Papers

Can robust portfolio optimisation help to build better portfolios?

Received: 14th December, 2006

Bernd Scherer

is Managing Director and Head of Quantitative Structured Products at Morgan Stanley Investment Management in London.

Managing Director, Morgan Stanley, IM-Alternative Investments 20 Bank Street, Canary Wharf, E14 4QW, London. Tel: +44 20 7425-4016; Fax: +44 20 7425-8763; E-mail: Bernd.Scherer@morganstanley.com

Abstract Estimation error has always been acknowledged as a substantial problem in portfolio construction. Various approaches exist that range from Bayesian methods with a very strong rooting in decision theory to practitioner-based heuristics with no rooting in decision theory at all as portfolio resampling. Robust optimisation is the latest attempt to address estimation error directly in the portfolio construction process. It will be shown that robust optimisation is equivalent to Bayesian shrinkage estimators and offer no marginal value relative to the former. The implied shrinkage that comes with robust optimisation is difficult to control. Consistent with the *ad hoc* treatment of uncertainty aversion in robust optimisation, it can be seen that out of sample performance largely depends on the appropriate choice of uncertainty aversion, with no guideline on how to calibrate this parameter or how to make it consistent with the more well-known risk aversion. *Journal of Asset Management* (2007) **7**, 374–387. doi:10.1057/palgrave.jam.2250049

Keywords: robust optimisation, Bayes, resampling, portfolio construction, estimation error

Introduction

Virtually all attempts to address estimation error in portfolio construction have been around the refinement of expected returns before they enter the portfolio construction process. The error maximising property of traditional portfolio optimisation (assets with positive estimation error are over-weighted, while assets with negative estimation error are under-weighted) has been felt as a major obstacle in achieving a more scientific approach to investing. Financial economists tried to control the variation in expected returns with some form of shrinkage to either equal returns (James-Stein approach) or implied market returns (Black-Litterman approach) in the hope to also control the variation in portfolio output (and hence to

arrive at less extreme and more stable solutions).

Success has been mixed. First, return estimates still show outlier dependency, whatever statistical method is used. Secondly, parameter ambiguity will always be present, even if we increase the amount of extra sample information. But this means that error maximisation still affects portfolio construction. Lately, engineers and operations research academics have become interested in the field of portfolio optimisation and suggested two variations to mainstream thinking. The first was the idea of robust statistics, which promotes the clever removal (or down-weighting) of what are thought to be extreme observations (outliers). While outliers are sometimes the only information we have got (eg, in hedge

fund returns, where one manager bets against extreme events), it has been broadly felt that outlier removal reduced portfolio risk, rather than increasing it, as we would expect in the face of model error. This runs against the intuition of most portfolio and risk managers. The second addition to mainstream finance has been robust optimisation. On an intuitive level, robust optimisation attempts at minimising the worst case return for a given confidence region (without confidence region the worst case return is always -100per cent) subject to the usual constraints. Practitioners feel this is a conservative and hence prudent form of portfolio construction, with estimation error directly built into the portfolio optimisation process. While in general this helps to dampen the error maximisation problem, we will show that what looks like an innovation can be written in terms of ordinary shrinkage estimation and that the efficient set (set of optimal portfolios across an efficient frontier) remains the same.

In the next section, we introduce the early work by Tüntücü and König (2004) on robust portfolio optimisation as we can use their fairly simple setting to address the strength and weakness that all robust methods share. The subsequent section continues start with a more sophisticated review of the robust portfolio optimisation set up, utilising the framework by Ceria and Stubbs (2005). We will then extend their model in the section In sample critique to formally prove that robust optimisation equals Bayesian shrinkage, where the weight given to the speculative portfolio (relative to the minimum variance portfolio that does not suffer from estimation error in expected returns) depends both on the number of observations and the required confidence level. The latter is difficult to consistently calibrate. If we for example assume estimation error aversion to be high, while at the same time risk aversion is low, we are likely to underperform out of sample. The reason for this is, that high estimation error aversion forces the investor into assets with little estimation and hence also little investment

risk, while at the same time his risk aversion demands aggressive portfolios. At a more philosophical level this shows how artificial the split between risk and uncertainty is.

All arguments in the fourth section are essentially in sample. We conclude the paper with a computational example in the final section to underline the previous points. Essentially we show that the optimality of robust optimisation critically depends on the complicated interplay between risk aversion and uncertainty aversion.

The Tüntücü and König (2004) approach

Suppose investors are ambiguous about the correct variance covariance matrix or the correct mean vector in a mean variancebased portfolio optimisation. Instead they have many possible candidates in mind. More precisely, there exists a set of mean vectors and covariance matrices $\mathbf{\mu} \in S_{\mu}, \Omega \in S_{\Omega}$ where S_{μ} is the set of all mean vectors and S_{Ω} is the set of all covariance matrices. All matrices are given equal importance no matter how unlikely they are in a probabilistic sense as it is assumed the decision maker cannot form probabilities. The optimisation problem now becomes

$$\max_{\mathbf{w}} \left(\min_{\boldsymbol{\mu} \in S_{\boldsymbol{\mu}}, \boldsymbol{\Omega} \in S_{\boldsymbol{\Omega}}} \mathbf{w}^{\mathrm{T}} \boldsymbol{\mu} - \frac{\lambda}{2} \mathbf{w}^{\mathrm{T}} \boldsymbol{\Omega} \mathbf{w} \right)$$
(1)

In (1) we want to maximise the worst case utility for all combinations of variance covariance matrices with respect to the portfolio weight vector \mathbf{w} . The idea is to provide 'good' solutions for all possible parameter realisations. We will see that this in reality means to be very pessimistic as the solution has to provide a 'good' outcome even if the worst parameter specification becomes true. Problems like this can be reformulated to fit traditional optimisation software.¹ Essentially we maximise the worst case utility. For a large number of securities and a large set of mean vectors and



covariance matrices, it becomes infeasible to solve (1). Tüntücü and König (2004) have, however, shown that under the assumption of a long only constraint (all asset must be held in non-negative quantities), Equation (1) can be replaced by

$$\max_{\mathbf{w}\geq 0} \mathbf{w}^{\mathrm{T}} \boldsymbol{\mu}_{l} - \frac{\lambda}{2} \mathbf{w}^{\mathrm{T}} \boldsymbol{\Omega}_{h} \mathbf{w}$$
(2)

where $\mathbf{\mu}_l$ is the worst case return vector and $\mathbf{\Omega}_h$ is the worst case covariance matrix. The reason we can readily identify the worst case inputs rests on the imposed long only constraint. For a long only position the worst case is a low expected return, so μ_l is the smallest element in S_{μ} , while the worst case for $\mathbf{\Omega}_h$ is little diversification so it must be the largest element in S_{Ω} . A high covariance for example would not be worst case for a long/ short position as this implies that short positions are risk reducing (hedging). The same is true for expected returns. Low expected returns would actually be best case for a short position as there is on average less to loose. Tüntücü and König (2004), therefore, use bootstrapping to elementwise construct $\boldsymbol{\mu}_l$ and $\boldsymbol{\Omega}_h$. For lets say 1,000 resamplings from the original inputs μ_0, Ω_0 , we get 1,000 mean vectors and covariance matrices.² We now look at the top left hand element (variance for asset 1) in the variance covariance matrix and select the 5 per cent largest entry across all 1,000 matrices. This procedure is repeated for all elements³ as well as for the mean vector. With respect to the later we select the lower 5 per cent entries. Note that as $\mathbf{\Omega}_h \geq \mathbf{\Omega}_0$ by construction it follows directly that $\mathbf{w}^{\mathrm{T}} \boldsymbol{\Omega}_{h} \mathbf{w} - \mathbf{w}^{\mathrm{T}} \boldsymbol{\Omega}_{0} \mathbf{w} =$ $\mathbf{w}^{\mathrm{T}}(\mathbf{\Omega}_{h}-\mathbf{\Omega}_{0})\mathbf{w}\geq 0$, that is $\mathbf{\Omega}_{h}$ is riskier. Also the dispersion of eigenvalues in Ω_h is much larger, that is larger fraction of variance is explained by a smaller number of factors. This should come as no surprise as the procedure to construct Ω_h created high covariances mimicking the presence of a dominating market factor.

How should we evaluate this framework? The main problem in the authors view is that

the above approach translates investment risk into estimation error. Most prominently we see this with cash. Cash has neither investment nor estimation risk and in the Tüntücü and König (2004) procedure it will be the highest return asset with the lowest risk entry (zero volatility and correlation). Suppose the cash return is 2 per cent, while a given risky asset is distributed with a 3 per cent risk premium and 20 per cent volatility. These numbers have been estimated with 60 monthly observations. If we identify the 'worst case' expected return as a three standard deviation event the respective entry in μ_l will become $5\% - 2 \cdot \frac{20\%}{\sqrt{60}} = -0.164\%$, ⁴ which is considerably beneath cash. For any optimisation based on (2) with an investment universe containing both risky assets and cash will end up with a 100 per cent cash holding as long as we look deep enough into the estimation error tail.⁵ This seems to be overly pessimistic.⁶ Moreover, we can see from (2) that the Tüntücü/König formulation is equivalent to very narrow Bayesian priors. Investors would get the same result by putting a 100 per cent weight on their priors about $\boldsymbol{\mu}_l$ and $\boldsymbol{\Omega}_h$. This is hardly a plausible proposition.

A more general objective function for robust portfolio construction

Suppose we are given an *m*-dimensional vector of true expected returns μ , that is distributed around a mean vector, $\bar{\mu}$, and a known covariance matrix of estimation errors, Σ .⁷ Suppose further the known variance covariance matrix of asset returns is given by the symmetric $m \times m$ matrix Ω . Note that we focus on errors in expected returns and assume the covariance matrix of asset returns to be known, such that $\Sigma = n^{-1} \mathbf{\Omega}$, where *n* denotes the number of observations used to estimate expected returns. We will maintain this interpretation unless otherwise mentioned.

It is well known from statistics that α per cent of the distribution of expected returns lie within an ellipsoid defined by

$$(\boldsymbol{\mu} - \bar{\boldsymbol{\mu}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}) \leq \kappa_{\alpha,m}^{2}$$
(3)

with $\kappa_{\alpha,m}^2 = \chi_m^2(1-\alpha)$, $\chi_n^2(1-\alpha)$ is the inverse of a chi-square distribution with *m* degrees of freedom. For $\alpha = 95$ per cent and m = 8 we can say that 95 per cent of all expected returns lie within a statistical distance of 15.5 as defined in (3). Moving alongside the ellipsoid covers all possible μ that are within a provided confidence band. We can use this relationship to assess how large the difference between estimated and realised portfolio return can become, given a particular confidence region and vector of portfolio weights. Analytically we maximise the difference between expected portfolio return $\mathbf{w}^{T}\bar{\boldsymbol{\mu}}$ and worst case statistically equivalent portfolio returns $\mathbf{w}^{\mathrm{T}}\boldsymbol{\mu}$, that is those inputs that are along the ellipsoid defined in (3). Hence, we solve the following optimisation problem

$$L(\bar{\boldsymbol{\mu}}, \theta) = \mathbf{w}^{\mathrm{T}} \bar{\boldsymbol{\mu}} - \mathbf{w}^{\mathrm{T}} \boldsymbol{\mu} - \frac{\theta}{2} ((\boldsymbol{\mu} - \bar{\boldsymbol{\mu}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}) - \kappa_{\alpha,m}^{2})$$
(4)

where θ defines the Lagrange multiplier associated with the ellipsoid constraint. Essentially we look for the maximum distance $\mathbf{w}^T \bar{\boldsymbol{\mu}} - \mathbf{w}^T \boldsymbol{\mu}$ using $\boldsymbol{\mu}$ as choice variable for any given allocation \mathbf{w} . First take derivatives of (4) with respect to $\boldsymbol{\mu}$ and θ

$$\frac{\mathrm{d}L}{\mathrm{d}\boldsymbol{\mu}} = -\mathbf{w} - \theta \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}) = 0 \tag{5}$$

$$\frac{\mathrm{d}L}{\mathrm{d}\theta} = (\boldsymbol{\mu} - \bar{\boldsymbol{\mu}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \bar{\boldsymbol{\mu}}) - \kappa_{\alpha,m}^2 = 0 \quad (6)$$

Solving (5) for $(\mu - \bar{\mu})$ we arrive at $(\mu - \bar{\mu}) = -(1/\theta)\Sigma w$. This can then be

substituted into (6) to get us

$$\left(-\frac{1}{\theta}\boldsymbol{\Sigma}\mathbf{w}\right)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left(-\frac{1}{\theta}\boldsymbol{\Sigma}\mathbf{w}\right) - \kappa_{\alpha,m}^{2} = 0$$
$$\frac{1}{\theta^{2}}\mathbf{w}^{\mathrm{T}}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Sigma}\mathbf{w} - \kappa_{\alpha,m}^{2} = 0 \tag{7}$$
$$\frac{1}{\theta^{2}}\sigma^{2} = \kappa_{\alpha,m}^{2}$$

We now solve for $1/\theta$ to substitute this back into (5)

$$\boldsymbol{\mu} = \bar{\boldsymbol{\mu}} - \left(\frac{\kappa_{\alpha,m}}{\sigma}\right) \boldsymbol{\Sigma} \mathbf{w}$$
(8)

Finally, multiply both sides with \mathbf{w}^{T} to arrive at an expression for the distance between expected portfolio return $\mathbf{w}^{\mathrm{T}}\bar{\boldsymbol{\mu}}$ and worst case statistically equivalent portfolio returns $\mathbf{w}^{\mathrm{T}}\boldsymbol{\mu}$,

$$\mathbf{w}^{\mathrm{T}}\bar{\boldsymbol{\mu}} - \mathbf{w}^{\mathrm{T}}\boldsymbol{\mu} = \frac{\kappa_{\alpha,m}}{\sigma}\mathbf{w}^{t}\boldsymbol{\Sigma}\mathbf{w}$$

$$\mathbf{w}^{\mathrm{T}}\boldsymbol{\mu} = \mathbf{w}^{\mathrm{T}}\bar{\boldsymbol{\mu}} - \kappa_{\alpha,m}\sigma$$
(9)

In other words: what is the lowest value for expected portfolio returns as we move along the α per cent-ellipsoid? The factor $\kappa_{\alpha,m}$ can be heuristically viewed as an aversion to estimation error (uncertainty), although we have calibrated it differently above.

Robust portfolio optimisation uses $\mathbf{w}^{\mathrm{T}}\bar{\boldsymbol{\mu}} - \kappa_{\alpha m}\boldsymbol{\sigma}$ instead of $\mathbf{w}^{\mathrm{T}}\bar{\boldsymbol{\mu}}$ as estimate for expected returns. The portfolio construction problem becomes then⁸

$$\mathbf{w}_{rob}^{*} = \arg\max_{w \in C} \mathbf{w}^{\mathrm{T}} \bar{\mathbf{\mu}} - \kappa_{\alpha,m} \sigma$$
$$-\frac{\lambda}{2} \mathbf{w}^{\mathrm{T}} \mathbf{\Omega} \mathbf{w}$$
(10)

instead of

$$\mathbf{w}_{m\nu}^{*} = \operatorname*{arg\,max}_{w \in C} w^{\mathrm{T}} \bar{\mathbf{\mu}} - \frac{\lambda}{2} \mathbf{w}^{\mathrm{T}} \Omega \mathbf{w} \qquad (11)$$

for Markowitz-based portfolio optimisation. Note that $\mathbf{w} \in \mathbf{C}$ serves as a shorthand for investment constraints (full investment, nonnegativity, sector neutrality, beta neutrality, etc).

Robust portfolio optimisation maximises the worst case expected portfolio return for a given confidence region subject to risk .∰. Scherer

return considerations. The computational difficulty with (10) is, that it can no longer be solved using quadratic programming (as it contains a square root) if the constraint set **C** contains non-negativity ($\mathbf{w} \ge 0$) constraints. We need to either apply second-order cone programming⁹ or use an optimiser that can handle general convex expressions.

Note that if $\kappa_{\alpha,m}$ is assumed to be large, this forces the optimal solution towards assets that are relatively free from estimation error. For $\Sigma = n^{-1} \Omega$, estimation and investment risk move hand in hand and a larger aversion to estimation risk also reduces investment risk. The optimal portfolio invests more heavily into less risky assets. In the extreme, cash is the only asset without estimation risk, as its return is known at the beginning of the period with certainty. We will exemplify these statements in the next sections.

Deriving optimal portfolio weights

In this section, we will look at the implicit assumptions and properties of the robust portfolio construction mechanism. This analysis is inherently in sample in nature as it does not place weight on the actual (out of sample) performance of constructed portfolios, but rather checks consistency with decision theory as well as evaluating the additional properties relative to established algorithms.

We start with deriving a closed form solution for (10) in order to better understand the mechanics of robust optimisation. For means of comparison, we first state the familiar solution to the traditional portfolio optimisation problem within our context and notation. The traditional optimisation problem is given by

$$L(\mathbf{w}, \theta) = \mathbf{w}^{\mathrm{T}} \bar{\boldsymbol{\mu}} - \frac{\lambda}{2} \mathbf{w}^{\mathrm{T}} \boldsymbol{\Omega} \mathbf{w} + \theta(\mathbf{w}^{\mathrm{T}} \mathbf{I} - 1)$$
(12)

where θ denotes the multiplier associated with the full investment constraint ($\mathbf{w}^{T}\mathbf{I} = 1$). After taking first-order derivatives with respect to the Lagrange multiplier and the vector of portfolio weights, solving for the Lagrange multiplier and substituting this back into the derivative with respect to portfolio weights, we arrive at the familiar solution:

$$\mathbf{w}_{m\nu}^{*} = \frac{1}{\lambda} \Omega^{-1} \left(\bar{\boldsymbol{\mu}} - \frac{\boldsymbol{\mu}^{\mathrm{T}} \Omega^{-1} \mathbf{1}}{\mathbf{1}^{\mathrm{T}} \Omega^{-1} \mathbf{1}} \mathbf{1} \right) + \frac{\Omega^{-1} \mathbf{1}}{\mathbf{1}^{\mathrm{T}} \Omega^{-1} \mathbf{1}}$$
(13)

The optimal Markowitz portfolio can be written as the combination of the minimum variance portfolio

$$w_{min} = \frac{\mathbf{\Omega}^{-1} \mathbf{1}}{\mathbf{1}^{\mathrm{T}} \mathbf{\Omega}^{-1} \mathbf{1}} \tag{14}$$

that is neither dependent on preferences (λ) nor expected returns ($\bar{\mu}$) and a speculative demand

$$w_{spec} = \frac{1}{\lambda} \Omega^{-1} \left(\bar{\boldsymbol{\mu}} - \frac{\boldsymbol{\mu}^{\mathrm{T}} \Omega^{-1} \boldsymbol{1}}{\boldsymbol{1}^{\mathrm{T}} \Omega^{-1} \boldsymbol{1}} \boldsymbol{1} \right)$$
(15)

that depends on those factors. Note that $\mu^T \Omega^{-1} 1/1^T \Omega^{-1} 1$ equals the return of the minimum variance portfolio. The speculative part increases if returns (opportunities) increase or risk aversion falls. This is the familiar two-fund separation.

Robust optimisation instead aims at trading off the minimum expected return for a given level of confidence against risk. Using the same notation as in (12), this problem can be written as maximising

$$L(\mathbf{w}, \theta) = \mathbf{w}^{\mathrm{T}} \bar{\mathbf{\mu}} - \kappa_{\alpha,m} n^{-\frac{1}{2}} \sigma_{p}$$
$$-\frac{\lambda}{2} \sigma_{p}^{2} + \theta(\mathbf{w}^{\mathrm{T}} \mathbf{1} - 1) \qquad (16)$$

where $\sigma_p^2 = \mathbf{w}^T \Omega \mathbf{w}$, **1** is a $m \times 1$ vector of ones an $n^{-1/2} \sigma_p = (\mathbf{w}^T n^{-1} \Omega \mathbf{w})^{1/2} = n^{-1/2}$ $(\sigma_p^2)^{1/2}$. The first-order condition with respect to **w** is given as

$$\frac{\mathrm{d}L}{\mathrm{d}\mathbf{w}} = \bar{\mathbf{\mu}} - \left(\frac{n^{-\frac{1}{2}}\kappa_{\alpha,m} + \lambda\sigma_p}{\sigma_p}\right) \mathbf{\Omega}\mathbf{w} - \lambda \mathbf{1} = 0$$
(17)

Note that the bracketed term in (17) is a scalar which allows us to solve for \mathbf{w}

$$\mathbf{w} = \left(\frac{\sigma_p}{n^{-\frac{1}{2}}\kappa_{\alpha,m} + \lambda\sigma_p}\right) \mathbf{\Omega}^{-1}(\bar{\mathbf{\mu}} - \theta \mathbf{I}) \quad (18)$$

Transpose both sides, multiply by $\mathbf{1}$ and use $\mathbf{w}^{\mathrm{T}}\mathbf{1} = 1$ to arrive at

$$\mathbf{w}^{\mathrm{T}}\mathbf{1} = \left(\frac{\sigma_{p}}{n^{-\frac{1}{2}}\kappa_{\alpha,m} + \lambda\sigma_{p}}\right) \quad (\bar{\mathbf{\mu}}^{\mathrm{T}} - \theta\mathbf{1}^{\mathrm{T}})\Omega^{-1}\mathbf{1}$$

$$= \left(\frac{\sigma_p}{n^{-\frac{1}{2}}\kappa_{\alpha,m} + \lambda\sigma_p}\right)(b - \theta a) = 1$$

where $b = \bar{\boldsymbol{\mu}}^{\mathrm{T}} \boldsymbol{\Omega}^{-1} \mathbf{1}$, $\mathbf{a} = \mathbf{1}^{\mathrm{T}} \boldsymbol{\Omega}^{-1} \mathbf{1}$. From (19), we can solve for θ

$$\theta = \frac{1}{a} \left(b - \frac{n^{-\frac{1}{2}} \kappa_{\alpha,m} + \lambda \sigma_p}{\sigma_p} \right)$$
(20)

Substituting (20) into (18) yields

$$\mathbf{w}_{rob}^{*} = \left(1 - \frac{n^{-\frac{1}{2}}\kappa_{\alpha,m}}{\lambda\sigma_{p}^{*} + n^{-\frac{1}{2}}\kappa_{\alpha,m}}\right) \frac{1}{\lambda} \mathbf{\Omega}^{-1} \\ \times \left(\bar{\mathbf{\mu}} - \frac{\bar{\mathbf{\mu}}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{1}}{\mathbf{1}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{1}}\mathbf{1}\right) + \frac{\mathbf{\Omega}^{-1}\mathbf{1}}{\mathbf{1}^{\mathrm{T}}\mathbf{\Omega}^{-1}\mathbf{1}} \\ = \left(1 - \frac{n^{-\frac{1}{2}}\kappa_{\alpha,m}}{\lambda\sigma_{p}^{*} + n^{-\frac{1}{2}}\kappa_{\alpha,m}}\right) \mathbf{w}_{spec}^{*} + \mathbf{w}_{min}$$

$$(21)$$

where σ_p^* denotes the standard deviation of the optimal robust portfolio.¹⁰ For low required confidence levels ($\kappa_{\alpha,m} \rightarrow 0$) as well as for many data ($n \rightarrow \infty$) the optimal portfolio converges to a mean variance efficient (frontier) portfolio as $(1-(n^{-1/2}\kappa_{\alpha,m})/\lambda\sigma_p^* + n^{-1/2}\kappa_{\alpha,m}) \rightarrow 1$, which results in¹¹

$$\mathbf{w}_{rob}^{*} = \frac{1}{\lambda} \Omega^{-1} \left(\bar{\boldsymbol{\mu}} - \frac{\boldsymbol{\mu}^{\mathrm{T}} \Omega^{-1} \mathbf{1}}{\mathbf{1}^{\mathrm{T}} \Omega^{-1} \mathbf{1}} \mathbf{1} \right) + \frac{\Omega^{-1} \mathbf{1}}{\mathbf{1}^{\mathrm{T}} \Omega^{-1} \mathbf{1}} = \mathbf{w}_{mv}^{*}$$
(22)

If $\kappa_{\alpha,m} \rightarrow \infty$, or if $n \rightarrow 0$ the robust portfolio, however, converges to the minimum

variance portfolio

$$\mathbf{w}_{nob}^{*} = \frac{\mathbf{\Omega}^{-1} \mathbf{1}}{\mathbf{1}^{\mathrm{T}} \mathbf{\Omega}^{-1} \mathbf{1}} = \mathbf{w}_{min}$$
(23)

We see that the robust portfolio ranges between a mean variance efficient portfolio (with speculative investment demand) and the minimum variance portfolio (that ignores the information in return estimates).

In sample critique

(19)

How well is uncertainty aversion rooted in decision theory?

The author argues, that what matters after all for investors is the predictive distribution of future returns (as it determines an investors expected utility) given by $p(\mathbf{\tilde{r}} | \mathbf{r}_{hist})$, where $\mathbf{\tilde{r}}$ denotes the future returns yet unknown. The distribution is conditioned only by the observed data \mathbf{r}_{hist} and not by any fixed realisation of the parameter vector $\mathbf{\theta}$ (covariances, means). We can express the predictive distribution as¹²

$$p(\mathbf{\tilde{r}}|\mathbf{r}_{hist}) = \int p(\mathbf{\tilde{r}}|\theta) p(\theta|\mathbf{r}_{hist}) \mathrm{d}\theta \qquad (24)$$

From (24) we can easily see, that it is irrelevant where the variation in future returns comes from. It could either come from estimation error $p(\boldsymbol{\theta} | \mathbf{r}_{hist})$ or from the conditional distribution of asset returns $p(\mathbf{\tilde{r}} | \mathbf{\theta})$. In this respect, it seems to make little sense to differentiate between model uncertainty and risk. Both are inseparable. In other words: if investors believe, returns could come from an array of different distributions with different parameters, they will use (24) to model the predictive distribution taking account for parameter uncertainty. If an investor is shown the predictive distribution he does not care how much of it is due to parameter uncertainty and how much to investment risk.

It should, however, be mentioned that the literature provides conflicting views on the above. Robust optimisation traces back to



Knight (1921) who distinguishes (without axiomatic foundations) between aversion to risk, where objective probabilities exist to guide investment decisions and aversion to uncertainty where decision makers cannot even define probabilities. Very much to the contrast, Savage (1951) showed that decision makers rationally act by placing a prior on the parameter space to maximise posterior expected utility as long as they satisfy a set of axioms on coherent behaviour. In fact, individuals use all available tools to calculate subjective probabilities for expected utility maximisation (SEU).¹³ This framework came under attack on behavioural grounds. Ellsberg (1961) observed ambiguity aversion in a series of experiments similar to the following.¹⁴ An urn contains 300 balls, with 200 being a mixture of blue and green and 100 being red. Participants receive 100 Euro if a random draw selects a ball from a prespecified colour. Participants are asked, whether they prefer this colour to be red or blue. Alternative participants receive 100 Euro if the selected ball is not from the prespecified colour. Again do you prefer red or blue? The most frequent response is red in both cases. If, however, red is preferred in the first case, the subjective probability for red must be higher than for blue. This must mean that you should prefer blue in the second experiment as the probability of not receiving blue (where you now receive money for) is higher than observing blue. A choice of red in both experiments is not coherent and therefore a violation of Savages SEU. What do we make from this? For a start this is merely empirical evidence that some investors might behave irrationally. Dismissing SEU on these grounds is similar to dismissing stochastic calculus because many students repeatedly fail in experiments called exams. In fact, scientists should help individuals to make better decisions, that is erecting a normative framework, rather than following a more descriptive approach that tries to expost rationalise the Ellsberg paradox. For example, Gilboa and

Schmeidler (1989) showed that 'inventing' a decision maker following the minmax principle (under a different set of axioms) could reconcile the Ellsberg Paradox. Despite the intellectual beauty of their work a major problem remains. Minmax preferences are not in any respect superior to those already established by maximising expected utility with subjective utilities. Not only do they violate Savage's sure thing principle (if decision makers prefer x to y in all possible states of the world, then they should also prefer x to y in any particular state of the world) but they can also lead to a Dutch Book outcome, a situation where someone agrees to a set of bets that cause him to lose money with probability one.¹⁵ In the authors mind these are more serious consequences than the Ellsberg paradox is for SEU.

How different is robust optimisation relative to already existing methods?

Let us interpret (21). The careful reader will realise that the previous result essentially views robust optimisation as shrinkage estimator that combines the minimum variance portfolio with a speculative investment portfolio, where the weighting factor is given by

$$\left(1 - \frac{n^{-\frac{1}{2}}\kappa_{\alpha,m}}{\lambda\sigma_p^* + n^{-\frac{1}{2}}\kappa_{\alpha,m}}\right) \le 1$$
(25)

Note that the weighting factor contains σ_p^* , that is the optimal volatility of robust portfolios, which is only known after the robust portfolio has been constructed. This makes 'robust shrinkage' a very intransparent and difficult to control process as the weighting factor is endogeneous. As long as estimation error aversion is positive, this term will always be smaller than one. Robust portfolio construction will not be different from a shrinkage estimator like Jorion (1986) as it simply interpolates between the minimum variance portfolio and the maximum Sharpe-Ratio portfolio.

Additionally, the efficient set (the set of all solutions, ie optimal portfolio) coincides with the mean variance efficient set. Solutions for investors with a particular risk aversion only differ to the extent lower weight is given to the speculative portfolio. An alternative way to see this is to rewrite (10) as $U = \mathbf{w}^{\mathrm{T}} \bar{\mathbf{\mu}} - \lambda_1 / 2 \sqrt{\mathbf{w}^{\mathrm{T}} \Omega \mathbf{w}} - (\lambda_2 / 2)$ $\mathbf{w}^{\mathrm{T}} \mathbf{\Omega} \mathbf{w}$. Taking first derivatives yields and solving for \mathbf{w}^* yields $\mathbf{w}^* = (1/(\lambda_1 1/\sigma + \lambda_2))$ $\mathbf{\Omega}^{-1}\bar{\mathbf{\mu}}$ which is essentially equal to $\mathbf{w}^* = 1/\lambda^* \mathbf{\Omega}^{-1} \bar{\mathbf{\mu}}$, where λ^* is the pseudo risk aversion that makes Markowitz and robust optimisation coincide. In other words: the appropriate choice of λ^* in mean variance optimisation will recover the robust optimisation result. Viewed this way robust optimisation offers nothing additional apart from increased in-transparency. The return adjustments are not user specified but determined during the optimisation. Also we note an increased ambiguity in parameter choice: How can we justify our choice of $\kappa_{\alpha,m}$ and how can we make it 'consistent' with risk aversion.¹⁶

What are the implicit return refinements made by robust optimisation?

We have seen in (8) that robust optimisations uses $\bar{\mu} - (\kappa_{\alpha,m}/\sigma) \Sigma \mathbf{w}$ instead of $\bar{\mu}$ as inputs for expected returns. The return adjustment consists of factor $\kappa_{\alpha,m}$ plus a measure for the marginal contribution to estimation risk. Essentially this describes a situation where assets with the highest marginal contribution to estimation error $((1/\sigma)\Sigma \mathbf{w})$ are more heavily penalised in terms of expected returns than assets with lower estimation error contributions. The effective expected return in robust optimisation thus depends on both estimation error and actual position. Assets that carry positive weights get a return subtraction (to make the overweight less attractive), while assets with negative weights get a return add on (to make the short position less attractive). While this is aimed at mitigating estimation error maximisation, it is not extremely pessimistic. Essentially

robust optimisation moves alongside the ellipsoid (3) and implicitly assumes that estimation error is always on the wrong side, overestimating the expected returns of assets that are over-weighted and underestimating the expected returns of under-weighted assets. Apart from an unsymmetric treatment of estimation error, it also creates a logical impossibility. Expected returns (before transaction costs) can never be dependent on position size or sign. In fact, return expectations are made separate from portfolio construction.

Out of sample critique

So far we have seen that robust optimisation is similar to Bayesian shrinkage, without its theoretical foundation or transparency. As such it is unclear why investors should prefer it to Bayesian analysis. Additionally, this section will show that robust optimisation methods can severely underperform traditional mean variance optimisation as the inability to consistently determine $\kappa_{\alpha,m}$ and λ can lead to severe out of sample underperformance of robust optimisation relative to naive optimisation.

Before we start with our out of sample testing example, we should summarise the key principles for out of sample evaluation as they apply to our case.

1. Out of sample testing is not equivalent to a rolling period analysis through a historical sample path: A particular sample path might have characteristics that put an unfair advantage (ie an advantage that is not universal across many sample path) to a particular method. For example: downside risk-based methodologies might find little downside risk in an upwards trending markets and hence overweight risky assets (based on their low downside risk) leading to an immediate advantage over mean variance-based measures that is spurious and does not generally hold. We therefore need to employ a large number of

simulations across many economic environments to evaluate a portfolio construction mechanism. This is best done with the use of Monte Carlo simulations where we have perfect control over the underlying processes.

2. Out of sample testing requires the evaluation of expected (out of sample) utility, as this is the only measure with a sound foundation in decision theory: What are alternatives to compare portfolio construction methodologies? Is it the probability of one method outperforming the other in terms of realised return? Is it the turnover generated as new information becomes available? The author believes that out of sample comparisons always need to be made on the basis of expected out of sample utility. This is the only way to ensure that comparable decisions are made across samples with different risk return trade offs. Minimising risk for a given return target does not meet these criteria as it leads to relatively risky portfolio for samples with depressed risk premium and relatively less risky portfolios across samples with high risk premium. This implies that risk aversions change dependent on the market risk premium, which is highly implausible.

Secondly, even if all the above is met, out of sample tests are purely statistical results.

3. Without underlying theoretical arguments, the results of out of sample tests are impossible to generalise and hence are highly data dependent: While we could to certain extent address this concern by sampling across a wide array of alternative covariance matrices, we can never generalise an argument that is essentially build upon pure empirical results. To put it differently: how can we be certain that it works for other than the tested situations if we have no theoretic underlying?

How should we design the out of sample testing for robust portfolio optimisation? The author suggests the following set up described in Figure 1.

1. Assume a true mean vector $\bar{\mu}$ and covariance matrix Ω . Draw s = 1, ..., S = 1,000 samples from

$$\bar{\boldsymbol{\mu}}_{s} \sim N(\bar{\boldsymbol{\mu}}, n^{-1}\boldsymbol{\Omega}) \tag{26}$$

with n = 60,120. Estimation error is equivalent to five or ten years of monthly data. Essentially this means that we focus on estimation error in means, while we assume the covariance matrix to be perfectly known.



Figure 1 Out of sample testing. We choose one vector of true expected returns (out of many alternatives, indicated by the grey area). From this, we resample 1,000 statistically equivalent vectors of expected returns for alternative data equivalents, that is for n = 60,120 according to $\bar{\mu}_s \sim N(\bar{\mu}, n^{-1}\Omega)$. For each s = 1, ..., 1,000 realisations we calculate optimal portfolios for both construction methodologies. These portfolios are then evaluated (out of sample) with the true distribution $\bar{\mu}, \Omega$.

- 2. For each realisation (sample draw), we construct traditional as well as robust portfolios using (10) and (11) for varying risk aversions ($\lambda = 0.01, 0.025, 0.05$) and confidence requirements ($\alpha = 99.99$, 97.5, 95 per cent).
- 3. We construct 1,000 portfolios for each method.¹⁷ Both algorithms (traditional mean variance as well as robust optimisation) adjust to the sampled data only. The true mean vector is not known to any method. In contrast to the previous section we add a non-negativity constraint on portfolio weights ($\mathbf{w} \ge 0$) for both methodologies. Each of these constructed portfolio is then evaluated under the true distribution, that is we calculate

$$U_{s}(\mathbf{w}_{m\nu,s}^{*}(\bar{\boldsymbol{\mu}}_{s}), \bar{\boldsymbol{\mu}}, \boldsymbol{\Omega})$$
$$= (\mathbf{w}_{m\nu,s}^{*})^{\mathrm{T}} \bar{\boldsymbol{\mu}} - \frac{\lambda}{2}$$
$$\times (\mathbf{w}_{m\nu,s}^{*})^{\mathrm{T}} \boldsymbol{\Omega}(\mathbf{w}_{m\nu,s}^{*}) \qquad (27)$$

$$U_{s}(\mathbf{w}_{rob,s}^{*}(\bar{\boldsymbol{\mu}}_{s}), \bar{\boldsymbol{\mu}}, \boldsymbol{\Omega})$$
$$= (\mathbf{w}_{rob,s}^{*})^{\mathrm{T}} \bar{\boldsymbol{\mu}} - \frac{\lambda}{2}$$
$$\times (\mathbf{w}_{rob,s}^{*})^{\mathrm{T}} \boldsymbol{\Omega}(\mathbf{w}_{mv,s}^{*})$$
(28)

In other words: we calculate the optimal response ($\mathbf{w}_{mv,s}^*, \mathbf{w}_{rob,s}^*$) for each s = 1, ..., 1000 draws and calculate the out of sample utility for the true return distribution, that is the utility we would experience out of sample.¹⁸ Averaging across all draws we get the expected utility for a given portfolio construction mechanism.

In order to appreciate the above procedure we go through a detailed example for $\lambda = 0.01$, a total of eight assets (m = 8) and a required confidence of $\alpha = 99.9$ per cent ($\kappa_{99.9 \text{ per cent},8} = 5.11$). We use the data from Michaud (1998) on global equity and fixed income markets. Running 1,000 draws will allow us to evaluate the 'robustness' of robust portfolios. After all the main perceived property of robust optimisation is to dampen the response of portfolio weights with respect to variations in expected returns. The results are summarised in Figure 2 for the first 100 draws. We see that robust optimisation indeed creates portfolios that are robust to changes in expected returns, that is optimal allocations vary much less as new and potentially noisy information comes along. Traditional portfolios seem much more concentrated in very few assets (sometimes even only one) hitting corner solutions in the optimisation process. Why should 'robustness' be a valuable property, however? We cannot infer this from Figure 2. After all robust portfolios might be overly diversified, that is not concentrated enough into high yielding assets for an aggressive investor. The only way to check the claim that robust optimisation delivers superior performance is to compare the expected (out of sample) utility from the portfolios in Figure 2 using the true return vector (ie the one we sampled from at the beginning). If we then find that

$$\frac{1}{S}\sum_{s=1}^{S}U_{s}(\mathbf{w}_{m\nu,s}^{*}(\bar{\boldsymbol{\mu}}_{s}),\bar{\boldsymbol{\mu}},\boldsymbol{\Omega}) = EU_{m\nu} \geq EU_{rob}$$
$$=\frac{1}{S}\sum_{s=1}^{S}U_{s}(\mathbf{w}_{rob,s}^{*}(\bar{\boldsymbol{\mu}}_{s}),\bar{\boldsymbol{\mu}},\boldsymbol{\Omega})$$
(29)

With statistical significance, we can say that robust optimisation outperforms traditional optimisation.¹⁹

Take a look at the distribution of utilities (for each s = 1, ..., 1000 draws) in Figures 3 and 4. Utilities across scenarios are positively correlated (Figure 3). If out of sample utility is high for traditional portfolios, it also tends to be high for robust portfolios. While out of sample utility for robust portfolios is different for every single sample, it seems to be stuck at certain levels for traditional portfolio construction. These are simply the corner portfolios, that is the optimisation is stuck at the same solution for a variety of inputs. The histograms in Figure 4 pick up the same





Figure 2 Robust versus traditional portfolio construction ($\lambda = 0.01$, n = 60, $\alpha = 99.99$ per cent, S = 1,000). Robust portfolio react less sensitive to changes in expected returns. Given the high required confidence of $\alpha = 99.99$ per cent, robust portfolios invest heavily in assets with little estimation error. This is entirely different with our intuition that error in return estimates becomes less and less important as we move towards the minimum risk portfolio. The data are taken from Michaud (1998) and the abbreviations used are FI.EU (fixed income Europe), FI.US (fixed income US), EQ.US (equity US), EQ.UK (equity UK), EQ.Jap (equity Japan), EQ.Ger (equity Germany), EQ.Fra (equity France) and EQ.Can (equity Canada).



Figure 3 Mean variance versus robust optimisation ($\lambda = 0.01$, n = 60, $\alpha = 99.99$ per cent, S = 1,000). We plot the utility from mean variance optimisation $U(\mathbf{w}_{mv}^*(\bar{\mu}_s),\bar{\mu},\Omega)$ versus the utility from robust optimisation $U(\mathbf{w}_{rob}^*(\bar{\mu}_s),\bar{\mu},\Omega)$. The vertical lines represent corner solutions, that is the optimiser arrives at the same solution for different set of inputs.

effect, where more weight is given to the more frequent corner portfolios. Moreover, we see that sample utility is less extremely distributed in the robust case, which is a direct consequence of robustness. For means of comparison we will calculate three statistics. First, we state the difference in expected utility between both construction methodologies. A difference of 0.1, for example, means a 10 basis points

٠Ж



Figure 4 Histogram of out of sample utilities ($\lambda = 0.01$, n = 60, $\alpha = 99.99$ per cent, S = 1,000). Out of sample utility traditional optimisation peaks around corner portfolios, while it seems to be much smoother distributed for robust optimisation. Both very low and very large utilities are reported under traditional optimisation.

(bps) return advantage (measured as security equivalent) per month. Secondly, we calculate the statistical significance of the difference in expected utility. You can think of this as the *t*-value on the intercept of a regression of utility differences against a constant. Finally, we also calculate the probability that robust optimisation outperforms traditional optimisation simply stated as percentage. In our example, we find that the difference in expected utility amounts to -26.24 bps per month, which adds to a return disadvantage of more than 300 bps per year. Not only is this a very sizeable result, but it also comes with a *t*-value of 16 and the probability of robust optimisation outperforming traditional is a mere 6.8 per cent. Robust optimisation underperforms traditional optimisations significantly for the above example. The reason for this is, that the high aversion to estimation error conflicts with the low risk aversion. This is in general the problem with addressing uncertainty aversion separately from risk aversion. Portfolios are shrunk too

much towards the minimum variance portfolio, to be consistent with low risk aversion. It would be unfair to base a comparison between two methodologies on a single parameter constellation. In order to get a more complete picture, we repeat the above exercise for many parameters and summarise the results in Table 1.

Robust optimisation will lead to inferior out of sample results, if investors show little risk aversion ($\lambda = 0.01, 0.025$), but high uncertainty aversion. For high values of $\kappa_{\alpha,m}$, the robust optimisation approach forces investors into portfolios that lean towards assets with little estimation risk. For the case $\Sigma = 1/n\Omega$, however, this is equal to portfolios with little investment risk. Out of sample this leads to a deterioration of expected utility. Robust portfolios are simply not aggressive enough.

While this interpretation holds qualitatively for the case with small (n = 60)as well a large (n = 120) estimation error, we see that robust does increasingly worse if estimation error is reduced. If the precision

| | α=95% | α=97.5 % | α=99.99% | | | | | |
|---------------------------|--------------|-----------------|-----------------|--|--|--|--|--|
| Small sample size (n=60) | | | | | | | | |
| λ=0.05 | 9.7 bps | 8.7 bps | 7.79 bps | | | | | |
| | (20.16) | (17.76) | (16.09) | | | | | |
| | 68.5% | 63.4% | 61.0% | | | | | |
| λ=0.025 | -3.13 bps | -5.54 bps | -6.96 bps | | | | | |
| | (-18.23) | (-14.84) | (-11.85) | | | | | |
| | 39.14% | 32.6% | 28.8% | | | | | |
| λ=0.01 | -19.8 bps | -24.09 bps | -26.14 bps | | | | | |
| | (-49,6) | (-63.33) | (-70.07) | | | | | |
| | 13.0% | 8.1% | 6.8% | | | | | |
| Large sample size (n=120) | | | | | | | | |
| λ=0.05 | 4.3 bps | 3.1 bps | 2.1 bps | | | | | |
| | (12.5) | (8.82) | (5.99) | | | | | |
| | 57.4% | 50.7% | 45.4% | | | | | |
| λ=0.025 | –5.7 bps | -8.7 bps | -10.5 bps | | | | | |
| | (–16.65) | (-26.71) | (-33.31) | | | | | |
| | 29.1% | 20.1% | 16.4%% | | | | | |
| λ=0.01 | -18.4 bps | -24.1 bps | -27.2 bps | | | | | |
| | (-44.8) | (-72.22) | (-89.2) | | | | | |
| | 12.3% | 5.0% | 3.6% | | | | | |

| Table 1 | Out of samp | ole pe | erformance | for | full |
|----------|----------------|--------|------------|-----|------|
| investme | ent universe (| (m=8) | | | |

Table shows the relative performance of robust portfolio optimisation relative to traditional portfolio optimisation. The first number is the difference in expected utility, which we can interpret in terms of a security equivalent (ie basis points of monthly performance). The second number (in round brackets) represents the *t*-value of the difference in expected utility (a value of about 2 would be significant at the 5% level, for a two-sided hypothesis), while the third number represents the percentage of runs, where robust optimisation generated a higher out of sample utility than traditional optimisation.

in estimates becomes larger (estimation error is removed), the traditional Markowitz-based approach is more powerful in selecting optimal portfolios. For very low estimation errors, it makes little sense to demand a high estimation error aversion.

Conclusions

Robust portfolio optimisation aims at explicitly incorporating estimation error into the portfolio optimisation process. The above contribution has formally shown that robust methods are equivalent to shrinkage estimators and leave the efficient set unchanged. In other words: they offer nothing new. All this, however, comes at the expense of computational difficulties (second-order cone programming). Moreover, the return adjustment process is largely in-transparent relative to Bayesian alternatives and suffers from a logical impossibility: Return estimates need to be independent from position sign and size. We constructed a simple but realistic example, that showed how severely robust optimisation (up to 300 bps in return, measured via security equivalent) will underperform even simple mean variance optimisation.

Notes

1. The mathematics has been developed by Halldorsson and Tüntücü (2003). Let us, for example, assume that there is only ambiguity about the mean vector and that the covariance matrix is known. Further assume we have 1,000 possible mean vector candidates from μ_1 to $\mu_{1,000}$. We can reformulate (1) into

$$\max_{\mathbf{w}, U_{\min}} (U_{\min}), \mathbf{w}^{\mathsf{T}} \boldsymbol{\mu}_{1} - \frac{\lambda}{2} \mathbf{w}^{\mathsf{T}} \boldsymbol{\Omega} \mathbf{w}$$

$$\geq U_{\min}, \dots, \mathbf{w}^{\mathsf{T}} \boldsymbol{\mu}_{1 \ 000} - \frac{\lambda}{2} \mathbf{w}^{\mathsf{T}} \boldsymbol{\Omega} \mathbf{w} \geq$$

Nuopt for S-Plus can deal with problems of this kind while other dedicated portfolio optimisers cannot.

 U_{\min}

- 2. Note that mean vector entries are uncorrelated with covariance entries.
- 3. As all covariance matrices are symmetric, it suffices to work through the upper or lower triangle.
- 4. For simplicity, we assumed σ to be known.
- 5. Also see Brinkmann (2005) for a review on Tüntücü and König (2004) as well as some out of sample tests. Using synthetic data with equal volatilities, he does not expose their method to this major deficiency and still gets only mixed results for Tüntücü/König.
- 6. Maxmin criteria are known to be overly pessimistic. Their use has recently be motivated by Gilboa and Schmeidler (1989) that try to capture ambiguity by applying maxmin to expected utility. See more on this in the section In sample critique.
- 7. This section draws on the work by Ceria and Stubbs (2005) and clean up some of their notation. We will extent their setting in the next section.
- 8. Ceria/Stubbs use the term $\|\Sigma^{1/2}\mathbf{w}\|$, which is the vector norm of a product that uses the square root of a matrix. This is computationally inefficient. Using the definition of a vector norm, we get $\|\Sigma^{1/2}\mathbf{w}\| = (\mathbf{w}^T\Sigma^{1/2}\Sigma^{1/2}\mathbf{w})^{1/2} = (\mathbf{w}^T\Sigma\mathbf{w})^{1/2} = \sigma$ which is much easier to interpret as estimation error and easier to implement.
- 9. See Ghaoui et al. (2003).
- 10. Note that σ_p^* is the solution to the polynomial $a_4\sigma_p^{\ 4} + a_3\sigma_p^{\ 3^p} + a_2\sigma_p^{\ 2} + a_1\sigma_p + a_0$ where the coefficients depend on the above model parameters. A proof is

available from the author upon request. As σ_p^* is determined endogenously we have little control over the degree of implied shrinkage.

- 11. We know that $(d\sigma_p^*)/(d\kappa_{\alpha,m}) < 0$, that is an increase in estimation error risk aversion will result in portfolios that carry less investment risk. This is needed to ensure that $1-(n^{-1/2}\kappa_{\alpha,m})/(\lambda\sigma_p^*+n^{-1/2}\kappa_{\alpha,m})$ converges to 0 as $\kappa_{\alpha,m}$ increases.
- 12. See Scherer (2004, p. 106)
- Given that the whole finance industry is devoting its resources to this task, this seems highly uncontroversial to the author.
- 14. See Kreps (1990).
- 15. See Sims (2001) for a critical view on minmax utility.
- 16. Recently there has been some work on this problem. See for example Maenhout (2004) and the quoted literature therein. The author, however, arrives at a similar result: a dramatic decrease in the demand for risky assets.
- Robust optimisation has been implemented in Nuopt for S-Plus. For more details, see Scherer and Martin (2005).
- 18. One might be tempted to argue that we cannot compare both methods on the basis of expected utility as an investor with uncertainty aversion actually exhibits a different utility function. This argument would, however, be misplaced as we investigate whether a mean variance investor can benefit from robust optimisation methods.
- 19. Note, that no reference is made here to transaction costs. If all transactions were free of cost, trading would have no impact on performance, but if transaction costs were substantial investors would be well advised to explicitly consider the costs of trading, rather than implicitly limit transactions.

References

Brinkmann, U. (2005) 'Robuste Portfoliooptimierung: Eine kritische Bestandsaufnahme und ein Vergleich alternativer Verfahren', in Haasis, H-D., Kopfer, H. and Schönberger, J. (eds.), Operations Research Proceeding, 2005, pp. 229–234.

- Ceria, S. and Stubbs, R. (2005) 'Incorporating Estimation Error into Portfolio Selection: Robust Efficient Frontiers, Axioma Working Paper.
- Ellsberg, D. (1961) 'Risk, Ambiguity and the Savage Axioms', The Quarterly Journal of Economics, 75, 643–669.
- Ghaoui, L., Oks, M. and Oustry, F. (2003) 'Worst Case Value-at-Risk and Robust Portfolio Optimization: A Conic Programming Approach', *Operations Research*, 51(4), 543–556.
- Gilboa, I. and Schmeidler, D. (1989) 'Maxmin Expected Utility with Non-Unique Prior', *Journal of Mathematical Economics*, 18, 141–153.
- Halldorsson, B. and Tüntücü, R. H. (2003) 'An Interiorpoint Method for a Class of Saddle Point Problems', *Journal of Optimization Theory and Applications*, 116(3), 559–590.
- Jorion, P. (1986) 'Bayes-Stein Estimation for Portfolio Analysis', Journal of Financial and Quantitative Management, 21, 279–291.
- Kreps, D. (1990) A Course in Microeconomic Theory, Prentice Hall, Englewood Cliffs, NJ.
- Knight, F. H. (1921) Risk, Uncertainty, and Profit, Hart, Schaffner & Marx, Boston.
- Maenhout, P. (2004) 'Robust Portfolio Rules an Asset Pricing', *Review of Financial Studies*, 17(4), 951–983.
- Michaud, R. (1998) *Efficient Asset Management*, Harvard Business School Press, Boston, MA.
- Savage, L. J. (1951) 'The Theory of Statistical Decision', Journal of the American Statistical Association, 46, 55–67.
- Scherer, B. (2004) Portfolio Construction and Risk Budgeting, 2nd edn, Riskwaters, London.
- Scherer, B. and Martin, D. (2005) Modern Portfolio Optimization with Nuopt for S-Plus, Springer, New York.
- Sims, C. (2001) 'Pitfalls of a Minmax Approach to Model Uncertainty. http://sims.princeton.edu/yftp/RobustPits/ RobustPitfalls.pdf.
- Tüntücü, R. H. and König, M. (2004) 'Robust asset allocation', Annals of Operations Research, 132, 132–157.