A practical guide to robust portfolio optimization[1]

20th November 2019

Chenyang Yin
is analyst in Multi-Asset team of the Quant Research Group at
BNP Paribas Asset Management
14 rue Bergère, 75009 Paris, France.
chenyang.yin@bnpparibas.com, Tel. +33 (0)1 58 97 13 02

Romain Perchet
is head of Multi-Asset team of the Quant Research Group at
BNP Paribas Asset Management
14 rue Bergère, 75009 Paris, France.
romain.perchet@bnpparibas.com, Tel. +33 (0)1 58 97 28 23

François Soupé
is co-head of the Quant Research Group at
BNP Paribas Asset Management
14 rue Bergère, 75009 Paris, France.
francois.soupe@bnpparibas.com, Tel. +33 (0)1 58 97 21 96

[1] We are grateful for valuable comments from Raul Leote de Carvalho, Benoit Bellone, Frederic Abergel and Koye Somefun.

A practical guide to robust portfolio optimization

20th November 2019

ABSTRACT

Robust optimization considers uncertainty in inputs to address the shortcomings of mean-variance optimization. We investigate the mechanisms by which robust optimization achieves its goal and give practical guidance regarding its parametrization. We show that quadratic uncertainty sets are preferred to box uncertainty sets, that a diagonal uncertainty matrix with only variances should be used, and that the level of uncertainty can be chosen based on Sharpe ratios. We use examples with the proposed parametrization to show that robust optimization efficiently overcomes the weaknesses of mean-variance optimisation and can be applied in real investment problems like multi-asset portfolio management or robo-advising.

Keywords: Robust optimization, Portfolio construction, Mean-variance optimization, Multi-asset, Asset Allocation

I. Introduction

The influential work of Markowitz (1952, 1959) laid the groundwork for the modern portfolio construction theory. However, the practical application of this framework in the asset management industry has been disappointing. In fact, the inputs (expected returns and covariance matrix) for Mean-Variance Optimization (MVO) needs to be estimated, either statistically from historical data or with a factor or valuation model. Chopra and Ziemba (1993) and KallBerg and Ziemba (1984) find that the uncertainty in expected returns are roughly ten times as important as that in the covariance matrix. Yam *et al.* (2016) consider the uncertainty effect by investigating different robust formulations and find that the uncertainty on the expected returns is more significant than the uncertainty on the covariance matrix for the sensitivity of the solution. Therefore, in this paper, we focus on the uncertainty in expected returns while assuming that the covariance of returns is known.

The main problem of MVO is that it not only fails to take into account the uncertainty in the estimation process of expected returns but also tends to amplify them. This issue is analytically reported and empirically tested by Best and Grauer (1991), Chopra and Ziemba (1993) and Jobson and Korkie (1983). In general, the correlation coefficients of asset returns are non-zero, so correlation matrix is different from identity matrix. The covariance matrix, resulting from a correlation matrix different from an identity matrix, contains small eigenvalues. The inverse of a matrix plagued by small eigenvalues, in the solution to MVO, accentuates the impact of uncertainty in expected returns on the final result (Roncalli 2013 and Bruder *et al.* 2013), leads to error-maximized and financially irrelevant investment portfolios (Michaud 1989) and increases the sensitivity of the MVO optimal portfolios to small changes in inputs (Black and Litterman 1990). Typically, if two assets are highly correlated, a tiny difference in expected returns may lead to a large long-short position in the MVO portfolio (Best and Grauer 1991). He and Litterman (1999) point out that many investment managers find MVO portfolio weights extreme and counter-intuitive.

In portfolio optimization literature, two approaches have been proposed to mitigate the previously mentioned drawbacks suffered by MVO. The first approach, embodied by the Black-Litterman model or more broadly by Bayesian shrinkage approaches, proposes making robust the estimation of expected returns before feeding them into MVO. The second one, exemplified by robust optimization (RO), takes into account the uncertainty in the optimization objective function and provides another promising alternative to MVO. In this paper, we provide a practical guide to implementing RO in the multi-asset investment universe.

As opposed to the MVO which treats its estimated expected returns in a deterministic manner, RO, introduced by El Ghaoui and Lebret (1997) and Ben-Tal and Nemirovski (1998), assumes that the estimated expected returns are random variables and seeks to find the optimal portfolio even when the realized values of inputs deviate from the estimated ones within some given set. The latter is

called uncertainty set and defines the degree of deviation one wishes to be protected from. In RO literature applied to portfolio construction, two forms of uncertainty set have been studied (Fabozzi *et al.* 2007). Goldfarb and Iyengar (2003) analyze the quadratic uncertainty set for expected returns $(\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}})^T \boldsymbol{\Omega}_{\boldsymbol{\mu}}^{-1} (\boldsymbol{\mu} - \widehat{\boldsymbol{\mu}}) \leq \kappa^2$, with $\boldsymbol{\mu}$ the $n \times 1$ expected returns vector, T transpose, $\widehat{\boldsymbol{\mu}}$ estimated expected returns vector, $\boldsymbol{\Omega}_{\boldsymbol{\mu}}$ the uncertainty matrix and κ the level of uncertainty. They find that the RO constructed with such an uncertainty set can be solved as a second-order cone program. Tütüncü and König (2004) obtain similar results with box uncertainty set $|\mu_i - \hat{\mu}_i| \leq \xi_i,\ i = 1,2, \ldots, n$, with μ_i the expected return of asset i, $\hat{\mu}_i$ the estimated expected return of asset i and ξ_i the level of uncertainty for the expected return estimation of asset i. As shown above, one needs to specify two additional parameters, namely the uncertainty matrix $\boldsymbol{\Omega}_{\boldsymbol{\mu}}$ and the level of uncertainty κ, in a quadratic uncertainty set. As for box uncertainty set, one has to determine the level of uncertainty ξ_i for each asset i.

A great deal has been written about the uncertainty matrix $\boldsymbol{\Omega}_{\boldsymbol{\mu}}$ in RO literature. Ceria and Stubbs (2006) argue that it is important to distinguish the covariance matrix of asset returns from the uncertainty matrix of the estimation error without providing further guidance on how to choose the uncertainty matrix. Scherer (2006) analyzes the uncertainty matrix proportional to the covariance matrix of asset returns. He demonstrates that the robust optimal portfolio can be expressed as a weighted average of the MVO portfolio and the minimum-variance portfolio and argues that RO provides no additional benefit compared with Bayesian shrinkage approaches. Fabozzi *et al.* (2007) and Stubbs and Vance (2005) argue that the uncertainty matrix should use sample variance as the estimation error; therefore, they suggest using the diagonal of the sample covariance matrix as the uncertainty matrix. Heckel *et al.* (2016) find the limiting portfolios of RO, formulated with different uncertainty matrices, for the highest and lowest uncertainty levels. However, these papers fail to evaluate the pros and cons of each uncertainty matrix, and they do not provide clear guidelines for the choice of the uncertainty matrix in the quadratic uncertainty set.

As regards to the level of uncertainty parameter κ, very little has been published. Cornuéjols *et al.* (2018) briefly discuss the fact that κ is related to the size of the uncertainty set and needs to be chosen based on the desired level of robustness. Most authors analyze κ from a purely probabilistic point of view, neglecting the fact that κ is also a parameter in an optimization problem. From a probabilistic point of view, κ represents the size of an uncertainty set, so it corresponds to the quantile yielded by the inverse cumulative distribution function. For practical purposes, κ should be determined to provide a wide enough safety margin for investors and its determination depends on the assumptions of the distribution of asset returns. Most RO empirical applications consists of simply varying κ from one extremum to another. For instance, Goldfarb and Iyengar (2003) carry out experiments on real data by varying κ from the first to the 99th percentile. They conclude that the estimation uncertainty is unknown a priori. Therefore, the correct choice of κ remains a vexing problem and they also suggest that κ should be adjusted dynamically. Scherer (2006) and Ceria

and Stubbs (2006) assume that the returns follow an elliptical distribution and that κ is derived as the inverse cumulative distribution function of the chi-squared distribution. Ceria and Stubbs (2006) and Santos (2010) run an out-of-sample performance comparison between RO and MVO and find that RO outperforms MVO. However, they provide no justification for performing the simulations with κ equal to 1, 3, 5 and 7.

The goal of this paper is to take RO from theory to application, by first arguing the preference of quadratic uncertainty set over box uncertainty set and by providing guidance in calibrating the two important elements of a quadratic uncertainty set: the uncertainty matrix as well as the level of uncertainty. It is important to note that the final objective of RO is to improve the MVO by reducing the sensitivity to inputs and by avoiding creating large arbitrage positions for highly correlated assets. From this perspective, the determination of the uncertainty set should be studied in the context of the optimization. Most research work in RO literature analyzes the uncertainty set as the confidence region of the estimated inputs. In this paper, we seek to enhance this purely probabilistic approach by treating the uncertainty set as an integrated part of the optimization, in the sense that it should satisfy the optimality condition of RO. Following this approach, we propose the optimal choices regarding the form of uncertainty set and the uncertainty matrix. Leveraging the fact that κ is also a parameter in the optimization function, we derive its upper bound limit and propose a rule of thumb for its calibration in terms of Sharpe ratios.

The paper is organized as follows: Section II.A studies two major forms of uncertainty set and demonstrate the superiority of quadratic uncertainty set over box uncertainty set by deriving and comparing the objective function of RO using both forms. We identify, in Section II.A, the origins of drawbacks of MVO by expressing its optimality condition in terms of risk budgets, Sharpe ratios and correlation matrix. We explain how RO, with the right parameters, can mitigate these drawbacks based on its optimality condition. Section II.B analyzes four uncertainty matrices to construct the quadratic uncertainty sets and advocates the use of a diagonal matrix of sample variances based on two criteria: reduction of sensitivity and keeping the original volatilities unchanged. Section II.C derives useful insights from the κ calibration using both analytical and empirical techniques for multi-asset portfolios. In particular, the upper bound limit of κ is obtained using the optimality condition of RO. Moreover, we propose a rule of thumb to calibrate κ based on the Sharpe ratios using simulations from returns of major asset classes. Finally, in section III, we provide two asset allocation examples to argue that RO reduces the sensitivity to inputs of the portfolio construction process, and leads to more intuitive and more diversified portfolios compared to MVO. Section III.A constructs an example from real historical data and aims to show that RO optimal portfolios are more diversified in risk and with less extreme long-short positions compared to the MVO. Section III.B details another example based on a simple case where assets have the same expected return, the same volatility, and correlation coefficients equal to zero. We introduce a minor estimation error in the expected return of Asset 1 and analyze its impact on the portfolios constructed by MVO and RO when the correlation coefficients between Asset 1 and Asset 2 varies from -99% to 99%.

II. Modelling of Uncertainty Set

The RO applied to portfolio construction can be reformulated by modifying the MVO through a max-min process (Scherer 2006): first, one finds, within the uncertainty set $\boldsymbol{U_\mu}$, the worst-case expected returns of assets. They are defined as the realized returns that deviate most negatively from the estimated expected returns $\widehat{\boldsymbol{\mu}}$. As discussed in the introduction, the drawbacks of the MVO exist regardless of the estimated expected returns used; we can assume, without loss of generality, that the expected returns are estimated by sample mean $\widehat{\boldsymbol{\mu}} = \bar{\boldsymbol{\mu}}$. Once the worst-case expected returns are obtained, the optimization process maximizes the portfolio returns, computed with the worst-case expected returns, under the risk constraint.

$$\max_{\boldsymbol{w}}\left(\min_{\boldsymbol{\mu}\in\boldsymbol{U_\mu}}\left(\boldsymbol{w}^T\boldsymbol{\mu}\right)-\frac{\lambda}{2}\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}\right)) \quad \text{(Equation 1.1)}$$

With $\boldsymbol{w}$ the $n \times 1$ vector of portfolio weights, $\boldsymbol{\mu}$ the $n \times 1$ vector of expected returns, λ the risk aversion parameter and $\boldsymbol{\Sigma}$ the $n \times n$ covariance matrix of asset returns.

Below, we advocate the use of a quadratic form with diagonal matrix of the sample variance of asset returns to define the uncertainty set in the robust portfolio optimization.

A. Form of Uncertainty Set

The discussion of RO's application in portfolio management revolves around the choice of uncertainty set and the calibration of its parameters. Fabozzi *et al.* (2007) presented two forms of uncertainty set: box uncertainty set and quadratic uncertainty set.

Box Uncertainty Set

The box uncertainty set $U_{\boldsymbol{\xi}}(\bar{\boldsymbol{\mu}})$ is the simplest way to express the uncertainty in inputs. Assume that there are n assets in the investment universe. The expected return $\boldsymbol{\mu}$ is estimated by the sample mean $\bar{\boldsymbol{\mu}}$. The uncertainty is assumed smaller than a constant vector $\boldsymbol{\xi} \geq \boldsymbol{0}$. Expressing this uncertainty set in mathematical form yields:

$$\boldsymbol{U_\mu} = \{\boldsymbol{\mu} \mid |\mu_i - \bar{\mu}_i| \leq \xi_i,\ i = 1,2,\dots,n\} \quad \text{(Equation 1.2)}$$

Note that the box uncertainty set, in fact, models the estimation error in the expected return of each asset separately. In other words, each asset's expected return has its individual confidence intervals around its own average.

Following Fabozzi *et al.* (2007) and Heckel *et al.* (2016), we derive the robust portfolio optimization problem with box uncertainty set:

$$\max_{\boldsymbol{w}}\left(\min_{\boldsymbol{\mu}\in\boldsymbol{U_\mu}}\left(\boldsymbol{w}^T\boldsymbol{\mu}\right)-\frac{\lambda}{2}\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}\right), \boldsymbol{U_\mu} = \{\boldsymbol{\mu} \mid |\mu_i - \bar{\mu}_i| \leq \xi_i,\ i = 1,2,\dots,n\} \quad \text{(Equation 1.3)}$$

Equation 1.3 can be reformulated into a robust version of MVO by minimizing the worst case portfolio return $\boldsymbol{w}^T\boldsymbol{\mu}$ and then putting it back into the MVO optimization.

Minimizing $\boldsymbol{w}^T\boldsymbol{\mu}$ for $\boldsymbol{\mu}$ within the uncertainty set defined by $|\mu_i - \bar{\mu}_i| \leq \xi_i,\ i = 1,2,\dots,n$ is equivalent to maximize $\boldsymbol{w}^T\bar{\boldsymbol{\mu}} - \boldsymbol{w}^T\boldsymbol{\mu}$ for $\boldsymbol{\mu}$ within the uncertainty set. The equivalence is validated by the fact that $\min_{\boldsymbol{\mu}\in\boldsymbol{U}_{\boldsymbol{\mu}}}(\boldsymbol{w}^T\boldsymbol{\mu}) = \max_{\boldsymbol{\mu}\in\boldsymbol{U}_{\boldsymbol{\mu}}}(-\boldsymbol{w}^T\boldsymbol{\mu})$. Note that the worst-case portfolio return would be obtained when $\mu_i - \bar{\mu}_i = -\xi_i, \forall\ i = 1,\dots,n$. Let $\kappa = \sum_{i=1}^{n}\xi_i$,

$$\boldsymbol{w}^T\bar{\boldsymbol{\mu}} - \boldsymbol{w}^T\boldsymbol{\mu} \leq \sum_{i=1}^{n}|\mu_i - \bar{\mu}_i|\max(|w_i|) \leq \kappa\max(|w_i|) \qquad \text{(Equation 1.4)}$$

The worst-case portfolio return $\boldsymbol{w}^T\boldsymbol{\mu}$ is achieved when the above inequality is saturated and $\boldsymbol{w}^T\boldsymbol{\mu} = \boldsymbol{w}^T\bar{\boldsymbol{\mu}} - \kappa\max(|w_i|)$. Inputting the $\boldsymbol{w}^T\bar{\boldsymbol{\mu}} - \kappa\max(|w_i|)$ into the objective function yields:

$$\max_{\boldsymbol{w}}\left(\boldsymbol{w}^T\bar{\boldsymbol{\mu}} - \kappa\max(|w_i|) - \frac{\lambda}{2}\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}\right) \qquad \text{(Equation 1.5)}$$

Quadratic Uncertainty Set

The quadratic uncertainty set takes a step further by including the uncertainty matrix, $\boldsymbol{\Omega}_{\boldsymbol{\mu}}$. It is assumed that the expected returns $\boldsymbol{\mu}$ are normally distributed with mean vector $\bar{\boldsymbol{\mu}}$**.** Hence, the uncertainty $\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}$ follow a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix of uncertainty in mean return $\boldsymbol{\Omega}_{\boldsymbol{\mu}}$. The uncertainty set around the estimated mean return can be written as follows:

$$\boldsymbol{U}_{\boldsymbol{\mu}} = \left\{\boldsymbol{\mu} \mid (\boldsymbol{\mu} - \bar{\boldsymbol{\mu}})^T\boldsymbol{\Omega}_{\boldsymbol{\mu}}^{-1}(\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}) \leq \kappa^2\right\} \qquad \text{(Equation 1.6)}$$

Here, the constant scalar κ^2 represents the level of uncertainty. The above uncertainty set covers all possible expected returns $\boldsymbol{\mu}$ within the level of uncertainty κ^2. Provided with this formulation, one can define the expected returns that deviate most negatively from the estimated returns within a certain level of uncertainty.

In the case of a quadratic uncertainty set, the robust portfolio optimization can be formulated as follows:

$$\max_{\boldsymbol{w}}\left(\min_{\boldsymbol{\mu}\in\boldsymbol{U}_{\boldsymbol{\mu}}}(\boldsymbol{w}^T\boldsymbol{\mu}) - \frac{\lambda}{2}\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}\right), \boldsymbol{U}_{\boldsymbol{\mu}} = (\boldsymbol{\mu} - \bar{\boldsymbol{\mu}})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}) \leq \kappa^2 \qquad \text{(Equation 1.7)}$$

The solution to the RO can be obtained in two steps. The first step involves finding the worst-case realized returns from the confidence region derived from the uncertainty set. Minimizing $\boldsymbol{w}^T\boldsymbol{\mu}$ for $\boldsymbol{\mu}$ within the uncertainty set defined by $\boldsymbol{U}_{\boldsymbol{\mu}}$ is equivalent to maximizing $\boldsymbol{w}^T\bar{\boldsymbol{\mu}} - \boldsymbol{w}^T\boldsymbol{\mu}$ for $\boldsymbol{\mu}$ within the uncertainty set:

$$\max_{\boldsymbol{\mu}}(\boldsymbol{w}^T\bar{\boldsymbol{\mu}} - \boldsymbol{w}^T\boldsymbol{\mu}) \quad s.t.\ (\boldsymbol{\mu} - \bar{\boldsymbol{\mu}})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}) \leq \kappa^2 \qquad \text{(Equation 1.8)}$$

Rewriting Equation 1.8 with the Lagrangian:

$$\underset{\boldsymbol{\mu}}{\mathcal{L}}(\bar{\boldsymbol{\mu}}, \delta) = \boldsymbol{w}^T\bar{\boldsymbol{\mu}} - \boldsymbol{w}^T\boldsymbol{\mu} - \delta((\boldsymbol{\mu} - \bar{\boldsymbol{\mu}})^T\boldsymbol{\Omega}^{-1}(\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}) - \kappa^2) \qquad \text{(Equation 1.9)}$$

Solving Equation 1.9 yields:

$$\boldsymbol{\mu} = \bar{\boldsymbol{\mu}} - \sqrt{\frac{\kappa^2}{\boldsymbol{w}^T\boldsymbol{\Omega}\boldsymbol{w}}}\,\boldsymbol{\Omega}\boldsymbol{w} \qquad \text{(Equation 1.10)}$$

Substituting the above formula for $\boldsymbol{\mu}$ into MVO transforms Equation 1.1:

$$\max_{\boldsymbol{w}} \left(\boldsymbol{w}^T\bar{\boldsymbol{\mu}} - \sqrt{\frac{\kappa^2}{\boldsymbol{w}^T\boldsymbol{\Omega}\boldsymbol{w}}}\,\boldsymbol{w}^T\boldsymbol{\Omega}\boldsymbol{w} - \frac{\lambda}{2}\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w} \right) \qquad \text{(Equation 1.11)}$$

We note the optimal robust portfolio weights as $\boldsymbol{w}^*_{rob}$:

$$\boldsymbol{w}^*_{rob} = \text{argmax}(\boldsymbol{w}^T\bar{\boldsymbol{\mu}} - \kappa\sqrt{\boldsymbol{w}^T\boldsymbol{\Omega}\boldsymbol{w}} - \frac{\lambda}{2}\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}) \qquad \text{(Equation 1.12)}$$

It is worth mentioning that the quadratic uncertainty set $U_\delta(\bar{\boldsymbol{\mu}})$ jointly models the errors of the expected returns of all assets. Depending on the choice of uncertainty matrix, different relationships between errors can be considered. If the uncertainty matrix is (proportional to) the covariance matrix, then, one implicitly assumes that the estimation errors among expected returns have the same correlation structure as the point estimates. If the uncertainty matrix is a diagonal matrix, then, the estimation errors are supposed to be uncorrelated.

Both box and quadratic uncertainty sets can be used to model the uncertainty set. However, Equation 1.5 illustrates the robust counterpart of MVO resulting from the box uncertainty set, and it is evident that the optimization penalizes only the mean return of the asset that has the largest absolute weight. This property is not desirable because the asset with the largest weight is not necessarily the asset that has the largest uncertainty in expected return estimation. Moreover, the presence of the absolute value operator makes Equation 1.5 not differentiable. On the other hand, the robust counterpart generated by the quadratic uncertainty set in Equation 1.11 provides a sounder representation of robustness because it penalizes the estimated returns of all assets jointly by taking into account the risks introduced by $\boldsymbol{\Omega}$. Thus, we prefer using quadratic form of uncertainty set.

There are other arguments, mentioned by authors in RO literature, for choosing the quadratic form of uncertainty set over the box uncertainty set:

1. Ben-Tal and Nemirovski (1998) review both box and quadratic uncertainty sets and argue that the latter leads to a tractable robust counterpart of the convex optimization problem while the robust counterpart induced by box uncertainty set is only tractable in linear programming.

2. Pachamanova and Fabozzi (2016) point out that the box uncertainty set assumes that all assets will achieve their worst-case return at the same time and this assumption is not verified in practice. They suggest that it may be more practical to assume that not all assets attain their worst-case returns at the same time and more informative to take into account the variance-covariance structure of the expected returns as formulated with a quadratic uncertainty set.
3. Goldfarb and Iyengar (2003) demonstrated analytically that the quadratic uncertainty set is generated naturally from the estimation process using regression when the expected returns are estimated with a linear factor model, which is quite common in the finance industry.

In conclusion, we advocate the use of quadratic uncertainty set in RO. In the remainder of the paper, we focus on the RO with quadratic uncertainty set.

RO with Quadratic Uncertainty Set and MVO

Now, let us take a closer look at the RO with a quadratic uncertainty set and the potential improvement it provides compared to MVO. At the optimum, the gradient of Equation 1.12 is equal to zero. It is important to note that the optimal robust weights $\boldsymbol{w}_{rob}^{*}$ satisfy the following equality when at least one of the optimal weights is different from zero:

$$\bar{\boldsymbol{\mu}} - \frac{\kappa}{\sqrt{\boldsymbol{w}_{rob}^{*T}\boldsymbol{\Omega}\boldsymbol{w}_{rob}^{*}}}\boldsymbol{\Omega}\boldsymbol{w}_{rob}^{*} - \lambda\boldsymbol{\Sigma}\boldsymbol{w}_{rob}^{*} = 0 \qquad \text{(Equation 1.13)}$$

Note that MVO optimal portfolio weights $\boldsymbol{w}_{MVO}^{*}$ can be obtained by setting to zero the derivative of $\boldsymbol{w}^{T}\bar{\boldsymbol{\mu}} - \frac{\lambda}{2}\boldsymbol{w}^{T}\boldsymbol{\Sigma}\boldsymbol{w}$, with respect to $\boldsymbol{w}$:

$$\bar{\boldsymbol{\mu}} - \lambda\boldsymbol{\Sigma}\boldsymbol{w}_{MVO}^{*} = 0 \qquad \text{(Equation 1.14)}$$

By rearranging the terms and inverting $\boldsymbol{\Sigma}$, we get:

$$\boldsymbol{w}_{MVO}^{*} = \frac{1}{\lambda}\boldsymbol{\Sigma}^{-1}\bar{\boldsymbol{\mu}} \quad \text{(Equation 1.15)}$$

Roncalli (2013) points out that the inversion of a covariance matrix $\boldsymbol{\Sigma}$ with small eigenvalues is the main cause of the high sensitivity to inputs and possible counter-intuitive long-short position suffered by MVO. However, in a covariance matrix $\boldsymbol{\Sigma}$, there are two elements: correlation coefficients and volatilities. Are they both responsible for creating small eigenvalues? To answer this question, note that optimizing on weights $\boldsymbol{w}$, $\boldsymbol{\Sigma}$ and $\bar{\boldsymbol{\mu}}$ is equivalent to optimizing on risk budgets $\boldsymbol{x} = \boldsymbol{\sigma} \times \boldsymbol{w}$, with $\boldsymbol{\sigma}$ the $n \times 1$ vector of volatilities of assets, $\times$ the element-wise multiplication operator and $\boldsymbol{x}$ the $n \times 1$ vector of risk budgets, $\boldsymbol{P} = \frac{\boldsymbol{\Sigma}}{\boldsymbol{\sigma}\boldsymbol{\sigma}^{T}}$ correlation matrix and $\overline{\boldsymbol{SR}} = \boldsymbol{\sigma}^{-1} \times \bar{\boldsymbol{\mu}}$ Sharpe ratios. The advantage of expressing MVO with risk budgets and correlation matrix is that the effect of correlation on the small eigenvalues of $\boldsymbol{\Sigma}$ is separated from that of volatilities. Equation 1.15 can be reformulated in terms of $\overline{\boldsymbol{SR}}$, $\boldsymbol{P}$ and $\boldsymbol{x}_{MVO}^{*}$, assuming $\lambda = 1$:

$$\boldsymbol{x}^*_{MVO} = \boldsymbol{P}^{-1}\overline{\boldsymbol{SR}} \qquad \text{(Equation 1.16)}$$

Equation 1.16 shows that the Sharpe ratios and the correlation matrix are two parameters that are responsible for the drawbacks of MVO mentioned earlier. In fact, MVO aims to exploit the differences in Sharpe ratios while taking into account the correlations among assets. Once the MVO optimal risk budgets are determined, the volatilities are there to yield the final portfolio weights. This last step is linear and does not involve any inversion of matrix.

Because **P** is symmetric and positive semi-definite, it can be decomposed into $\boldsymbol{P} = \mathbf{ZL}\boldsymbol{Z}^T$, with **Z** the matrix of eigenvectors of **P** and **L** the diagonal matrix with eigenvalues of **P** on the diagonal. Equation 1.16 can be transformed as:

$$\boldsymbol{x}^*_{MVO} = \boldsymbol{ZL}^{-1}\boldsymbol{Z}^T\overline{\boldsymbol{SR}} \qquad \text{(Equation 1.17)}$$

By expressing Equation 1.17 in the spaces spanned by the eigenvectors of **P**, we get:

$$\ddot{\boldsymbol{x}}^*_{MVO} = \boldsymbol{L}^{-1}\overline{\ddot{\boldsymbol{SR}}} \qquad \text{(Equation 1.18)}$$

With $\ddot{\boldsymbol{w}}_{\boldsymbol{MVO}}$, the optimal MVO weights expressed in the spaces spanned by the eigenvectors of $\boldsymbol{\Sigma}$ and $\overline{\ddot{\boldsymbol{SR}}}$ the expected returns of eigenvectors of $\boldsymbol{\Sigma}$. [2] The expected returns on eigenvectors are closely related to the expected returns of the assets.

Equation 1.18 shows two origins of the drawbacks of the MVO:

1. Inversion of small eigenvalues in $\boldsymbol{L}^{-1}$ which is the diagonal matrix of eigenvalues of correlation matrix **P**.

2. Non-negligible expected returns in $\overline{\ddot{\boldsymbol{SR}}}$ of the eigenvectors of **P** associated with small eigenvalues.

Once the origins of the drawbacks are determined, the ways that RO with a quadratic uncertainty set improves MVO become clearer. Equation 1.13 sheds light on the modification of the MVO optimality condition induced by the introduction of uncertainty in the objective function. There are two ways to interpret Equation 1.13 compared to Equation 1.14. These two interpretations represent two ways in which RO mitigates the drawbacks of the MVO.

Modification of $\boldsymbol{\Sigma}$: $$\bar{\boldsymbol{\mu}} - \lambda\left(\frac{\kappa}{\lambda\sqrt{\boldsymbol{w}^{*T}_{rob}\boldsymbol{\Omega}\boldsymbol{w}^*_{rob}}}\boldsymbol{\Omega} + \boldsymbol{\Sigma}\right)\boldsymbol{w}^*_{rob} = 0 \qquad \text{(Equation 1.19)}$$

[2] In view of the fact that eigenvectors correspond to the weights on the assets, each vector in the matrix of eigenvectors can be viewed as a portfolio formed from the original assets. Henceforth, when multiplying the eigenvectors by the Sharpe ratios, we can get the expected return of the eigenvectors. Recall that in a setting of correlation matrix and Sharpe ratios, the latter play the same role as expected returns in a setting of covariance matrix and expected returns.

By factoring $\boldsymbol{w}_{rob}^{*}$, Equation 1.13 illustrates the modification of covariance matrix when the uncertainty is introduced in the objective function. It is evident that the choice of the uncertainty matrix $\boldsymbol{\Omega}$ can have a huge impact on the final covariance matrix that will be inverted. As shown in the introduction as well as in Equation 1.14, the inverse of an ill-conditioned matrix is responsible for the instabilities in MVO. The robustness of the RO depends on the choice of $\boldsymbol{\Omega}$. Based on Equation 1.19, we analyze four major uncertainty matrices later in this subsection.

Modification of $\bar{\boldsymbol{\mu}}$: $\left(\bar{\boldsymbol{\mu}} - \frac{\kappa}{\sqrt{\boldsymbol{w}_{rob}^{*T}\boldsymbol{\Omega}\boldsymbol{w}_{rob}^{*}}}\boldsymbol{\Omega}\boldsymbol{w}_{rob}^{*} \right) - \lambda\boldsymbol{\Sigma}\boldsymbol{w}_{rob}^{*} = 0$ (Equation 1.20)

By grouping the first two terms on the left-hand side, Equation 1.13 represents the modification of expected returns by the uncertainty. In this formulation, the original covariance matrix is not modified, however, the $\bar{\boldsymbol{\mu}}$ are adjusted so that the expected returns of eigenvectors of $\boldsymbol{\Sigma}$ associated with small eigenvalues are neutralized. Inspired by Equation 1.20, we discuss the choice of κ in Section II.C.

B. Choice of Uncertainty Matrix in Quadratic Uncertainty Set

Equation 1.19 shows that the effectiveness of RO with a quadratic uncertainty set to improve MVO depends heavily on the uncertainty matrix. In the RO literature, four types of uncertainty matrices are proposed. On the one hand, Scherer (2006) used an uncertainty matrix for the estimation error $\boldsymbol{\Omega}$ that is proportional to the covariance matrix of asset returns $\boldsymbol{\Sigma}$. For the sake of simplicity, we analyze the case $\boldsymbol{\Omega} = \boldsymbol{\Sigma}$ in this subsection. This simplification can be done without loss of generality. On the other hand, Stubbs and Vance (2005) suggested that when working with uncertainty at the asset level, retaining only the diagonal part of $\boldsymbol{\Sigma}$, i.e., $\boldsymbol{\Omega} \propto \mathrm{diag}(\boldsymbol{\Sigma})$ is preferable, however they did not provide any justification. More recently, Heckel *et al.* (2016) studied the optimal portfolios yielded by two other uncertainty matrices: the identity matrix and the diagonal matrix of sample volatilities. All four uncertainty matrices will be examined in this subsection. The choice of the uncertainty matrix will be based on its effectiveness in mitigating the drawbacks of the MVO mentioned in the introduction: high sensitivity to inputs and possible counter-intuitive positions in the optimal portfolios.

The analysis of uncertainty matrix starts with Equation 1.19, since at the optimum $\sqrt{\boldsymbol{w}_{rob}^{*T}\boldsymbol{\Omega}\boldsymbol{w}_{rob}^{*}}$ is just a number. By noting $\frac{\kappa}{\sqrt{\boldsymbol{w}_{rob}^{*T}\boldsymbol{\Omega}\boldsymbol{w}_{rob}^{*}}}$ as β and $\frac{\beta}{\lambda+\beta}$ as η, we get:

$\bar{\boldsymbol{\mu}} = (\beta\boldsymbol{\Omega} + \lambda\boldsymbol{\Sigma})\boldsymbol{w}_{rob}^{*}$ (Equation 1.21)

$\frac{\bar{\boldsymbol{\mu}}}{(\lambda+\beta)} = (\eta\boldsymbol{\Omega} + (1-\eta)\boldsymbol{\Sigma})\boldsymbol{w}_{rob}^{*}$ (Equation 1.22)

Instead of inverting $\boldsymbol{\Sigma}$, the solution to RO with a quadratic uncertainty set requires the inversion of a modified covariance matrix $\eta\boldsymbol{\Omega} + (1-\eta)\boldsymbol{\Sigma}$.

CASE 1: $\boldsymbol{\Omega} = \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_1^2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \vdots & \ddots & \vdots \\ \rho_{1n}\sigma_1\sigma_n & \cdots & \sigma_n^2 \end{pmatrix}$

By replacing $\boldsymbol{\Omega}$ by $\boldsymbol{\Sigma}$ in the new modified covariance matrix, we get:

$\eta\boldsymbol{\Omega} + (1-\eta)\boldsymbol{\Sigma} = \boldsymbol{\Sigma}$ (Equation 1.23)

There is no change to the original covariance matrix. The RO, formulated with this uncertainty matrix, cannot mitigate the drawbacks of the MVO. Note that this uncertainty matrix was studied by Scherer (2006).

CASE 2: $\boldsymbol{\Omega} = \text{diag}(\boldsymbol{\Sigma}) = \begin{pmatrix} \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n^2 \end{pmatrix}$ with σ_1^2, σ_n^2 the variances of asset 1 and asset n

$$\eta\boldsymbol{\Omega} + (1-\eta)\boldsymbol{\Sigma} = \eta\begin{pmatrix} \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n^2 \end{pmatrix} + (1-\eta)\begin{pmatrix} \sigma_1^2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \vdots & \ddots & \vdots \\ \rho_{1n}\sigma_1\sigma_n & \cdots & \sigma_n^2 \end{pmatrix} \quad \text{(Equation 1.24)}$$

With ρ_{1n} the correlation coefficient between asset 1 and asset n.

The new covariance matrix, in Equation 1.22, is now a weighted average between the original covariance matrix and the diagonal matrix of sample variances.

CASE 3: $\boldsymbol{\Omega} = \boldsymbol{I}_n$, with $\boldsymbol{I}_n$ the $n \times n$ identity matrix

$$\eta\boldsymbol{\Omega} + (1-\eta)\boldsymbol{\Sigma} = \eta\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} + (1-\eta)\begin{pmatrix} \sigma_1^2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \vdots & \ddots & \vdots \\ \rho_{1n}\sigma_1\sigma_n & \cdots & \sigma_n^2 \end{pmatrix} \quad \text{(Equation 1.25)}$$

The new covariance matrix, in Equation 1.22, is a weighted average between the original covariance matrix and the identity matrix.

CASE 4: $\boldsymbol{\Omega} = \text{sqrt}\big(\text{diag}(\boldsymbol{\Sigma})\big) = \begin{pmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n \end{pmatrix}$

$$\eta\boldsymbol{\Omega} + (1-\eta)\boldsymbol{\Sigma} = \eta\begin{pmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n \end{pmatrix} + (1-\eta)\begin{pmatrix} \sigma_1^2 & \cdots & \rho_{1n}\sigma_1\sigma_n \\ \vdots & \ddots & \vdots \\ \rho_{1n}\sigma_1\sigma_n & \cdots & \sigma_n^2 \end{pmatrix} \quad \text{(Equation 1.26)}$$

The new covariance matrix, in Equation 1.22, is a weighted average between the original covariance matrix and the diagonal matrix of sample volatilities.

Uncertainty Matrices Selection: Criteria

Equation 1.16 shows that the part of the covariance matrix that is responsible for the high sensitivity to inputs is the correlation matrix. Reducing the sensitivity to inputs consists of eliminating small eigenvalues from the correlation matrix according to Equation 1.18. Appendix A demonstrates the equivalence between eliminating small eigenvalues from the correlation matrix and shrinking the correlation coefficients towards zero.

In contrast to the correlation coefficients, according to Equation 1.16, the role of volatilities in the solution to MVO in terms of risk budgets is merely a scaling factor to determine the final portfolio weights. Therefore, the original volatilities are not accountable for the high sensitivity to inputs suffered by MVO. Moreover, Equation 1.22 shows that the expected returns of all assets are scaled by the same factors in all robust optimization settings; hence, RO does not change the relative magnitude of the expected returns. Therefore, if volatilities are unchanged, it can be guaranteed that the relative magnitude of the Sharpe ratios is preserved.

From the above remarks, we propose the following criteria to select the uncertainty matrix:

1. The ideal uncertainty matrix is expected to reduce the sensitivity to inputs by shrinking the original correlation coefficients towards zero.

2. The ideal uncertainty matrix should keep the original volatilities unchanged.

Uncertainty Matrices Selection: Result

From the reduction of sensitivity to inputs perspective, we note that on the one hand, case 1 with the uncertainty matrix equal to the covariance matrix provides no reduction of sensitivity to inputs of the optimization solution because the correlation coefficients are unmodified. On the other hand, cases 2, 3 and 4 all consist of a weighted average between an uncertainty matrix that is diagonal and covariance matrices. Given the fact that all the off-diagonal terms in the uncertainty matrices are zero in cases 2, 3 and 4, they all shrink the original correlation coefficients towards zero.

Regarding the second criterion, note that the diagonal terms of the uncertainty matrices in case 1 and 2 are equal to those of the original covariance matrix. Thus, the introduction of uncertainty matrices specified as in case 1 and case 2 leaves the diagonal terms of the new covariance matrix unchanged while the diagonal terms of both uncertainty matrices from case 3 and 4 are different from those of the original covariance matrix. Henceforth, using uncertainty matrices from case 3 and 4 would lead to changes in volatilities while the introduction of uncertainty matrices specified as in case 1 and 2 would preserve the original volatilities. Moreover, the changes in volatilities with case 3 and 4 are difficult to control as the weighting factor, which depends on $\boldsymbol{w}_{rob}^{*}$, is endogenous.

Table 1: Uncertainty Matrix Characteristics

	Case 1: $\boldsymbol{\Omega} = \boldsymbol{\Sigma}$	Case 2: $\boldsymbol{\Omega} =$ diag($\boldsymbol{\Sigma}$)	Case 3: $\boldsymbol{\Omega} = \mathbf{I}$	Case 4: $\boldsymbol{\Omega}$ = sqrt(diag($\boldsymbol{\Sigma}$))
Reducing Sensitivity	No	Yes	Yes	Yes
Preserving Volatilities	Yes	Yes	No	No

Note:
$\boldsymbol{\Omega}$ = uncertainty matrix
$\boldsymbol{\Sigma}$ = estimated variance covariance matrix
sqrt(diag($\boldsymbol{\Sigma}$)) = the diagonal matrix with volatilities on the main diagonal
$\mathbf{I}$ = Identity Matrix

Summarizing the above analyses, it is evident that the diagonal matrix with variances on the main diagonal is the best candidate for the uncertainty matrix. Unlike the covariance matrix, it is effective in shrinking the correlation coefficients towards zero, which a) reduces the condition number of the correlation matrix, b) eliminates the small eigenvalues and c) attenuates the sensitivity of the solution to inputs. Moreover, unlike the identity matrix or the diagonal matrix with the volatility on the main diagonal, the uncertainty set constructed from $\boldsymbol{\Omega} = \text{diag}(\boldsymbol{\Sigma})$ achieves the aforementioned amelioration without distorting the volatility structure of the risk model. Thus, we prefer the diagonal matrix with variances on the main diagonal as the uncertainty matrix, corresponding to Case 2 in Table 1.

C. Choice of Level of Uncertainty Parameter: κ

Analytical Framework

Once the choice of the uncertainty matrix is determined, we concentrate on the level of uncertainty parameter κ. From now on, we fix the diagonal matrix of sample variances as the uncertainty matrix. Bearing in mind that the diagonal matrix of sample variances does not change the volatilities of the new covariance matrix used in RO, we can analyze the optimization problem in terms of risk budget and correlation matrix.

Many RO literature contributors analyze κ solely from a probabilistic and statistical point of view and treat it as the size of the confidence region around the expected returns. For instance, Scherer (2006) argues that the $\kappa_{\alpha,n}^2 = \chi_n^2(1-\alpha)$, where $\chi_n^2(1-\alpha)$ is the inverse of a chi-squared cumulative distribution with n degrees of freedom. Fabozzi *et al.* (2007) propose that κ represents the level of the scaled deviation of realized returns from the forecasts against which one wish to be protected. However, the authors tend to ignore another aspect of κ, namely, that κ is a key parameter in an optimization problem and it should satisfy the first order condition at optimum. In this section, we investigate the manner by which to determine the right parameter of robustness. We also provide a **rule of thumb** for the calibration of κ in the multi-asset investment universe.

Upper Bound for κ

For the sake of notational convenience, we reformulate the RO stated in Equation 1.12 in terms of the estimated Sharpe ratios $\overline{\boldsymbol{SR}}$, risk budget $\boldsymbol{X}$ and correlation matrix $\mathbf{P}$. For simplicity and

without loss of generality, we assume that λ is equal to 1. We note the optimal robust risk budget as $\boldsymbol{X}_{rob}^{*}$:

$$\boldsymbol{X}_{rob}^{*} = \text{argmax}(\boldsymbol{X}^{T}\overline{\boldsymbol{SR}} - \kappa\sqrt{\boldsymbol{X}^{T}\boldsymbol{I}_{n}\boldsymbol{X}} - \frac{1}{2}\boldsymbol{X}^{T}\mathbf{P}\boldsymbol{X}) \quad \text{(Equation 1.27)}$$

Deriving the optimality condition:

$$\overline{\boldsymbol{SR}} - \left(\frac{\kappa}{\sqrt{\boldsymbol{X}_{rob}^{*}{}^{T}\boldsymbol{X}_{rob}^{*}}}\boldsymbol{I}_{n} + \mathbf{P}\right)\boldsymbol{X}_{rob}^{*} = 0 \quad \text{(Equation 1.28)}$$

In addition, rearranging it yields the following expression:

$$\overline{\boldsymbol{SR}} = \left(\frac{\kappa}{\sqrt{\boldsymbol{X}_{rob}^{*T}\boldsymbol{X}_{rob}^{*}}}\boldsymbol{I}_{n} + \mathbf{P}\right)\boldsymbol{X}_{rob}^{*} \quad \text{(Equation 1.29)}$$

The above formulation sheds light on the role of κ as the parameter that tackles the high sensitivity to inputs suffered by MVO. In fact, the greater κ is, the more $\frac{\kappa}{\sqrt{\boldsymbol{X}_{rob}^{*T}\boldsymbol{X}_{rob}^{*}}}\boldsymbol{I}_{n} + \mathbf{P}$ shifts towards $\boldsymbol{I}_{n}$. The shift of the modified correlation matrix towards $\boldsymbol{I}_{n}$ helps to reduce the high sensitivity caused by the small eigenvalues but the benefit does not come without cost: a large κ may distort completely the correlation structure that makes assets indistinguishable from a risk perspective.

Taking the L2-Norm on both sides' yields:

$$\overline{\boldsymbol{SR}}^{T}\overline{\boldsymbol{SR}} = \kappa^{2} + \boldsymbol{X}_{rob}^{*T}\mathbf{P}^{T}\mathbf{P}\boldsymbol{X}_{rob}^{*} + 2 \times \frac{\kappa}{\sqrt{\boldsymbol{X}_{rob}^{*T}\boldsymbol{X}_{rob}^{*}}}\boldsymbol{X}_{rob}^{*T}\mathbf{P}\boldsymbol{X}_{rob}^{*} \quad \text{(Equation 1.30)}$$

Note that the second term on the right-hand side $\boldsymbol{X}_{rob}^{*T}\mathbf{P}^{T}\mathbf{P}\boldsymbol{X}_{rob}^{*}$, as the square of the L2 Norm of $\mathbf{P}\boldsymbol{X}_{rob}^{*}$, is non-negative; the third term on the right-hand side is also non-negative because $\mathbf{P}$ is positive semi-definite, the L2 Norm of $\boldsymbol{X}_{rob}^{*}$ is non-negative and κ^{2} is always non-negative. Therefore, the following upper bound for κ holds:

$$\kappa \leq \sqrt{\overline{\boldsymbol{SR}}^{T}\overline{\boldsymbol{SR}}} \quad \text{(Equation 1.31)}$$

Note that if κ is set higher than the upper bound, the first order derivative of the optimization with respect to $\boldsymbol{X}$ will always be negative. Therefore, the solution to the RO will be no-investment, i.e., $\boldsymbol{X} = \mathbf{0}$.

Rule of Thumb for Calibrating κ in a Multi-Asset Investment Universe

The upper bound can help us narrow the range for κ, but it is still not sufficient to determine its suitable value. A more efficient approach can be related to the nature of the high sensitivity of an

MVO solution to small changes in inputs. Recall that the solution to an MVO, viewed on the basis defined by the eigenvectors of the correlation matrix and assuming λ is equal to 1, is given by: $\ddot{\boldsymbol{X}}_{\boldsymbol{MVO}} = \boldsymbol{L}^{-1}\ddot{\overline{\boldsymbol{SR}}}$. It is evident that the sensitivity of an MVO to inputs is caused by the small eigenvalues, which amplify the expected returns of the corresponding eigenvectors. From this perspective, both the uncertainty matrix and κ address the sensitivity to inputs but from different angles: the introduction of a diagonal matrix of sample variances as the uncertainty matrix leads to an attenuation of the dispersion of eigenvalues, while κ can essentially be viewed as reducing the expected returns of the eigenvectors corresponding to the small eigenvalues.

We first provide an analytical illustration of the way κ can be used to reduce the aforementioned expected returns. Then, we propose a rule of thumb for determining κ based on a practical multi-asset example, which indicates that, the choice of κ does not depend on the number of assets in the optimization universe. Note that we do not pretend to find an exact formula to calibrate κ, rather, we propose a rule of thumb that can be helpful in practical application of RO in a multi-asset portfolio.

As shown at the beginning of this section, RO is a max-min process. The objective is to maximize the objective function even under the worst return realization. Thus, the RO uses penalized returns ($\boldsymbol{\mu}$) instead of the traditional expected returns from the sample mean ($\bar{\boldsymbol{\mu}}$):

$$\boldsymbol{\mu} = \bar{\boldsymbol{\mu}} - \sqrt{\frac{\kappa^2}{\boldsymbol{w}^T\boldsymbol{\Omega}\boldsymbol{w}}}\,\boldsymbol{\Omega}\boldsymbol{w}, \text{ with } \boldsymbol{\Omega} = \text{diag}(\boldsymbol{\Sigma}) \quad \text{(Equation 1.32)}$$

Once again, re-expressing just the above equation in terms of the Sharpe ratios and risk budget yields the following expression:

$$\boldsymbol{SR} = \overline{\boldsymbol{SR}} - \frac{\kappa}{\|\boldsymbol{X}\|_2}\boldsymbol{X} \quad \text{(Equation 1.33)}$$

The expected returns on the eigenvectors can be found easily when we apply the L2 Norm of $\boldsymbol{SR}$ expressed in the spaces spanned by the eigenvectors of correlation matrix $\mathbf{P}$.

Consider the eigenvalues-eigenvectors decomposition of the correlation matrix $\mathbf{P} = \mathbf{ZLZ}^T$, with $\mathbf{Z}$ the matrix of the eigenvectors and $\mathbf{L}$ the diagonal matrix with the eigenvalues on the main diagonal, the expected returns on the eigenvectors $\ddot{\overline{\boldsymbol{SR}}}$ of $\mathbf{P}$ can be found with the upper limit of $\ddot{\boldsymbol{SR}}$:

$$\left\|\ddot{\boldsymbol{SR}}\right\|_2 = \sqrt{\ddot{\overline{\boldsymbol{SR}}}^T\ddot{\overline{\boldsymbol{SR}}} + \kappa^2 - 2 * \kappa\frac{\ddot{\overline{\boldsymbol{SR}}}^T\ddot{\boldsymbol{X}}}{\left\|\ddot{\boldsymbol{X}}\right\|_2}} \leq \sqrt{\ddot{\overline{\boldsymbol{SR}}}^T\ddot{\overline{\boldsymbol{SR}}} - \kappa^2} \quad \text{(Equation 1.34)}$$

With $\ddot{\boldsymbol{SR}} = \mathbf{Z}^T\boldsymbol{SR}$, $\ddot{\overline{\boldsymbol{SR}}} = \mathbf{Z}^T\overline{\boldsymbol{SR}}$ and $\ddot{\boldsymbol{X}} = \mathbf{Z}^T\boldsymbol{X}$, see Appendix B for details.

$$\left\|\ddot{\boldsymbol{SR}}\right\|_2 \leq \sqrt{(\boldsymbol{Z}^T\boldsymbol{SR})^T(\boldsymbol{Z}^T\boldsymbol{SR}) - \kappa^2} \quad \text{(Equation 1.35)}$$

Note that the right-hand side of Equation 1.35 can be rewritten in two ways:

$$\sqrt{(\boldsymbol{Z}^T\boldsymbol{SR})^T(\boldsymbol{Z}^T\boldsymbol{SR}) - \boldsymbol{\kappa^2}} = \sqrt{(\boldsymbol{Z}_1^T\boldsymbol{SR})^2 + (\boldsymbol{Z}_2^T\boldsymbol{SR})^2 + \cdots (\boldsymbol{Z}_n^T\boldsymbol{SR})^2 - \boldsymbol{\kappa^2}} \quad \text{(Equation 1.36)}$$

$$\sqrt{(\boldsymbol{Z}^T\boldsymbol{SR})^T(\boldsymbol{Z}^T\boldsymbol{SR}) - \boldsymbol{\kappa^2}} = \sqrt{\boldsymbol{SR}^T\boldsymbol{Z}\boldsymbol{Z}^T\boldsymbol{SR} - \boldsymbol{\kappa^2}} = \sqrt{\boldsymbol{SR}^T\boldsymbol{SR} - \boldsymbol{\kappa^2}} \quad \text{(Equation 1.37)}$$

With $\boldsymbol{Z}_1$ the eigenvector that corresponds to the largest eigenvalue. The vectors of $\boldsymbol{Z}$ are ordered following the order of eigenvalues (from the largest to smallest).

The key for calibrating κ is to make use of the equivalence between Equation 1.36 and Equation 1.37: when κ is calibrated in terms of the Sharpe ratios (Equation 1.37), it is able to reduce or even neutralize the cumulative sum of "returns" on eigenvectors that correspond to the small eigenvalues (Equation 1.36). Namely, $\kappa = \sqrt{(\boldsymbol{Z}_i^T\boldsymbol{SR})^2 + \cdots (\boldsymbol{Z}_n^T\boldsymbol{SR})^2}$ with i to n the indices that correspond to small eigenvalues. Small is a rather abstract term and it does not tell us how to choose the cut-off number i. In the empirical experiment below, we propose a rule of thumb to help us determine the cut-off number and thus to calibrate κ.

Multi-Asset Universe Example

Let us now proceed with a simulation in a multi-asset universe. The universe studied consists of 23 indices encompassing the major asset classes.

All data series are extracted from Bloomberg in net total returns and in local currency. The period covered runs from February 2003 to April 2019, with a monthly frequency. We carry out the simulation from a European investor perspective, so the net total returns in local currency of non-EUR assets are transformed into returns in EUR with the following formula: $(1 + R_{EUR}) = (1 + R_{Local})(1 + R_{FX})$, with R_{EUR} the return in EUR, R_{Local} the return in local currency of non-EUR asset and R_{FX} the return of exchange. To yield excess net total returns in EUR, we subtract EONIA from the net total return in EUR. The correlation matrix is then computed using the monthly excess net total returns in EUR. The list of asset names, the corresponding Bloomberg tickers, the correlation matrix as well as the long term Sharpe ratios can be found in Appendix C.

In order to show that the result of the calibration of κ in a multi-asset environment is almost invariant with respect to the size of the investment universe, we let the number of assets vary from 3 to 21. We exclude the cases where numbers of assets equal to 22 and 23 because the correlation matrices are practically the same for different simulations. For each number of assets, we simulate 1000 random combinations of assets and build 1000 different correlation matrices. We first assume that $\overline{\boldsymbol{SR}} = \mathbf{1}$, with $\mathbf{1} = (1, \ldots, 1)^T$ as the long term Sharpe ratios for major asset classes are quite close to each other except for cash. Later on, we will release this restriction by using the long term Sharpe ratios from Ilmanen (2011). For each number of assets, we carry out the following computation:

- Step1: For each correlation matrix simulated for this specific number of assets n, we compute the squared expected returns of the eigenvectors by assuming that all assets have a Sharpe ratio equal to 1. The squared expected returns of the eigenvectors are given by the following expression:$(\boldsymbol{Z}_1^T\mathbf{1})^2, \ldots, (\boldsymbol{Z}_n^T\mathbf{1})^2$, with $\boldsymbol{Z}_1^T$ the transpose of the eigenvector that corresponds to the largest eigenvalue, $\boldsymbol{Z}_n^T$ the transpose of the eigenvector that corresponds to the smallest eigenvalue and $\mathbf{1}$ the vector of 1.
- Step 2: We calculate the cumulative sum of the squared expected returns on the eigenvectors from the one that corresponds to the smallest eigenvalue to the one that corresponds to the largest eigenvalue: $(\boldsymbol{Z}_n^T\mathbf{1})^2, (\boldsymbol{Z}_n^T\mathbf{1})^2 + (\boldsymbol{Z}_{n-1}^T\mathbf{1})^2, \ldots, (\boldsymbol{Z}_n^T\mathbf{1})^2 + (\boldsymbol{Z}_{n-1}^T\mathbf{1})^2 + \cdots + (\boldsymbol{Z}_1^T\mathbf{1})^2$. For instance, for an investment universe of five assets, there are five cumulative sums. These cumulative sums are calculated for all the 1000 correlation matrices simulated.
- Step 3: We take the average of the squared root of each cumulative sum of squared returns of eigenvectors computed with 1000 correlation matrices for a specific number of assets n $\frac{1}{1000}\sum_{j=1}^{1000}\sqrt{(\boldsymbol{Z}_n^T\mathbf{1})^2}_j$, $\frac{1}{1000}\sum_{j=1}^{1000}\sqrt{(\boldsymbol{Z}_n^T\mathbf{1})^2 + (\boldsymbol{Z}_{n-1}^T\mathbf{1})^2}_j, \ldots, \frac{1}{1000}\sum_{j=1}^{1000}\sqrt{(\boldsymbol{Z}_n^T\mathbf{1})^2 + \cdots + (\boldsymbol{Z}_1^T\mathbf{1})^2}_j$.

Figure 1: Calibration of κ in terms of Sharpe Ratio

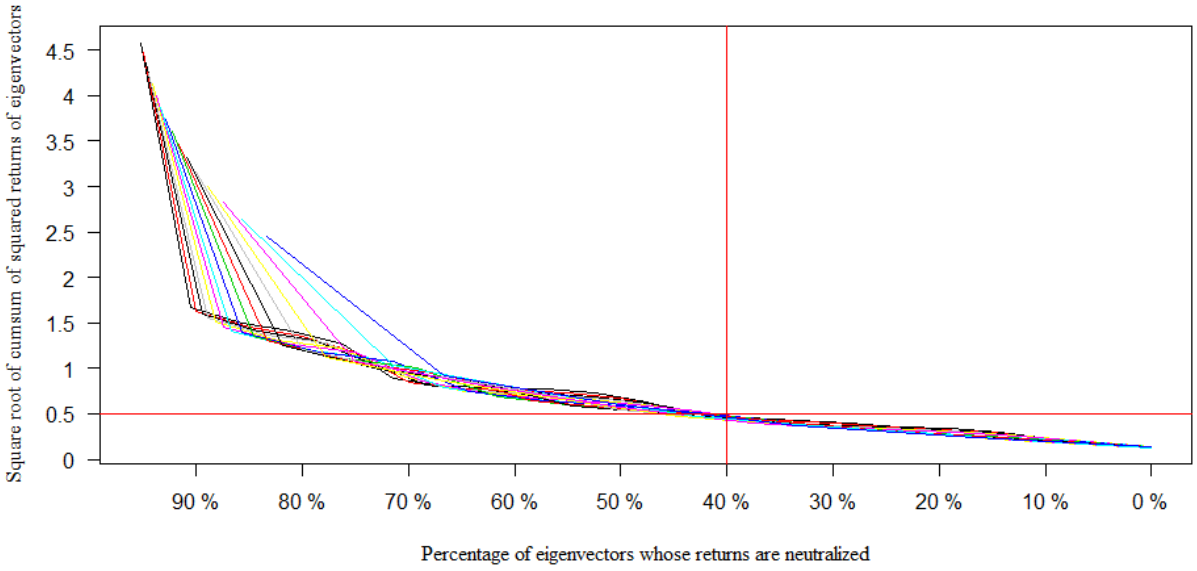


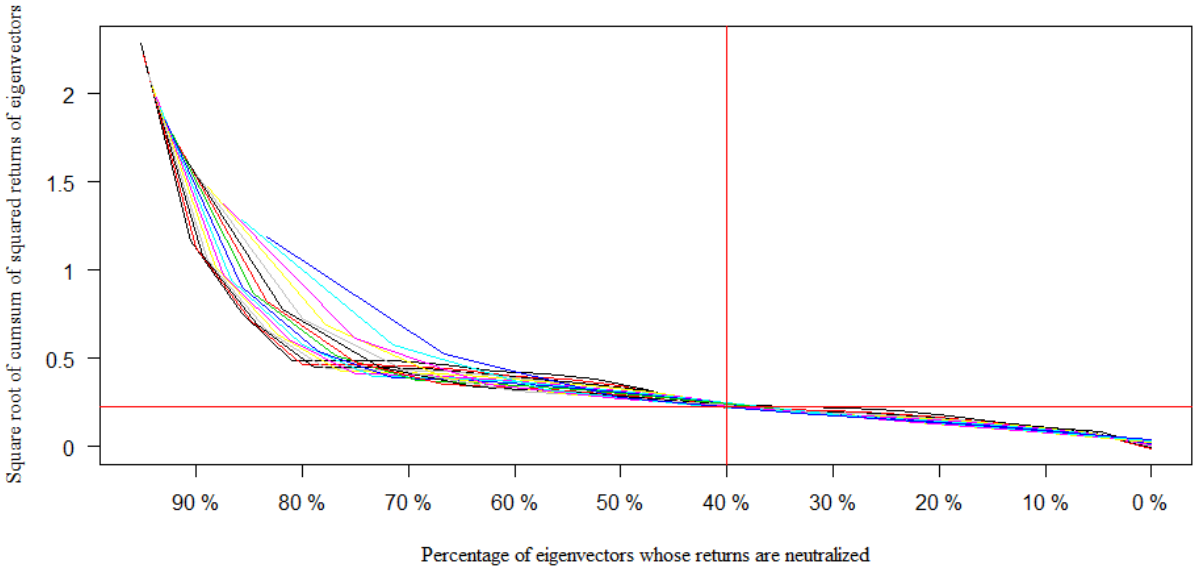
Figure 1 plots the squared root of cumulative sums of the squared returns of eigenvectors against the number of eigenvectors involved in each cumulative sum divided by the number of assets. Each colored line represents a specific number of assets, from 3 to 21. Viewed from the vertical axis, the data points on each of 18 colored lines represent $\frac{1}{1000}\sum_{j=1}^{1000}\sqrt{(\boldsymbol{Z}_n^T\mathbf{1})^2}_j$, $\frac{1}{1000}\sum_{j=1}^{1000}\sqrt{(\boldsymbol{Z}_n^T\mathbf{1})^2 + (\boldsymbol{Z}_{n-1}^T\mathbf{1})^2}_j, \ldots, \frac{1}{1000}\sum_{j=1}^{1000}\sqrt{(\boldsymbol{Z}_n^T\mathbf{1})^2 + \cdots + (\boldsymbol{Z}_1^T\mathbf{1})^2}_j$ for the investment universe of a specific number of assets. On the horizontal axis, data points for each colored line correspond to the percentage of eigenvectors involved in the cumulative sum.

Figure 1 illustrates that when κ is chosen to be half of the average of Sharpe ratios, which equals 0.5; it corresponds to the squared root of cumulative sums of the squared returns of eigenvectors that correspond to 40% smallest eigenvalues. We conclude that when κ is equal to half of the average of Sharpe ratios, it allows the neutralization of the returns of eigenvectors that correspond to 40% smallest eigenvalues; this conclusion is valid regardless of the number of assets included in the investment universe. The cut-off threshold is chosen at 40% because, according to Figure 1, 40% is the highest percentage we can choose while keeping the rule of thumb valid for different numbers of assets.

Robustness Check of Rule of Thumb

We obtained our rule of thumb of κ calibration by assuming all assets have the same Sharpe ratio. To examine whether our rule of thumb still applies if assets had different Sharpe ratios, we conduct a simple robustness check of our rule of thumb with the long term Sharpe ratios proposed by Ilmanen (2011). We calculate accordingly the squared expected returns of the eigenvectors in Step 1: $(\boldsymbol{Z}_1^T \boldsymbol{SR})^2, \ldots, (\boldsymbol{Z}_n^T \boldsymbol{SR})^2$, with $\boldsymbol{SR}$ the vector of long term Sharpe ratios extracted from Ilmanen (2011) and detailed in Appendix C. The average of the Sharpe ratios is equal to 0.46 now. Again, we find our rule of thumb: when κ is equal to half of the average of Sharpe ratios (0.23), it allows the neutralization of the returns of eigenvectors that correspond to 40% smallest eigenvalues. This conclusion remains valid regardless of the number of assets included in the investment universe.

Figure 2: Calibration of κ Robustness Check



To summarize our analysis, the proposed rule of thumb consists of choosing κ as half of the average of Sharpe ratios. This rule of thumb applies for multi-asset portfolios regardless of the number of assets they comprise and regardless of the assumptions on Sharpe ratios.

III. Asset Allocation Examples

In Section II.B, we mentioned that there are two ways to analyze the first order optimality condition of RO given by Equation 1.13. RO can be viewed as an MVO applied to a modified covariance matrix according to Equation 1.19 or as an MVO applied to modified expected returns following Equation 1.20 in Section II.B. RO, formulated with the uncertainty sets and the level of uncertainty we advocate in Section II, leads to portfolios that are well diversified in risks and that have less extreme long-short weight. Moreover, the high sensitivity to inputs suffered by MVO is reduced in RO too. In this section, we provide two examples to illustrate these claims.

A. RO and MVO: Practical Asset Allocation

The investment universe consists of four assets: US Equity, US Small Cap, US Sovereign and US Investment Grade (IG). The indexes used, in this example, can be found in Appendix C. We assume an identical Sharpe ratio that equals 0.46, the average of long term Sharpe ratio, for all assets. We prefer to use identical Sharpe ratio to show that MVO leads to large arbitrage positions in similar assets even when they have the same Sharpe ratio. The volatilities are equal to 19.14% for US Equity, 23.70% for US Small Cap, 9.89% for US Sovereign and 10.24% for US IG. They are estimated using monthly net total excess returns in EUR, already mentioned in Section II.C. The correlation matrix here is extracted from the original correlation matrix used in Section II.C.

In this example, we assume that κ is equal to 0.23, which corresponds to half of the average of Sharpe ratios. There are neither minimum or maximum weights constraints, nor a full investment constraint. For all portfolio optimizations, we impose the constraint that the volatility of optimal portfolio cannot exceed 10%, which facilitates comparison among optimization results.

The RO is performed with the uncertainty matrix equal to the diagonal of covariance matrix as proposed in Section II.B. We present also the results of the MVO in Table 3. Both portfolio weights and contributions to risk (CTR) are presented. The contributions to risk for an asset is given by the following formula: $CTR_i = w_i \frac{\boldsymbol{\Sigma w}}{\sqrt{\boldsymbol{w^T \Sigma w}}}$, with w_i the weight of asset i, $\boldsymbol{\Sigma}$ the covariance matrix and $\boldsymbol{w}$ the vector of portfolio weights.

Table 2: Portfolio weights of different optimizations

	Volatility	MVO		RO with $\boldsymbol{\Omega}$ = diag($\boldsymbol{\Sigma}$)	
		Weights	CTR	Weights	CTR
US Equity	19.14%	10.14%	1.45%	14.90%	2.32%
US Small Cap	23.70%	23.82%	4.23%	15.53%	2.75%
US Sovereign	9.89%	110.11%	8.15%	37.73%	2.77%
US IG	10.24%	-49.95%	-3.83%	25.24%	2.16%

Note: $\boldsymbol{\Omega}$ = uncertainty matrix
diag($\boldsymbol{\Sigma}$) = Diagonal of variance covariance matrix
CTR = Contribution to risks

Table 2 displays that MVO generates a very counter-intuitive portfolio with large long-short positions in US Sovereign and US IG. The long-short position in two similar assets is particularly undesirable when one wants to define the strategic asset allocation. In contrast to MVO, RO yields a portfolio with no large long-short arbitrage positions, even without any constraints on weights. The CTR calculations shed more light on the comparison between a robust optimal portfolio and an MVO portfolio. In RO, the CTR are shrinking towards an equal risk budget portfolio, which is intuitive since the correlation matrix is shrinking to zero and the SR are the same for all assets. As shown by Leote de Carvalho et al (2012), equal risk budget is optimal in the limit of correlations coefficients converge to zero and SR are the same for all assets.

In Tables 3, 4, 5 and 6, we present the original covariance matrix and correlation matrix as well as their modified versions in RO. Note that the formula of the modified covariance matrices is given by Equation 1.19. The RO with uncertainty matrix equal to the diagonal of covariance matrix shrinks the correlation coefficients towards zero while keeping the volatilities unchanged. We provide a concrete example to support our choice of uncertainty matrix in Section II.B.

Table 3: Original correlation matrix

Correlation	US Equity	US Small Cap	US Sovereign	US IG
US Equity	100.00%	87.00%	26.00%	43.00%
US Small Cap	87.00%	100.00%	15.00%	29.00%
US Sovereign	26.00%	15.00%	100.00%	93.00%
US IG	43.00%	29.00%	93.00%	100.00%

Table 4: New correlation matrix used in RO with $\boldsymbol{\Omega} = \text{diag}(\boldsymbol{\Sigma})$

Correlation	US Equity	US Small Cap	US Sovereign	US IG
US Equity	100.00%	48.30%	14.44%	23.87%
US Small Cap	48.30%	100.00%	8.33%	16.10%
US Sovereign	14.44%	8.33%	100.00%	51.63%
US IG	23.87%	16.10%	51.63%	100.00%

Table 5: Original covariance matrix

	US Equity	US Small Cap	US Sovereign	US IG
US Equity	3.66%	3.95%	0.49%	0.84%
US Small Cap	3.95%	5.62%	0.35%	0.70%
US Sovereign	0.49%	0.35%	0.98%	0.94%
US IG	0.84%	0.70%	0.94%	1.05%

Table 6: New covariance matrix used in RO with $\boldsymbol{\Omega} = \text{diag}(\boldsymbol{\Sigma})$

	US Equity	US Small Cap	US Sovereign	US IG
US Equity	3.66%	2.19%	0.27%	0.47%
US Small Cap	2.19%	5.62%	0.20%	0.39%
US Sovereign	0.27%	0.20%	0.98%	0.52%
US IG	0.47%	0.39%	0.52%	1.05%

Table 7 illustrates that RO can reduce the high dispersion in eigenvalues created by the correlation coefficients. The high dispersion in eigenvalues means that small eigenvalues exist for the covariance matrix. The dispersion in eigenvalues is measured by the condition number, the details of which can be found in Appendix A. Belsley *et al.* (1980) state that a condition number higher than 5 indicates the presence of the collinearity in the covariance matrix, which would create high sensitivity to inputs of the regression result. Like MVO, linear regression requires also the inversion of a covariance matrix to get the result. Here, in the case of the original covariance matrix, the largest condition number is higher than 10. The extreme weights given by the MVO are due to an inversion of a covariance matrix suffering from collinearity. As shown in Table 7, RO, with the previously mentioned uncertainty matrix configuration, brings the condition number under the threshold and reduces the sensitivity to inputs.

Table 7: Eigenvalues and condition numbers of different covariance matrices

	Original $\boldsymbol{\Sigma}$	New $\boldsymbol{\Sigma}$ used in RO with $\boldsymbol{\Omega}$ = diag($\boldsymbol{\Sigma}$)
First Eigenvalue	8.92%	7.12%
Second Eigenvalue	1.82%	2.30%
Third Eigenvalue	0.52%	1.41%
Fourth Eigenvalue	0.05%	0.48%
Condition Number 1	12.93	3.84
Condition Number 2	4.16	2.25
Condition Number 3	2.21	1.76

Note: Condition Number 1 = sqrt(Biggest Eigenvalue / Smallest Eigenvalue)
Condition Number 2 = sqrt(Biggest Eigenvalue / Second Smallest Eigenvalue)
Condition Number 3 = sqrt(Biggest Eigenvalue / Third Smallest Eigenvalue)

Table 8, 9 and Table 10 illustrate that RO reduces the sensitivity to inputs by neutralizing the expected returns given to the eigenvectors that correspond to the small eigenvalues.

In Table 8, we show the modified expected returns caused by RO with an uncertainty matrix equal to a diagonal matrix of sample variance as well as the original expected returns used in MVO. The formula used to compute the modified expected returns comes from Equation 1.20. The expected returns detailed in Table 8 are used to calculate the expected returns on eigenvectors illustrated in Table 10.

Table 8: Original Expected Returns in MVO and Modified Expected Returns in RO

Expected Returns	MVO	RO with $\boldsymbol{\Omega}$ = diag($\boldsymbol{\Sigma}$)
US Equity	8.80%	6.87%
US Small Cap	10.90%	7.82%
US Sovereign	4.55%	3.24%
US IG	4.71%	3.77%

Note: $\boldsymbol{\Omega}$ = uncertainty matrix
diag($\boldsymbol{\Sigma}$) = Diagonal of variance covariance matrix

Table 9 displays the eigenvectors of the original covariance matrix. The first eigenvector represents the common trend in the four assets; it can be interpreted as short market/equity risk factor. The second eigenvector has also a clear meaning; it represents short duration risk factor. The third and the last eigenvectors are just arbitrage portfolios without any meaningful interpretation. In particular, the last eigenvector, which corresponds to the smallest eigenvalue, consists of a large long-short portfolio in US Sovereign and US IG. The latter is responsible for the large arbitrage positions in US Sovereign and US IG in the MVO optimal portfolio.

Table 9: Eigenvectors of original covariance matrices

	First Eigenvector	Second Eigenvector	Third Eigenvector	Fourth Eigenvector
US Equity	-61.19%	-7.16%	78.15%	-9.85%
US Small Cap	-77.23%	26.30%	-57.78%	2.24%
US Sovereign	-8.94%	-68.59%	-21.95%	-68.80%
US IG	-14.53%	-67.47%	-8.50%	71.86%

Table 10 illustrates the expected returns of eigenvectors of $\mathbf{\Sigma}$. With $\kappa = 0.23$, RO preserves the expected returns for the first two eigenvectors, which are not just the noises, as much as possible while significantly shrinking the expected return of the last two eigenvectors; in fact, the expected return of the last eigenvector in RO is only 5%, in absolute term, of that used in MVO. From previous results regarding the eigenvalues of $\mathbf{\Sigma}$, the smallest eigenvalues, which are just noises, are the most harmful ones in the inversion of $\mathbf{\Sigma}$. By shrinking significantly the returns on the eigenvectors (here, the last two ones) corresponding to the smallest eigenvalues, RO reduces the sensitivity to inputs and limits the creation of arbitrage positions in highly correlated assets. The fact that only the expected returns of the last two eigenvalues are reduced significantly is not surprising. We mentioned, in Section II.C, when κ equals half of the average of Sharpe ratios, RO reduces the returns of 40% the smallest eigenvectors. Here, they correspond to the last two eigenvectors.

Table 10: Expected Returns on the Eigenvectors of Original Covariance Matrix

	MVO	RO with $\mathbf{\Omega}$ = diag($\mathbf{\Sigma}$)	Expected Return RO / Expected ReturnMVO
First Eigenvector	-14.90%	-11.08%	74.38%
Second Eigenvector	-4.06%	-3.21%	78.98%
Third Eigenvector	-0.82%	-0.18%	21.69%
Fourth Eigenvector	-0.37%	-0.02%	5.75%

Note: $\mathbf{\Omega}$ = uncertainty matrix
diag($\mathbf{\Sigma}$) = Diagonal of variance covariance matrix

B. RO and MVO: Sensitivity to Input Changes

The objective of this second example is to show that RO attenuates the sensitivity to inputs compared to MVO. This example is based on three fictive assets. We assume that all three assets have the same Sharpe ratio ($SR = 0.46$), the same volatility (15%) and the same correlation ($\rho_{12} = 0, \rho_{13} = 0, \rho_{23} = 0$). We perform the MVO and the RO with diagonal matrix of sample variance as the uncertainty matrix, subject to a full investment constraint. There is no constraint on maximum volatility and the aversion to risk λ is assumed equal to 1. In this example, we assume that κ is equal to 0.23, which corresponds to half of the average of Sharpe ratios.

Table 11 exhibits the optimal portfolios of MVO and RO with the above inputs and settings. Note that both RO and MVO give rise to an equally weighted portfolio with 33.33% for each asset. This result is not surprising. When all assets are identical from both risk and return perspectives and

their correlation is 0, then, the MVO optimal portfolio coincides with the risk-based portfolio. The latter in this case is an equally weighted portfolio (Leote de Carvalho *et al.* 2012).

Table 11: MVO and RO for 3 Assets with identical risk-return characteristics

	MVO	RO
Asset 1	33.33%	33.33%
Asset 2	33.33%	33.33%
Asset 3	33.33%	33.33%

To assess the sensitivity to inputs of both MVO and RO, we increase the expected return of asset 1 by 0.15% (which corresponds to an increase of 0.01 in Sharpe ratio) and we vary ρ_{12}, the correlation coefficient between asset 1 and asset 2 from -99% to 99%. The change in expected return is made deliberately small to show that MVO could create large arbitrage portfolio even with very small discrepancy in expected returns. In Appendix D, we provide the result of the same exercise when the expected return of asset 1 is increased by 1.5%, which corresponds to an increase of 0.1 in its Sharpe ratio.

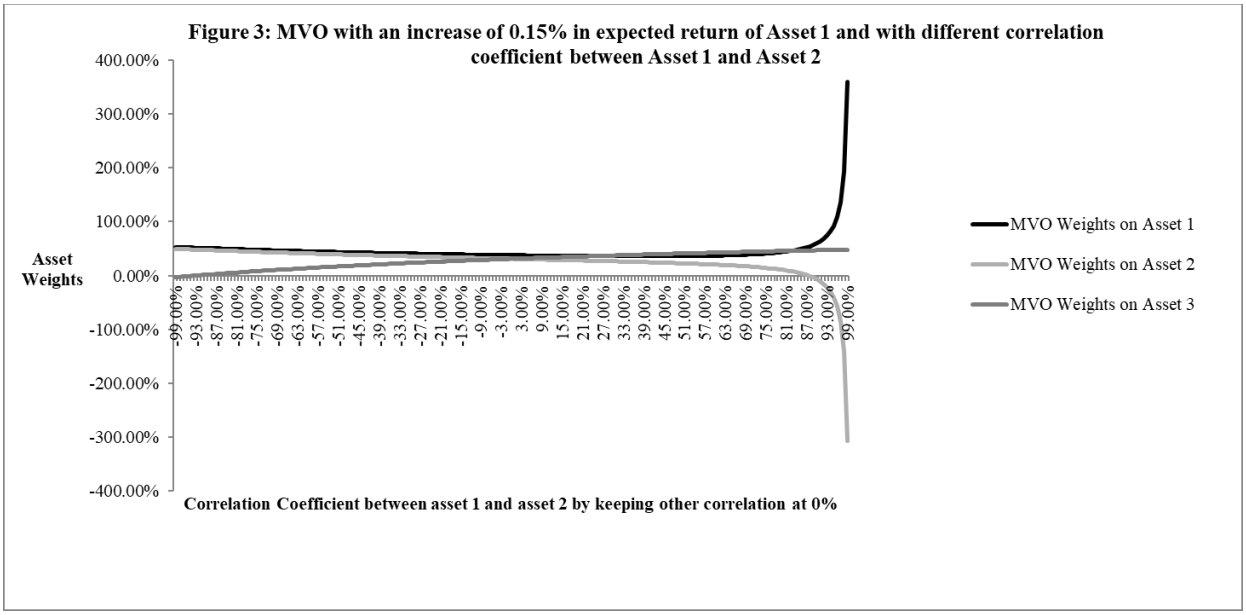


Figure 3 shows that the MVO optimal portfolio is highly sensitive to changes in expected returns and correlation coefficients. Recall that the increase in expected return of Asset 1 is only 0.15%, which corresponds to an increase of 0.01 in its Sharpe ratio with a volatility of 15%. This minor change is sufficient to create arbitrage positions in Asset 1 and Asset 2 when their correlation coefficient becomes larger than 80%. A correlation coefficient higher than 80% is quite common within equity or fixed income assets.

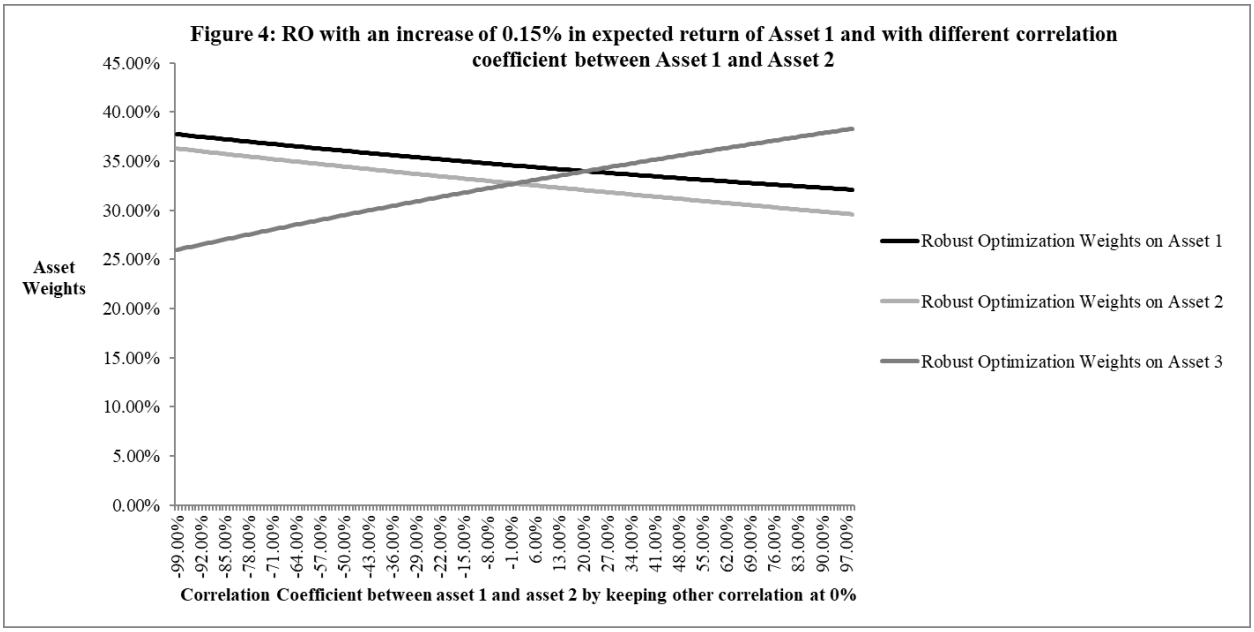


Figure 4 illustrates that RO, with a diagonal matrix of sample variance as the uncertainty matrix, is less sensitive to the estimation errors in the expected returns compared to the MVO process, even when the correlation coefficient becomes extremely high. There is no large long-short arbitrage position created between Asset 1 and Asset 2. From this example, we also show that RO, with a diagonal matrix of sample variance as the uncertainty matrix, is more suitable in defining the strategic asset allocation than MVO.

IV. Conclusion

In this paper, we propose a new approach to calibrating the three important elements of a RO uncertainty set: the form of uncertainty set, the uncertainty matrix as well as the level of uncertainty in quadratic uncertainty set. Previously, RO literature tends to treat them as standalone parameters. The form of uncertainty set and the uncertainty matrix are chosen rather arbitrarily. The level of uncertainty is determined from a solely probabilistic or statistical point of view. In this paper, we consider the choice of these three elements as an integrated part of the optimization process. To our knowledge, we are the first to use such a calibration philosophy in the RO literature.

In the first subsection of section II, We discuss the choice of the form of uncertainty set by deriving the robust counterparts of MVO for both box and quadratic uncertainty sets and advocate for the use of quadratic uncertainty set. Then, we show that there are two sources of the sensitivity of MVO to inputs: the inversion of small eigenvalues and the non-negligible expected returns given to the eigenvectors associated with these small eigenvalues. From the two formulations of the optimality condition of RO, we propose two ways that RO, with carefully chosen uncertainty matrices as the uncertainty set and the level of uncertainty, could overcome the shortcomings of MVO, namely, eliminating small eigenvalues and reducing significantly the expected returns given to the associated eigenvectors. Next, we review four main uncertainty matrices proposed in the

RO literature. We find that the diagonal matrix of sample variances provides the best trade-off between reduction of the sensitivity of an MVO solution and keeping volatilities unchanged. We derive a relationship between the L2 Norm of the Sharpe ratios and the level of uncertainty, κ, from the optimality condition of RO. We yield the upper bound limit of κ relying on this relationship. We obtain a rule of thumb for the calibration of κ in terms of the average of Sharpe ratios in a multi-asset investment universe including major asset classes. In fact, if κ is set to half of the average of Sharpe ratios, it can neutralize the returns of eigenvectors that correspond to 40% smallest eigenvalues. We use examples with the proposed parametrization to show that robust optimization efficiently overcomes the weaknesses of mean-variance optimisation and can be applied in real investment problems like multi-asset portfolio management or robo-advising.

Thus, we bring RO from theory to application; we provide guidance in the determination of the characteristics of a RO uncertainty set in a multi-asset environment, which allows RO to be applied directly in real-life portfolio construction problems. The scope of application for RO is extended considerably with the rise of robo-advisors in the finance industry. Indeed, to build automated robo-advisors, one needs a portfolio optimization algorithm that creates, without human intervention, well-balanced portfolios that are not highly sensitive to changes in expected return forecasts. From what we argued in this article, RO is clearly a better candidate for endorsing the role of portfolio optimizer in a robo-advisor than MVO.

Disclaimer

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Reference

Black, F., and Litterman, R., 1990, Asset allocation: combining investor views with market equilibrium, *Goldman Sachs Fixed Income Research.*

Belsley, D. A., Kuh, E., and Welsch, R. E., 1980, Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. *Wiley*.

Ben-Tal, A., and Nemirovski, A., 1998, Robust convex optimization. *Math. Oper. Res*. 23, 769–805.

Best, M., and Grauer R., 1991, On the Sensitivity of Mean Variance Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results, *Review of Financial Studies 4*, pp. 314–42.

Bruder, B., Gaussel, N., Richard, J-C., and Roncalli, T., 2013, Regularization of Portfolio Allocation, *SSRN, www.ssrn.com/abstract=2767358*.

Ceria, S., and Stubbs, R. A., 2006. Incorporating estimation errors into portfolio selection: robust portfolio construction. *Journal of Asset Management* 7(2), 109–127.

Chopra, V. and Ziemba, W. T., 1993, The Effects of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice, *Journal of Portfolio Management* (*Winter)* 6–11.

Cornuéjols, G., PEÑA, J., and Tütüncü, R.H., 2018, Optimization Methods in Finance. *Cambridge University Press*.

Fabozzi, F.J., Kolm P.N., Pachamanova D.A., and Focardi S.M., 2007, Robust Portfolio Optimization. *The Journal of Portfolio Management, Vol. 33, No. 3 (Spring)*, pp. 40-48.

Goldfarb, D., and Iyengar, G., 2003, Robust Portfolio Selection Problems. *Mathematics of Operations Research* 28(1), 1-38.

He G., and Litterman, R., 1999, The Intuition Behind Black-Litterman Model Portfolios, *Goldman Sachs Investment Management Research.*

Heckel, T., Leote de Carvalho, R., Lu, X., and Perchet, R., 2016, Insights into robust optimization: decomposing into mean-variance and risk-based portfolios, *Journals of Investment Strategies* 6(1), 1-24.

Ilmanen A., 2011, Expected Returns an Investor's Guide to Harvesting Market Rewards, *John Wiley & Sons, Inc.*

Jobson, J., and Korkie B., 1983, Statistical Inference in Two Parameter Portfolio Theory with Multiple Regression Software, *Journal of Financial and Quantitative Analysis 18*, 189–97.

Kallberg, J. G., and Ziemba W.T., 1984, Mis-specification in Portfolio Selection Problems, *Risk and Capital: Lecture Notes in Economics and Mathematical Systems (Springer)*.

Leote de Carvalho, R., Lu, X., and Moulin, P., 2012, Demystifying equity risk-based strategies: a simple alpha plus beta description. *Journal of Portfolio Management* 38(3), 56–70.

Markowitz, H., 1952, Portfolio selection, *Journal of Finance* 7 (1), 77-91.

Markowitz, H. M., 1959, Portfolio Selection: Efficient Diversification of Investments, *Yale University Press, New Haven, CT*.

Michaud, R., 1989, The Markowitz Optimization Enigma: Is Optimization Optimal? *Financial Analysts Journal* 45(1), 31-42.

Pachamanova D. A., and Fabozzi F. J., 2016, Portfolio Construction and Analytics, *John Wiley & Sons, Inc.*

Roncalli, T., 2013, Introduction to Risk Parity and Budgeting, *Chapman & Hall/CRC Financial Mathematics Series.*

Santos, A. A. P., 2010, The out-of-sample performance of robust portfolio optimization. *Brazilian Review of Finance*, 8(2), 141–166.

Scherer, B., 2006, Can robust portfolio optimization help to build better portfolios? *Journal of Asset Management*, Vol. 7, No. 6, 374-387.

Scherer, B., 2010, Portfolio Construction and Risk Budgeting, *Risk Books*.

Stubbs R. and Vance P., 2005, Computing Return Estimation Error Matrices for Robust Optimization*, Report Axioma*.

Tütüncü, R.H., and König M., 2004, Robust Asset Allocation. *Annals of Operations Research*, *Vol. 132, No. 1-4*, 157-187.

APPENDIX A

In this appendix, we show that eliminating the small eigenvalues is equivalent to reducing the Frobenius norm of the correlation matrix. Reducing the Frobenius norm is equivalent to shrinking the correlation coefficients towards zero.

To understand this, note that it is not the absolute magnitude but the relative magnitude of these eigenvalues that would determine the stability of the solution and its sensibility with respect to the estimated Sharpe ratios and correlation matrix. The relative magnitude of eigenvalues is captured by the $condition\ number = \sqrt{\frac{\beta_{max}}{\beta_{min}}}$ (Belsley, Kuh, and Welsch 1980) with β_{max} being the maximum eigenvalue of **P** and β_{min} being the minimum eigenvalue, or by the variance of eigenvalues. The more dispersed the eigenvalues, the more sensitive to the inputs the MVO solution is.

In the case of two assets, the link between eigenvalues and the correlation coefficient is straightforwardly given by the following formula:

$\beta = 1 \pm \rho$ (Equation 1.36)

With β being the eigenvalues and ρ being the correlation coefficient. Note that when $|\rho| \longrightarrow 1$, the condition number$\longrightarrow +\infty$, the variance of eigenvalues $\longrightarrow 2$ (the maximum) and the solution becomes less stable or even unsolvable. For $n \geq 5$, calculating eigenvalues directly requires solving a fifth degree polynomial. The Abel–Ruffini theorem states that there is no algebraic expression for a general fifth degree polynomial; therefore, there is no analytical solution for the eigenvalues. However, the Frobenius norm could be computed to get to the bottom of the

relationship between correlation coefficients and the relative magnitude of eigenvalues for a matrix of dimension higher than 5.

In fact, the Frobenius norm of a correlation matrix P of $n \times n$ dimension for n assets is given by the following formula:

$$\|\boldsymbol{P}\|_F = \sqrt{trace(\boldsymbol{P}^T\boldsymbol{P})} = \sqrt{n + 2 * \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\rho_{ij}^2} \quad \text{(Equation 1.37)}$$

With ρ_{ij} being the correlation coefficient between asset i and asset j. Given that ρ_{ij} ranges from -1 to 1, then the Frobenius norm of a correlation matrix attains its maximum n when $\rho_{ij} = \pm 1$ and its minimum $\sqrt{n}$ when $\rho_{ij} = 0$, for all $i, j = 1, \cdots, n \; and \; i \neq j$. So, the Frobenius norm of a correlation matrix with $n * n$ dimension ranges from $\sqrt{n}$ to n. Eigenvalue decomposition provides another expression for the Frobenius norm:

$$\|\boldsymbol{P}\|_F = \sqrt{trace(\boldsymbol{P}^T\boldsymbol{P})} = \sqrt{\sum_{i=1}^{n}\beta_i^2} \quad \text{(Equation 1.38)}$$

With β_i being eigenvalues.

Recall that the variance of the eigenvalues is given by: $var(eigenvalues) = \frac{1}{n}(\sum_{i=1}^{n}\beta_i^2 - n)$ and the squared L2 Norm of the correlation coefficient is given by $\|\rho\|_2^2 = 2 * \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\rho_{ij}^2$ and from the equality between the two formulas for the Frobenius norm, we have:

$$\sqrt{n + 2 * \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\rho_{ij}^2} = \sqrt{\sum_{i=1}^{n}\beta_i^2} \quad \text{(Equation 1.39)}$$

By squaring both sides of Equation 1.39, we get:

$$\sum_{i=1}^{n}\beta_i^2 - n = 2 * \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\rho_{ij}^2 \quad \text{(Equation 1.40)}$$

Note that $var(\beta) = \frac{1}{n}\sum_{i=1}^{n}\beta_i^2 - n$ and $\|\rho\|_2^2 = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\rho_{ij}^2$, Equation 1.40 is actually:

$$var(\beta) = \frac{2}{n}\|\rho\|_2^2 \quad \text{(Equation 1.41)}$$

When $\rho_{ij} = \pm 1$ for all $i, j = 1, \cdots, n \; and \; i \neq j$, the Frobenius norm attains its maximum and this corresponds to the largest relative magnitude among all the eigenvalues. In this case, the eigenvalues would be n for the first one and 0 for the rest of $n - 1$ eigenvalues when they are arranged from the largest to the smallest, the condition number for this correlation structure is infinite and the variance of eigenvalues also attains its maximum: $n - 1$. This correlation structure is the worst case for mean-variance optimization given that there is not even a solution. The more ρ_{ij} approaches ± 1 for all $i, j = 1, \cdots, n \; and \; i \neq j$, the more dispersed the eigenvalues, measured by condition number and variance, are.

When $\rho_{ij} = 0$ for all$i, j = 1, \cdots, n \; and \; i \neq j$, the Frobenius norm attains its minimum and this corresponds to the smallest possible condition number, which equals 1, and to the lowest possible variance of eigenvalues which equals 0. In this case, the eigenvalues would all be equal to 1. This correlation structure is the best case for mean-variance optimization in terms of stability.

In this way, we demonstrate the equivalence among improving stability, reducing the Frobenius norm and shrinking correlation coefficients towards zero.

APPENDIX B

In this appendix, we derive the upper limit of the L2 Norm of Sharpe ratios used in the RO in terms of the "returns" on the eigenvectors of the correlation matrix. Recall that the Sharpe ratio used for the RO is written as follows:

$$\boldsymbol{SR} = \overline{\boldsymbol{SR}} - \frac{\kappa}{\|\boldsymbol{X}\|_2}\boldsymbol{X} \quad \text{(Equation 1.42)}$$

Applying the change of basis according to the coordinate defined by the eigenvectors $\boldsymbol{Z}$ of the correlation matrix **P**, we get:

$$\boldsymbol{Z}^T\boldsymbol{SR} = \boldsymbol{Z}^T\overline{\boldsymbol{SR}} - \frac{\kappa}{\|\boldsymbol{Z}^T\boldsymbol{X}\|_2}\boldsymbol{Z}^T\boldsymbol{X} \quad \text{(Equation 1.43)}$$

$$\ddot{\boldsymbol{SR}} = \overline{\ddot{\boldsymbol{SR}}} - \kappa\frac{\ddot{\boldsymbol{X}}}{\|\ddot{\boldsymbol{X}}\|_2} \quad \text{(Equation 1.44)}$$

Note that $\boldsymbol{Z}^T\boldsymbol{SR}$ can be viewed as the "return" of the eigenvectors. Again, the exact solution for κ is not feasible because the above equation involves $\ddot{\boldsymbol{X}}$, which is the solution of the RO itself. However, it is important to note that κ should be chosen so that the "returns" of the eigenvectors that correspond to the small eigenvalues could be reduced. Following this guideline, we consider the two terms on the right-hand side of the above equation separately by taking the L2 Norm.

$$\left\|\overline{\ddot{\boldsymbol{SR}}}\right\|_2 = \sqrt{\overline{\ddot{\boldsymbol{SR}}}^T\overline{\ddot{\boldsymbol{SR}}}} = \sqrt{\overline{\ddot{SR}}_1^2 + \overline{\ddot{SR}}_2^2 + \ldots + \overline{\ddot{SR}}_n^2} \quad \text{(Equation 1.44)}$$

The $\overline{\ddot{SR}}_i^2, for\ i = 1 \cdots n,$ follow the order of eigenvalues, for instance, the last $\overline{\ddot{SR}}_n^2$ corresponds to the "return" of the eigenvector that corresponds to the smallest eigenvalue.

$$\left\|\kappa\frac{\ddot{\boldsymbol{X}}}{\|\ddot{\boldsymbol{X}}\|_2}\right\|_2 = \kappa \quad \text{(Equation 1.45)}$$

We take the L2 Norm on both side of equation 1.45 as follows:

$$\|\ddot{\boldsymbol{SR}}\|_2 = \left\|\overline{\ddot{\boldsymbol{SR}}} - \kappa\frac{\ddot{\boldsymbol{X}}}{\|\ddot{\boldsymbol{X}}\|_2}\right\|_2 = \sqrt{\overline{\ddot{\boldsymbol{SR}}}^T\overline{\ddot{\boldsymbol{SR}}} + \kappa^2 - 2 * \kappa\frac{\overline{\ddot{\boldsymbol{SR}}}^T\ddot{\boldsymbol{X}}}{\|\ddot{\boldsymbol{X}}\|_2}} \quad \text{(Equation 1.46)}$$

Recall that the optimality condition of the RO mentioned earlier is written as:

$$\left(\frac{\kappa}{\sqrt{\boldsymbol{X}^T\boldsymbol{X}}}\,\boldsymbol{I}_n + \mathbf{P}\right)\boldsymbol{X} = \overline{\boldsymbol{SR}} \quad \text{(Equation 1.47)}$$

Multiplying both sides of equation 1.47 by $\boldsymbol{X}^T$, we get

$$\kappa\sqrt{\boldsymbol{X}^T\boldsymbol{X}} + \boldsymbol{X}^T\mathbf{P}\boldsymbol{X} = \boldsymbol{X}^T\overline{\boldsymbol{SR}} \quad \text{(Equation 1.48)}$$

Expressing both sides in the space spanned by the eigenvectors, our calculations could read as follows:

$$\kappa\sqrt{\ddot{\boldsymbol{X}}^T\ddot{\boldsymbol{X}}} + \ddot{\boldsymbol{X}}^T\mathbf{P}\ddot{\boldsymbol{X}} = \ddot{\overline{\boldsymbol{SR}}}^{\boldsymbol{T}}\ddot{\boldsymbol{X}} \quad \text{(Equation 1.49)}$$

Diving both sides by $\sqrt{\ddot{\boldsymbol{X}}^T\ddot{\boldsymbol{X}}} = \left\|\ddot{\boldsymbol{X}}\right\|_2$, the equivalent to the previous equation is given by:

$$\kappa + \frac{\ddot{\boldsymbol{X}}^T\mathbf{P}\ddot{\boldsymbol{X}}}{\left\|\ddot{\boldsymbol{X}}\right\|_2} = \frac{\ddot{\overline{\boldsymbol{SR}}}^{\boldsymbol{T}}\ddot{\boldsymbol{X}}}{\left\|\ddot{\boldsymbol{X}}\right\|_2} \quad \text{(Equation 1.50)}$$

As $\ddot{\boldsymbol{X}}^T\mathbf{P}\ddot{\boldsymbol{X}} \geq 0$ and $\left\|\ddot{\boldsymbol{X}}\right\|_2 > 0$, the optimality condition gives rise to the following inequality:

$$\kappa \leq \frac{\ddot{\overline{\boldsymbol{SR}}}^{\boldsymbol{T}}\ddot{\boldsymbol{X}}}{\left\|\ddot{\boldsymbol{X}}\right\|_2} \quad \text{(Equation 1.51)}$$

The inequality just derived can be used to yield an upper limit of $\left\|\ddot{\boldsymbol{SR}}\right\|_2$:

$$\left\|\ddot{\boldsymbol{SR}}\right\|_2 = \sqrt{\ddot{\overline{\boldsymbol{SR}}}^{\boldsymbol{T}}\ddot{\overline{\boldsymbol{SR}}} + \kappa^{\mathbf{2}} - 2 * \kappa\frac{\ddot{\overline{\boldsymbol{SR}}}^{\boldsymbol{T}}\ddot{\boldsymbol{X}}}{\left\|\ddot{\boldsymbol{X}}\right\|_2}} \leq \sqrt{\ddot{\overline{\boldsymbol{SR}}}^{\boldsymbol{T}}\ddot{\overline{\boldsymbol{SR}}} - \kappa^{\mathbf{2}}} \quad \text{(Equation 1.52)}$$

End of proof.

APPENDIX C

The asset classes used in the simulation application for calibrating κ, as well as their Bloomberg tickers and their long term Sharpe ratios.

To stay as objective as possible, we prefer using Ilmanen ((2012), page 25, 44 and 48) as the source of long term Sharpe ratio. When the data for Non-US assets do not exist, we take the Sharpe ratio of the equivalent US assets. The Bloomberg tickers chosen are also in line with Ilmanen (2012).

Table 12: Indices Used for Calibrating κ and Robustness Test

Asset Class Name	Bloomberg Ticker	Long Term Sharpe Ratio
Equity Europe EMU	NDDLEURO Index	0.39
Equity Europe EMU Small Cap	NCLDEMU Index	0.39
Equity Europe UK	NDDLUK Index	0.29
Equity North America USA	NDDUUS Index	0.37
Equity North America USA Small Cap	RU20INTR Index	0.37
Equity Pacific Japan	NDDLJN Index	0.28
Equity Emerging Global	NDUEEGF Index	0.39
Bond EUR Sovereign	LEATTREU Index	0.66
Bond EUR Inflation Linked	BCEE1T Index	0.66
Bond EUR Investment Grade	LECPTREU Index	0.84
Bond EUR High Yield	LF88TREU Index	0.50
Bond USD Sovereign	LUATTRUU Index	0.66
Bond USD Inflation Linked	BCIT1T Index	0.66
Bond USD Investment Grade	LUACTRUU Index	0.84
Bond USD High Yield	LF89TRUU Index	0.50
Bond JPY Sovereign	G0Y0 Index	0.61
Bond Emerging Market Hard Currency Sovereign Global	JPGCCOMP Index	0.58
Bond Emerging Market Local Currency Sovereign Global	JGENVUUG Index	0.58
Real Estate Pan Europe	TRNHUE Index	0.36
Real Estate USA	TRNUSU Index	0.36
Real Estate Asia Pacific	TRNHPU Index	0.36
Commodity Global	BCOMXAL Index	0.13
Cash EUR	DBDCONIA Index	0.00

Correlation matrix used in the simulation application for calibrating κ. The correlation matrix is estimated with Pearson estimator using the full sample net total returns in EUR from 2003 to 2019.

Table 13: Correlation Matrix for Calibrating κ and Robustness Test

Equity Europe EMU	100%																						
Equity Europe EMU Small Cap	88%	100%																					
Equity Europe UK	81%	75%	100%																				
Equity North America USA	64%	56%	70%	100%																			
Equity North America USA Small Cap	57%	57%	60%	87%	100%																		
Equity Pacific Japan	44%	45%	54%	59%	52%	100%																	
Equity Emerging Global	66%	69%	67%	57%	53%	54%	100%																
Bond EUR Sovereign	-5%	-6%	-5%	-5%	-13%	3%	-2%	100%															
Bond EUR Inflation Linked	18%	16%	16%	7%	2%	14%	21%	78%	100%														
Bond EUR Investment Grade	26%	29%	35%	24%	13%	24%	31%	68%	70%	100%													
Bond EUR High Yield	70%	71%	64%	42%	39%	33%	54%	5%	30%	56%	100%												
Bond USD Sovereign	-21%	-23%	-3%	26%	15%	26%	-7%	35%	11%	16%	-30%	100%											
Bond USD Inflation Linked	-15%	-16%	5%	29%	18%	31%	5%	36%	25%	27%	-20%	93%	100%										
Bond USD Investment Grade	-1%	-1%	19%	43%	29%	38%	13%	37%	21%	39%	0%	93%	91%	100%									
Bond USD High Yield	32%	30%	48%	70%	59%	51%	37%	7%	11%	35%	39%	62%	64%	80%	100%								
Bond JPY Sovereign	-39%	-33%	-25%	-5%	-11%	6%	-16%	32%	11%	10%	-38%	69%	67%	59%	28%	100%							
Bond Emerging Market Hard Currency Sovereign Global	20%	17%	36%	51%	37%	45%	37%	36%	32%	47%	25%	71%	75%	85%	83%	41%	100%						
Bond Emerging Market Local Currency Sovereign Global	33%	33%	43%	42%	33%	42%	65%	24%	30%	39%	32%	26%	34%	40%	49%	15%	67%	100%					
Real Estate Pan Europe	64%	68%	66%	48%	47%	38%	51%	23%	39%	46%	58%	-1%	7%	17%	35%	-13%	34%	37%	100%				
Real Estate USA	34%	31%	41%	63%	63%	39%	37%	22%	29%	33%	32%	34%	40%	46%	57%	16%	55%	43%	57%	100%			
Real Estate Asia Pacific	49%	51%	59%	53%	46%	65%	75%	17%	27%	40%	47%	11%	22%	29%	43%	3%	47%	60%	55%	54%	100%		
Commodity Global	19%	22%	41%	26%	27%	25%	40%	-12%	9%	10%	17%	-1%	14%	9%	24%	-6%	16%	20%	15%	12%	29%	100%	
Cash EUR	12%	16%	6%	-1%	2%	6%	8%	-3%	4%	-7%	6%	-10%	-9%	-11%	-8%	-11%	-8%	-8%	15%	1%	7%	4%	100%

Appendix D

The results of example in III.B with an increase of 1.5% in expected return of Asset 1. From Figure 5, we observe that an increase of 1.5% in expected creates arbitrage positions in Asset 1 and Asset 2 when their correlation coefficient becomes larger than 20%. The threshold in correlation coefficient to create arbitrage positions is likely to be reduced with larger estimation errors is reduced from 80% to 20% when the increase changes from 0.15% to 1.5%.

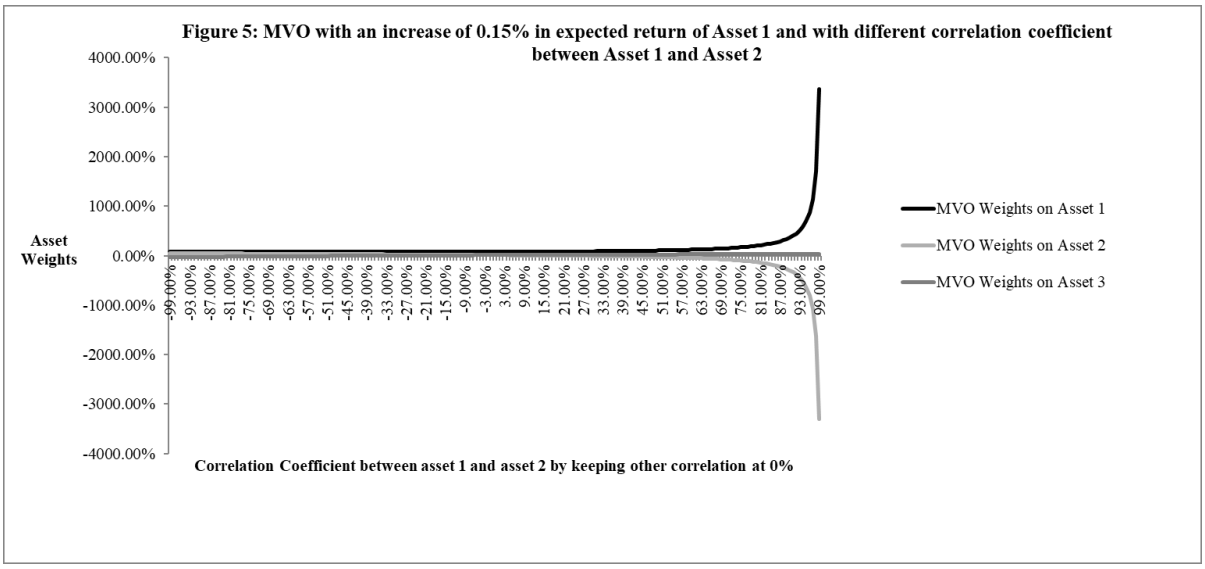


Figure 6 shows the result with RO. Note that due to the increase of 1.5% in expected return of Asset 1, RO optimal portfolios allocate more weights to asset 1. However, the weights are quite stable even when the correlation coefficients vary.