$ | 1.01 | 16.31 | 6.32 (0.001) | 2.58 (0.8252) | 1.58 |
| Min R-QL .75 | $C = 2 * 10^{-2}$ | 1.06 | 14.78 | 5.97 (0.001) | 2.47 (0.8242) | 2.45 |
| Min R-QL .75 | $C = 5 * 10^{-2}$ | 1.31 | 12.24 | 5.62 (0.001) | 2.17 (0.4436) | 4.57 |
| Min QL .75 | $C = 10^{-2}$ | 1.05 | 16.43 | 5.96 (0.001) | 2.75 (0.5355) | 2.61 |
| Min QL .75 | $C = 2 * 10^{-2}$ | 1.21 | 15.00 | 5.65 (0.001) | 2.65 (0.7922) | 4.56 |
| Min QL .75 | $C = 5 * 10^{-2}$ | 1.65 | 13.71 | 5.39 (0.001) | 2.53 (0.9880) | 9.59 |
| EW | | 1 | 18.30 | 7.21 | 2.53 | 0.82 |
| NS | | 1 | 13.07 | 5.35 (0.001) | 2.44 (0.9041) | 7.12 |
| GMV | | 3.56 | 1.39 | 7.63 (0.352) | 0.18 (0.0509) | 127.90 |
| Shrink | | 2.48 | 19.90 | 6.56 (0.193) | 3.03 (0.6334) | 38.03 |

is more likely for the daily return to fall into the target range based on the existing portfolio. Therefore, all range-based risk minimizing portfolios demonstrating low turnover aligns with Propositions 2 and 3. On the other hand, during the test period of 2006-2009, the market is severely volatile as evidenced by the 43.97% annual volatility of EW portfolio. The volatile stock market makes it much harder for daily return to fall within the target range. Since the premise of proposition 2 and 3 cannot be fulfilled, range-based risk minimizing portfolios do not necessarily demonstrate low daily turnover.

### 6.2. *Improvement on portfolio's volatility and sharpe ratio*

The Sharpe ratios of all portfolios for the recession period of 2007-2009 are negative and thus lack interpretability. As explained in the previous subsection, during the recession period, the daily return is volatile and falls outside the target range. However, we can observe that the annual returns and volatilities of range-based risk minimizing portfolios are comparable to that of the corresponding traditional risk minimizing portfolios.

In this subsection, we focus on the interpretation of the results in table 5 for the boom period of 2016-2017. It can be seen that the minimum RRM portfolios' variances are lower than that of EW portfolio significantly as evidenced by the low $p$-values. We also see that most of the range-based risk minimizing portfolios outperform the EW portfolio in terms of the Sharpe ratio but the differences are not significant. It is also noteworthy that the Shrink portfolio, which uses the shrinkage approach in Ledoit and Wolf (2004) to estimate the covariance matrix, performs the best in terms of the Sharpe ratio. However, when the transaction cost is taken into account, almost all the profits made by Shrink portfolio is eroded away. It is also common in portfolios considered in DeMiguel *et al.* (2009) that some portfolios perform well but the turnover rates are high, which make them impractical.

In table 7, we extract the portfolios in table 5 with best tuning parameters and compute their average return, volatility, and Sharpe ratio based on the net out-of-sample test returns after deduction of proporitional transaction cost with the rate of 30 basis point (bp). In the last column of table 7, we compute the transaction cost rate, under which the portfolio has the same Sharpe ratio (Net) as that of the EW portfolio. Hence, if the real transaction cost rate is below this unit cost, the portfolio will outperform the EW portfolio in terms of the Sharpe ratio.

Even with this conservative estimate of the transaction cost rate (30 bp), the net return of the Shrink portfolio sharply reduces to −8.30% and thus its Sharpe ratio becomes negative. All range-based risk minimizing portfolios strictly outperform the Shrink portfolio in terms of all metrics after deducting the transaction cost. When the unit transaction cost is higher, the EW strategy tend to dominate all other strategies due to its low daily turnover. Amongst all portfolios, the Min R-Var and Min R-QL .25 portfolios allow the unit transaction

Table 6. Empirical results with S&P 500 components data in a recession period (2007-2009). In parentheses next to Annual volatility (resp. Sharpe ratio) is the $p$-value of the difference between the variance (resp. Sharpe ratio) of each portfolio from that of the EW portfolio, which is computed using the methodology in Ledoit and Wolf (2011) (resp. Ledoit and Wolf 2008). Note that the annual returns, volatilities, and turnovers are reported in percentage.

| Portfolio | Tuning parameter | Gross exposure | Annual return (%) | Annual volatility (%) | Annual Sharpe ratio | Daily turnover (%) |
|---|---|---|---|---|---|---|
| Min R-MAD | $C = 10^{-7}$ | 1 | − 5.93 | 42.49 (0.001) | − 0.13 (0.3217) | 2.01 |
| Min R-MAD | $C = 10^{-6}$ | 1.03 | − 9.50 | 32.46 (0.001) | − 0.29 (0.2747) | 2.32 |
| Min R-MAD | $C = 10^{-5}$ | 1.99 | − 14.84 | 19.13 (0.001) | − 0.77 (0.3167) | 8.31 |
| Min R-Var | $C = 10^{-5}$ | 1.00 | − 9.32 | 37.77 (0.001) | − 0.24 (0.1049) | 1.98 |
| Min R-Var | $C = 10^{-4}$ | 1.31 | − 15.86 | 24.96 (0.001) | − 0.63 (0.1349) | 3.60 |
| Min R-Var | $C = 10^{-3}$ | 2.05 | − 17.69 | 18.64 (0.001) | − 0.94 (0.2438) | 8.12 |
| Min Var | $\lambda = 0.002$ | 1.36 | − 16.58 | 19.06 (0.001) | − 0.86 (0.3906) | 9.06 |
| Min Var | $\lambda = 0.001$ | 1.61 | − 16.00 | 17.76 (0.001) | − 0.90 (0.4146) | 13.45 |
| Min Var | $\lambda = 0.0005$ | 2.04 | − 14.24 | 17.20 (0.001) | − 0.82 (0.5205) | 21.75 |
| Ridge Min Var | $\lambda = 1$ | 1.25 | − 16.69 | 26.18 (0.001) | − 0.63 (0.1768) | 3.20 |
| Ridge Min Var | $\lambda = 0.5$ | 1.49 | − 17.52 | 22.66 (0.001) | − 0.77 (0.1868) | 4.57 |
| Ridge Min Var | $\lambda = 0.2$ | 1.94 | − 17.13 | 19.35 (0.001) | − 0.88 (0.2627) | 7.39 |
| Min R-QL .25 | $C = 5 * 10^{-3}$ | 1.19 | − 20.86 | 28.79 (0.001) | − 0.72 (0.0529) | 3.15 |
| Min R-QL .25 | $C = 10^{-2}$ | 1.45 | − 24.04 | 24.03 (0.001) | − 1.00 (0.0779) | 4.92 |
| Min R-QL .25 | $C = 2 * 10^{-2}$ | 1.83 | − 23.77 | 20.79 (0.001) | − 1.14 (0.1219) | 7.53 |
| Min QL .25 | $C = 5 * 10^{-3}$ | 1.22 | − 20.02 | 28.19 (0.001) | − 0.71 (0.0709) | 3.44 |
| Min QL .25 | $C = 10^{-2}$ | 1.51 | − 23.45 | 23.23 (0.001) | − 1.00 (0.0799) | 5.61 |
| Min QL .25 | $C = 2 * 10^{-2}$ | 1.97 | − 24.36 | 20.35 (0.001) | − 1.19 (0.1788) | 9.35 |
| Min R-QL .75 | $C = 5 * 10^{-3}$ | 1.19 | − 18.75 | 29.48 (0.001) | − 0.61 (0.0959) | 3.06 |
| Min R-QL .75 | $C = 10^{-2}$ | 1.39 | − 18.99 | 25.68 (0.001) | − 0.73 (0.1479) | 4.51 |
| Min R-QL .75 | $C = 2 * 10^{-2}$ | 1.72 | − 19.17 | 22.39 (0.001) | − 0.85 (0.1788) | 6.85 |
| Min QL .75 | $C = 5 * 10^{-3}$ | 1.22 | − 18.34 | 29.15 (0.001) | − 0.62 (0.1429) | 3.43 |
| Min QL .75 | $C = 10^{-2}$ | 1.46 | − 20.43 | 25.32 (0.001) | − 0.80 (0.1538) | 5.34 |
| Min QL .75 | $C = 2 * 10^{-2}$ | 1.87 | − 20.03 | 21.94 (0.001) | − 0.91 (0.1658) | 8.65 |
| EW | | 1 | − 5.45 | 43.97 | − 0.12 | 2.05 |
| NS | | 1 | − 10.76 | 21.34 (0.001) | − 0.50 (0.2865) | 5.15 |
| GMV | | 7.67 | − 23.14 | 23.10 (0.001) | − 1.00 (0.4076) | 154.17 |
| Shrink | | 5.13 | − 20.52 | 17.45 (0.001) | − 1.17 (0.3636) | 40.92 |

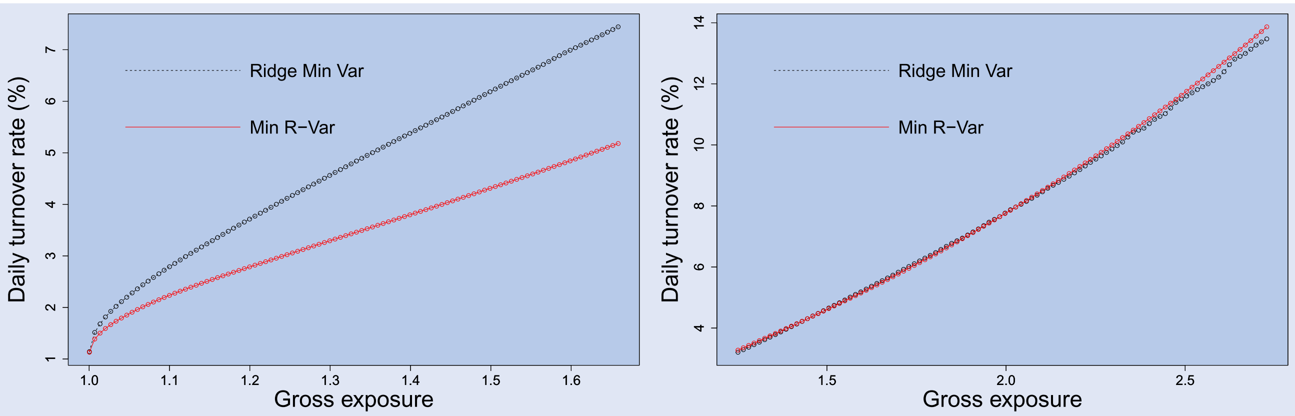


Figure 2. Daily turnover rate of Min R-Var and Ridge Min Var portfolios from 2016 to 2017 (left panel) and from 2006 to 2009 (right panel).

cost to be as high as 39 and 42 bp for them to have the same Sharpe ratio as EW strategy. Tables 5 and 7 confirm that the introduction of range-based risk measures helps reducing the turnover rate without sacrificing the effect of diversification, leading to the improved performance when the transaction cost is taken into account.

### 6.3. *Portfolio targets*

We record the endogenous portfolio targets of all portfolios over the test periods and they are reported in figure 3. We make two major observations. First, the introduction of range-based risk concept stabilizes the target of Min Var portfolios. As shown in figure 3(a), (c), the target returns of Min R-MAD and Min R-Var are consistently close to zero while the targets of LASSO and Ridge Min Var portfolios fluctuate in a much wider range. The targets of Min R-MAD and Min R-Var portfolios remain stable across both boom and recession periods. In contrast, conventional portfolios have postive target returns during 2016–2017 test period whereas that number changes from positive to negative during the 2006–2009 recession.

Second, the targets of portfolios that minimize higher quantile loss are higher than their counterpart portfolios that

Table 7. Net annual return and volatility in a boom period (2016–2017) with proportional transaction cost (30 bp). * the tuning parameter that produces the highest Sharpe ratio without considering the transaction cost is chosen. ** the transaction cost rate, at which level the portfolio has the same Sharpe ratio (Net) as that of the EW portfolio.

| Portfolio | Tuning parameter* | Gross exposure | Net Annual Return | Net Annual Volatility | Net Annual Sharpe ratio | Unit Cost** |
|---|---|---|---|---|---|---|
| Min R-MAD | $C = 10^{-6}$ | 1.00 | 15.10 | 6.41 | 2.35 | 31 bp |
| Min R-Var | $C = 10^{-4}$ | 1.01 | 15.13 | 6.37 | 2.37 | 39 bp |
| LASSO Min Var | $\lambda = 0.002$ | 1.05 | 12.98 | 6.12 | 2.12 | 10 bp |
| Ridge Min Var | $\lambda = 0.2$ | 1.07 | 13.33 | 5.71 | 2.33 | 28 bp |
| Min R-QL .25 | $C = 10^{-2}$ | 1.02 | 14.92 | 6.22 | 2.39 | 42 bp |
| Min QL .25 | $C = 10^{-2}$ | 1.05 | 14.04 | 6.04 | 2.32 | 27 bp |
| Min R-QL .75 | $C = 10^{-2}$ | 1.01 | 14.63 | 6.32 | 2.31 | 20 bp |
| Min QL .75 | $C = 10^{-2}$ | 1.05 | 14.05 | 5.96 | 2.35 | 31 bp |
| EW | | 1 | 16.89 | 7.18 | 2.34 | – |
| NS | | 1 | 7.59 | 5.40 | 1.40 | < 0 |
| GMV | | 3.56 | – 94.58 | 8.24 | – 11.47 | < 0 |
| Shrink | | 2.48 | – 8.30 | 6.70 | – 1.23 | 5 bp |

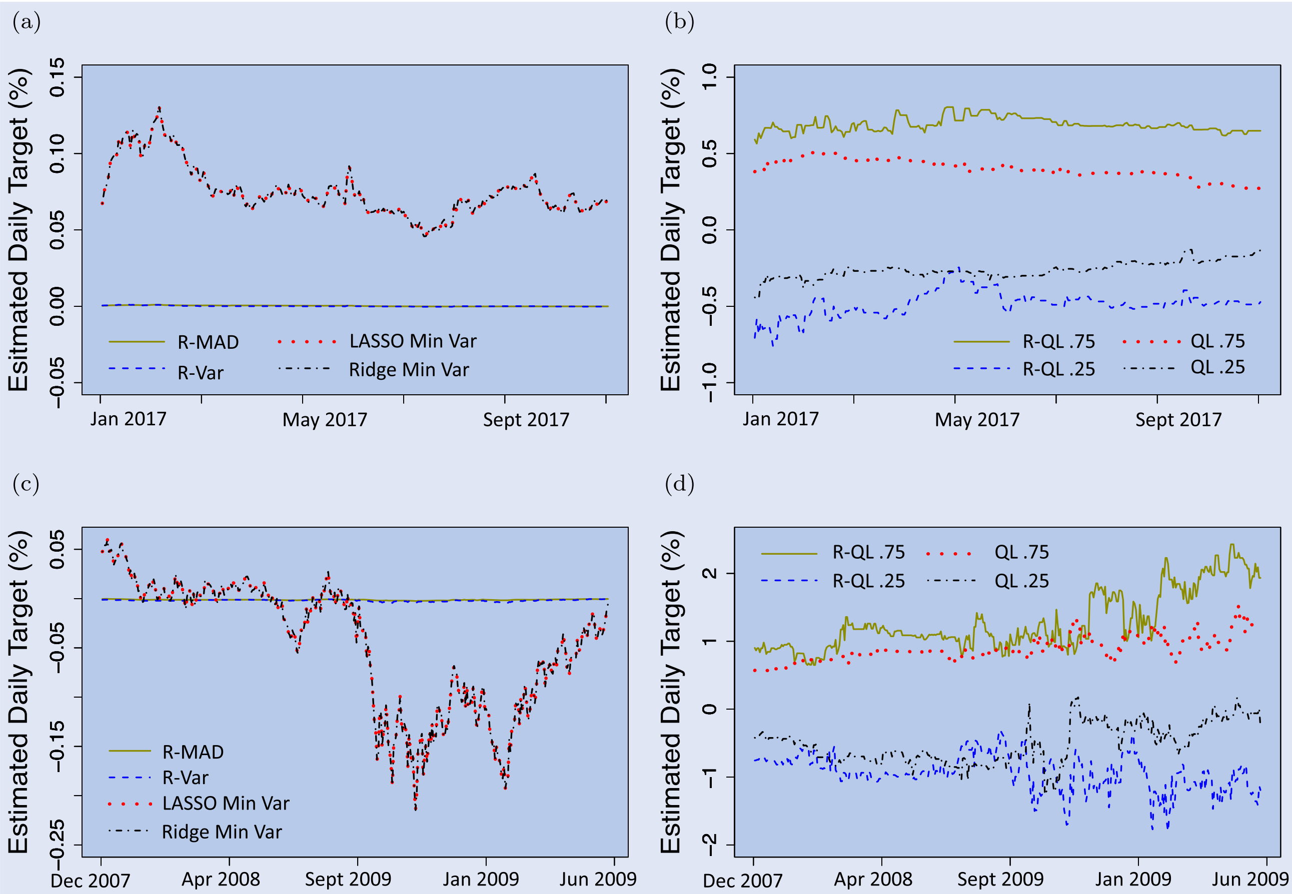


Figure 3. Endogenous portfolio targets. The values of $C$ and $\lambda$ are the same as those listed in table 7. (a) Portfolio targets of Min R-MAD, R-Var, LASSO and Ridge Min Var in a boom period (2016–2017). (b) Portfolio targets of Min R-QL.75, R-QL .25, QL .75 and QL .25 in a boom period (2016–2017). (c) Portfolio targets of Min R-MAD, R-Var, LASSO and Ridge Min Var in a recession period (2006–2009) and (d) Portfolio targets of Min R-QL.75, R-QL .25, QL .75 and QL .25 in a recession period (2006–2009).

minimize lower quantile loss. Figure 3(b), (d) show that the .75 quantile target is higher than the .25 quantile target for both minimum DRM and RRM portfolios. Moreover, though it is hard to visualize which portfolio's target is more stable, Min R-QL .75 and .25 portfolios have several consecutive days that have the same target return whereas Min QL .75 and .25 portfolios do not have any of that.

### 6.4. *Further discussions on the choice of parameters*

Two tuning parameters are crucial to the performances of range-based risk minimizing portfolios. The acceptance range is dependent on the parameter $\epsilon$ whereas the regularization term $C$ determines the gross exposure of the portfolios. According to our simulation and empirical studies, we are able to provide insights about the choices of these two parameters.

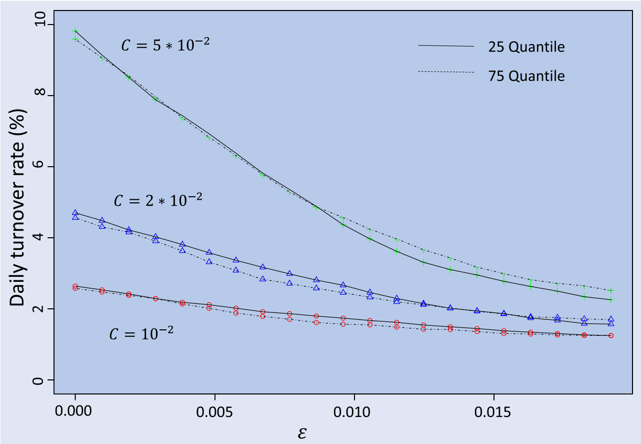


Figure 4. Change of daily turnover rate against $\epsilon$ with different $C$.

**6.4.1. Choice of $\epsilon$.** The $\epsilon$ parameter determines the number of daily returns that fall within the target range. In our paper, we adopt the method recommended by Cherkassky and Ma (2002), Cherkassky and Ma (2004) to determine the $\epsilon$ with the training data. Such a choice of $\epsilon$ produces portfolios with plausible returns and turnover rate. All range-based risk minimizing portfolios including Min R-Var, Min R-QL .25 and Min R-QL .75 portfolios demonstrate comparable annual return, annual volatility and Sharpe ratio to their counterpart (Ridge Min-Var, Min QL .25 and Min QL .75). They also consistently demonstrate lower turnover rate as compared to their counterparts. In figure 4, we can see that the turnover rate is generally decreasing with $\epsilon$. Interestingly, the decrease is only significant up until $\epsilon = 0.01$ and our method gives an estimate of $\epsilon$ at around 0.0095 for the empirical datasets we considered.

**6.4.2. Choice of $C$.** According to our simulation and empirical studies, the Min R-MAD portfolio has a good performance when its tuning parameter C is equal to $10^{-6}$ while the Min R-Var portfolio tends to perform well when C is equal to $10^{-4}$. As for Min R-QL .25 and Min R-QL .75 portfolios, optimal $C$ values tend to vary from $10^{-2}$ to $5 * 10^{-2}$ depending on the market condition.

## 7. Conclusion

This paper provides a general way of improving the risk-minimizing portfolio's performance in terms of lower transaction cost. Contrary to conventional minimum risk portfolios, range-based risk minimizing portfolios aim to minimize the expected deviation of daily return from *a range of* values instead of a *single* target value. The novel risk measure has a few advantages over traditional ones. Firstly, portfolio construction with ranged-based risk measures aligns with financial intuition as most investors do not make major changes to portfolios on daily basis. Secondly, the range-based risk-minimizing portfolios incur less transaction cost as shown by both theoretical results and empirical tests. Thirdly, after drawing the connection between range-based risk measures and SVR, we managed to apply RRM to a large universe of assets, even when number of assets is larger than the number of observations.

The range-based risk minimizing portfolios can be obtained through efficient algorithms of support vector regression and support vector quantile regression. With the dual-primal form of the SVR problem, we manage to prove the theoretical robustness of the Min R-MAD, Min R-QL and Min R-Var portfolios: when both the exiting and entering data are within the acceptable set, the portfolio-vector weights remain unchanged. Both the simulation and empirical studies numerically show that the range-based risk minimizing portfolios typically have lower daily turnover rate at a similar gross exposure level. Moreover, although turnover rate is reduced, the portfolio's performance is even better than the traditional risk minimizing portfolios and some other competitors.

In this paper, we only consider amendment to the deviation risk measures and the corresponding portfolio selection. Future research can extend the idea of range-based risk to other risk measures, such as coherent risk measures, and find its linkage with set-valued risk measures in Jouini *et al.* (2004).

## 8. Open Data

This article has earned the Center for Open Science badge for Open Data. The data are openly accessible at http://dx.doi.org/10.17632/ndxfrshm74.3.


## Acknowledgments

Chi Seng Pun gratefully acknowledges Ministry of Education (Singapore), AcRF Tier 2 grant (Reference No: MOE2017-T2-1-044) and Nanyang Technological University Start-up Grant (Reference No: 04INS000248C230) for the funding of this research.


## Disclosure statement

No potential conflict of interest was reported by the authors.


## Funding

Chi Seng Pun gratefully acknowledges Ministry of Education (Singapore), AcRF Tier 2 grant (Reference No: MOE2017-T2-1-044) and Nanyang Technological University, Start-up Grant (Reference No: 04INS000248C230) for the funding of this research.

## ORCID

*Chi Seng Pun* http://orcid.org/0000-0002-7478-6961

## References

Bonaccolto, G., Caporin, M. and Paterlini, S., Asset allocation strategies based on penalized quantile regression. *Comput. Manage. Sci.*, 2017, **15**, 1–32.
Broadie, M., Computing efficient frontiers using estimated parameters. *Ann. Oper. Res.*, 1993, **45**, 21–58.
Brodie, J., Daubechies, I., Mol, C.D., Giannone, D. and Loris, I., Sparse and stable Markowitz portfolios. *Proc. Nat. Acad. Sci.*, 2009, **106**, 12267–12272.
Cherkassky, V. and Ma, Y., Selection of meta-parameters for support vector regression. In *Artificial Neural Networks–ICANN 2002*, pp. 687–693, 2002 (Springer: Berlin Heidelberg).
Cherkassky, V. and Ma, Y., Practical selection of SVM parameters and noise estimation for SVM regression. *Neural. Netw.*, 2004, **17**, 113–126.
Chiu, M.C., Pun, C.S. and Wong, H.Y., Big data challenges of high-dimensional continuous-Time mean-Variance portfolio selection and a remedy. *Risk. Anal.*, 2017, **37**, 1532–1549.
Chopra, V.K. and Ziemba, W.T., The effect of errors in means, variances, and covariances on optimal portfolio choice. In *Handbook of the Fundamentals of Financial Decision Making*, pp. 365–373, 2013 (World Scientific).
Clarke, R., de Silva, H. and Thorley, S., Minimum-Variance portfolio composition. *J. Portfolio Manage.*, 2011, **37**, 31–45.
Cortes, C. and Vapnik, V., Support-Vector networks. *Mach. Learn.*, 1995, **20**, 273–297.
Davis, M.H.A. and Norman, A.R., Portfolio selection with transaction costs. *Math. Oper. Res.*, 1990, **15**, 676–713.
DeMiguel, V., Garlappi, L., Nogales, F.J. and Uppal, R., A generalized approach to portfolio optimization: improving performance by constraining portfolio norms. *Manage. Sci.*, 2009, **55**, 798–812.
DeMiguel, V., Garlappi, L. and Uppal, R., Optimal versus naive diversification: how inefficient is the 1/N portfolio strategy? *Rev. Financ. Stud.*, 2009, **22**, 1915–1953.
Drucker, H., Burges, C.J.C., Kaufman, L., Smola, A.J. and Vapnik, V., Support vector regression machines. In *Advances in Neural Information Processing Systems 9*, edited by M.C. Mozer, M.I. Jordan and T. Petsche, pp. 155–161, 1997 (MIT Press).
Fan, J., Fan, Y. and Lv, J., High dimensional covariance matrix estimation using a factor model. *J. Econom.*, 2008, **147**, 186–197.
Fan, J., Zhang, J. and Yu, K., Vast portfolio selection with gross-exposure constraints. *J. Am. Stat. Assoc.*, 2012, **107**, 592–606.
Ho, C.H. and Lin, C.J., Large-scale linear support vector regression. *J. Mach. Learn. Res.*, 2012, **13**, 3323–3348.
Jagannathan, R. and Ma, T., Risk reduction in large portfolios: Why imposing the wrong constraints helps. *J. Finance.*, 2003, **58**, 1651–1683.
Jouini, E., Meddeb, M. and Touzi, N., Vector-valued coherent risk measures. *Finance Stoch.*, 2004, **8**.
Kecman, V., Support vector machines – an introduction. In *Support Vector Machines: Theory and Applications*, pp. 1–47, 2005 (Springer: Berlin Heidelberg).
Kempf, A. and Memmel, C., Estimating the global minimum variance portfolio. *Schmalenbach Business Rev.*, 2006, **58**, 332–348.
Klopfenstein, Q. and Vaiter, S., Linear Support Vector Regression with Linear Constraints. arXiv:1911.02306, 2019.
Konno, H. and Yamazaki, H., Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Manage. Sci.*, 1991, **37**, 519–531.
Kuhn, H.W. and Tucker, A.W., Nonlinear programming. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*, pp. 481–492, 1951 (University of California Press: Berkeley and Los Angeles).
Ledoit, O. and Wolf, M., Honey, I shrunk the sample covariance matrix. *J. Portfolio Manage.*, 2004, **30**, 110–119.
Ledoit, O. and Wolf, M., Robust performance hypothesis testing with the sharpe ratio. *J. Empirical Finance*, 2008, **15**, 850–859.
Ledoit, O. and Wolf, M., Robust performances hypothesis testing with the variance. *Wilmott*, 2011, **2011**, 86–89.
Markowitz, H., Portfolio selection. *J. Finance.*, 1952, **7**, 77–91.
Mei, X. and Nogales, F.J., Portfolio selection with proportional transaction costs and predictability. *J. Bank. Financ.*, 2018, **94**, 131–151.
Michaud, R.O., The markowitz optimization enigma: Is 'Optimized' optimal? *Financial Anal. J.*, 1989, **45**, 31–42.
Muthuraman, K., A computational scheme for optimal investment - consumption with proportional transaction costs. *J. Econom. Dyn. Control*, 2007, **31**, 1132–1159.
Olivares-Nadal, A.V. and DeMiguel, V., Technical note—A robust perspective on transaction costs in portfolio optimization. *Oper. Res.*, 2018, **66**, 733–739.
Plyakha, Y., Uppal, R. and Vilkov, G., Why does an equal-weighted portfolio outperform value- and price-weighted portfolios? *SSRN Electron. J.*, 2012.
Pun, C.S., Low- and High-dimensional Asset Prices Data. Mendeley Data, v3, 2018.
Pun, C.S., Time-consistent mean-variance portfolio selection with only risky assets. *Econ. Model.*, 2018, **75**, 281–292.
Pun, C.S. and Wong, H.Y., Resolution of degeneracy in Merton's portfolio problem. *SIAM J. Financ. Math.*, 2016, **7**, 786–811.
Pun, C.S. and Wong, H.Y., A linear programming model for selection of sparse high-dimensional multiperiod portfolios. *Eur. J. Oper. Res.*, 2019, **273**, 754–771.
Pun, C.S. and Ye, Z., Dynamically optimal multiperiod mean-variance portfolio subject to proportional transaction costs and no-shorting constraint. Working Paper of NTU Singapore, 2019.
Seok, K.H., Cho, D., Hwang, C. and Shim, J., Support vector quantile regression using asymmetric e-insensitive loss function. In *Proceedings of the 2010 2nd International Conference on Education Technology and Computer*, Jun. 2010 (IEEE).
Smola, A.J. and Schlkopf, B., A tutorial on support vector regression. *Stat. Comput.*, 2004, **14**, 199–222.
Still, S. and Kondor, I., Regularizing portfolio optimization. *New. J. Phys.*, 2010, **12**, 075034.
van den Doel, K., Ascher, U.M. and Haber, E., The lost honor of $\ell$2-based regularization. In *Large Scale Inverse Problems: Computational Methods and Applications in the Earth Sciences*, 2012 (De Gruyter).
Vapnik, V.N., *The Nature of Statistical Learning Theory*, 2000 (Springer: New York).

## Appendix. dual problem of R-QL minimization with a fixed target

Here, we consider problem (11) with a fixed target, i.e. $a_t \equiv \mathcal{T}$, where $\mathcal{T}$ is a pre-specified constant. The Lagrangian of this constrained minimization problem is defined as

$$\begin{aligned} L = {} & \frac{1}{2}\widetilde{w}_t'\widetilde{w}_t + C\sum_{s=t-M}^{t-1}[\eta\xi_s + (1-\eta)\xi_s^*] \\ & + \sum_{s=t-M}^{t-1}[\alpha_s(Y_s - X_s'\widetilde{w}_t - \mathcal{T} - \epsilon - \xi_s)] \\ & + \sum_{s=t-M}^{t-1}[\alpha_s^*(-Y_s + X_s'\widetilde{w}_t + \mathcal{T} - \epsilon - \xi_s^*)] \\ & - \sum_{s=t-M}^{t-1}\beta_s\xi_s - \sum_{s=t-M}^{t-1}\beta_s^*\xi_s^* \end{aligned} \tag{A1}$$

Then the first-order optimality conditions for (A1) are as follows:

$$\frac{\partial L}{\partial \widetilde{w}_t} = \widetilde{w}_t - \sum_{s=t-M}^{t-1} \alpha_s X_s + \sum_{s=t-M}^{t-1} \alpha_s^* X_s = 0,$$
$$\frac{\partial L}{\partial \xi_s} = C\eta - \alpha_s - \beta_s = 0, \quad \frac{\partial L}{\partial \xi_s^*} = C(1-\eta) - \alpha_s^* - \beta_s^* = 0. \tag{A2}$$

Substituting the first-order conditions in (A2) (form of $\tilde{w}$) into the Lagrangian in (A1) yields that

$$\begin{aligned}
L &= \frac{1}{2} \sum_{r,s=t-M}^{t-1} (\alpha_r - \alpha_r^*)(\alpha_s - \alpha_s^*) X_r' X_s \\
&\quad + \sum_{s=t-M}^{t-1} \alpha_s (Y_s - X_s' \widetilde{w}_t - \mathcal{T} - \epsilon) \\
&\quad + \sum_{s=t-M}^{t-1} \alpha_s^* (-Y_s + X_s' \widetilde{w}_t + \mathcal{T} - \epsilon) \\
&= -\frac{1}{2} \sum_{r,s=t-M}^{t-1} (\alpha_r - \alpha_r^*)(\alpha_s - \alpha_s^*) X_r' X_s \\
&\quad + \sum_{s=t-M}^{t-1} \alpha_s (Y_s - \mathcal{T} - \epsilon) + \sum_{s=t-M}^{t-1} \alpha_s^* (-Y_s + \mathcal{T} - \epsilon),
\end{aligned}$$

The dual problem is to maximize this Lagrangian subject to inequality constraints:

$$0 \le \alpha_s \le \eta C \text{ and } 0 \le \alpha_s^* \le (1-\eta) C \quad \text{for } s \in \{t-M, \ldots, t-1\}.$$

Again, this constrained quadratic program can be solved with R package ‘quadprog’.