Robust Portfolio Optimization with Multiple Experts

Frank Lutgens and Peter Schotman*

April 29, 2004

Abstract

The success of quantitative approaches to portfolio choice crucially depends on the considered return model. Experts however do not agree on which return model is most appropriate. This controversy about the model specification introduces uncertainty in the optimal portfolio choice.

We will not meddle in the discussion on which model specification is most appropriate. Instead we consider the advice of various experts and aim for a portfolio choice which is robust to the advice of all the experts. More specifically we consider the portfolio choice with maximal performance for the most pessimistic advice.

We examine the effects of robust mean-variance portfolio choice: the difference with non-robust portfolio choice and its mean-variance performance. We test the robust approach empirically in the context of international portfolio choice with an investment set consisting of 81 portfolios from 9 European countries. Advice regarding the appropriate return model will be obtained from alternative experts advocating the CAPM, the international CAPM, the international Fama and French factor model. A bootstrap experiment demonstrates the merits - both in terms of expected performance as well as worst case performance - of the robust approach.

1 Introduction

Quantitative approaches to portfolio choice crucially rely on the employed return model. Each alternative return model supplies the decision maker with a specific set of parameters such as the expected asset returns and the covariance matrix of the asset returns. The optimal portfolio depends on these parameters and may thus differ among alternative return models. Experts however do not agree on which return model is most appropriate. The experts will therefore recommend different optimal portfolio choices.

Apart from relying on one particular expert, investors have also considered combining the information of alternative experts; they would use a weighted average of the models advocated by the different experts. However, the weighted average model is not automatically more appropriate than each of the individual models. Also the weights for the alternative models need to be chosen which requires additional modelling.

We consider portfolio choice that is robust to the advice of multiple experts who advocate different return models. Robust portfolio choice maximizes the performance over the least favorable return model. Hence the robust portfolio will - at least - obtain this robust performance, regardless of which alternative return model is the actual underlying model. The robust approach assumes no weights on the alternative return models but treats each alternative return model as plausible and will therefore fully account for each model in the decision. We will refer to this approach as the robust approach to model uncertainty.

Moreover, even if the correct return model specification would be known with certainty, there would still be estimation uncertainty: uncertainty about the true values of the model's parameters. The main source of

*This research is supported by Inquire Europe

estimation uncertainty is the limited amount of data upon which one can base the estimates. This causes the point estimates for the parameters, even in correctly specified models, to be (slightly) different from their true values. In the second part of our study we will extend the robust approach to account for estimation uncertainty.

We study robust portfolio choice in the context of Markowitz (1952) mean-variance portfolio optimization. The Markowitz optimization model measures portfolio return by the expected value of random portfolio return and risk by the variance of the random portfolio return. The portfolio choice aims at maximizing a given trade-off between the expected portfolio return and the portfolio variance.

The Markowitz (1952) problem presents an ideal basis to test techniques which deal with uncertainty in financial decision making: Michaud (1989), Jobson and Korkie (1981), and Jorion (1986), among others, have shown that disregarding uncertainty leads to poor and unstable mean-variance portfolio allocations. Model and estimation uncertainty lead to uncertainty in the estimates for expected returns and variance. The problem is that small changes in the expected return estimates may lead to fundamentally different portfolio choice. Different portfolios have a different actual performance. Moreover the perturbed estimates are empirically indistinguishable from the standard estimates. Therefore it is unclear what portfolio is best or even reasonable.

The plan of this paper is as follows. We start with a derivation of the robust mean-variance portfolio in section 2. We will consider the implications of model uncertainty and the added value of a robust approach. In section 3 we study the uncertainty in some return models that are commonly used in international portfolio choice. The results of this section will be used for the computation of the robust portfolios for the empirical test in the context of international portfolio choice (section 4). Next to a robust approach to model uncertainty, we will also consider a robust approach to estimation uncertainty in section 5. We will compare the importance of robustness to model and estimation uncertainty in section 6 and section 7 concludes.

2 Portfolio robust to a diversity of expert advice

We consider an investor with mean variance preferences over single-period portfolio choice. Her problem is to find the optimal allocation w to risky assets with expected excess returns and covariances given by the N-vector μ and $N \times N$ covariance matrix Σ respectively. The objective function is

$$Q(w) = \mu' w - \frac{1}{2}\gamma w' \Sigma w,$$

with γ measuring risk aversion. When μ and Σ are known, the optimal investment in the risky asset is

$$w^* = \frac{1}{\gamma}\Sigma^{-1}\mu.$$

In practice, the investor does not know the true values of μ and Σ. She obtains advice from J experts who provide her with their personal estimate (μ_j, Σ_j). The investor is convinced that these estimates are the only possible values for the true parameters. She is, however, not able to make a quantitative assessment as to which of the J parameter sets is more likely. A robust investor will consider the worst case and maximizes

$$\begin{aligned} Q_R(w) &= \min_j Q_j(w) \\ &= \min_j \mu_j' w - \frac{1}{2}\gamma w' \Sigma_j w \end{aligned}$$

This corresponds to what Rustem, Becker and Marty (2000) refer to as a model with rival return and rival risk scenarios.

Suppose the robust decision maker obtains advice (μ_j, Σ_j) from J experts. Her optimal portfolio which maximizes $Q_R(w)$ is (theorem 1, Appendix A):

$$w_R^* = \left(\sum_{j=1}^{J} \lambda_j \Sigma_j \right)^{-1} \left(\sum_j \lambda_j \mu_j \right) \tag{1}$$

where λ_j are constants satisfying $0 \leq \lambda_j \leq 1$ and $\sum \lambda_j = 1$.

The robust investor uses a weighted average of the opinions of the experts. In that sense the optimal portfolio is similar to the portfolio that a Bayesian investor would choose. The difference is that the Bayesian investor a priori assigns probabilities to the advice of each of the experts. The pseudo probabilities of the robust investor are derived endogenously from solving the maximization problem.

In view of (17b), positive weights λ_j refer to the models (pairs (μ_j, Σ_j)) which form(s) the worst case model for the robust portfolio. Equation 1 provides a partial solution in the sense that the λ_j are not given explicitly. A complete analytical solution for the minimization (2) with general μ_j and Σ_j is not available. Nevertheless special cases are analytically tractable and provide us with the intuition on robust portfolio choice. The next subsection considers the special case with one risky asset and two experts.

The effect of uncertainty about the correct model specification hinges on the question: *What happens if the model is incorrect?*. In that case the portfolio choice is suboptimal. The *loss function* expresses the utility loss that results from selecting a portfolio w which differs from the optimal portfolio (based on the actual parameters (μ, σ^2)). Formally,

$$\begin{aligned} L(\mu, \Sigma|w) &= \max_w Q(w) - Q(w) \\ &= \frac{1}{2\gamma}\mu'\Sigma^{-1}\mu - \mu'w + \frac{1}{2}\gamma w'\Sigma w \\ &= \frac{1}{2\gamma}\left(\mu - \gamma\Sigma w\right)'\Sigma^{-1}\left(\mu - \gamma\Sigma w\right) \end{aligned} \tag{2}$$

We can interpret $Q(w)$ as the risk adjusted expected return of a portfolio. Hence the loss function measures the expected return difference (after risk adjustment) between the optimal portfolio and a sub-optimal portfolio. The alternative interpretation of the loss function is the utility loss when the investor decides on portfolio w while the actual parameters are (μ, σ^2).

One risky asset, two experts

We illustrate the implications of robust portfolio choice with a simple, analytically tractable example. Consider an investor whose choice is limited to one risky asset and the risk-free asset. We assume that the expected excess return on the single risky asset is positive and consequently the optimal risky investment is non-negative. When we allow for borrowing at the risk free rate, the optimal portfolio could be $w > 1$. The quadratic objective function (2) reduces to $Q(w) = \mu w - \frac{1}{2}\gamma\sigma^2 w^2$ where σ^2 is the variance of the risky asset and w is a scalar in this case. The optimal portfolio w^* in (2) reduces to $w^* = \frac{\mu}{\gamma\sigma^2}$ with objective value

$$Q(w^*) = \frac{\mu^2}{2\gamma\sigma^2} \tag{3}$$

Also observe that the objective is zero at $w = 0$ and $w = 2w^*$ and is positive for $0 < w < 2w^*$.

The investor does not know the true values of μ and σ^2. She gets advice from two experts who have different opinions: (μ_1, σ_1^2) and (μ_2, σ_2^2). The investor is convinced that these are the only possible values for the true parameters. She is, however, not able to make a quantitative assessment of which parameter set (μ_j, σ_j^2) is more likely.

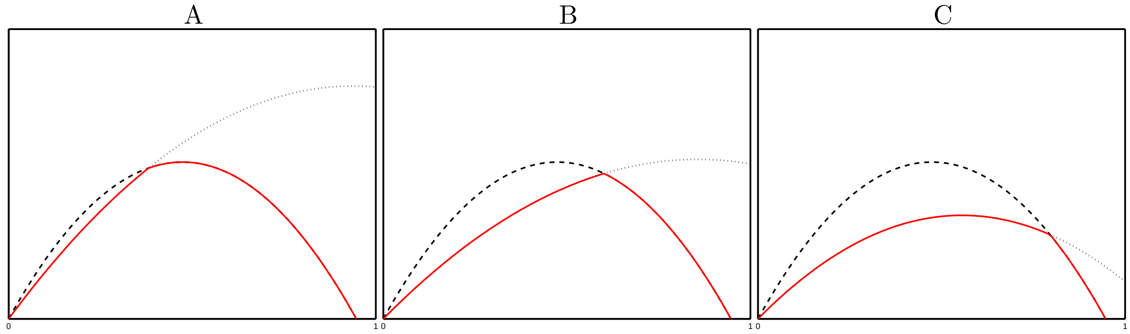
Suppose she considers a robust investment decision and maximizes the worst case utility

$$Q_R(w) = \min_{j=1,2} Q_j(w). \tag{4}$$

Before explicitly solving the robust portfolio problem, let us first distinguish different cases for (μ_j, σ_j^2). First, if expert 1 is more optimistic than expert 2, i.e. she expects higher returns with lower risk, the worst case utility coincides with the utility function of expert 2. The robust optimal portfolio is the optimal portfolio according to expert 2. The opinion of expert 1 does not affect the decision at all. Conversely, if expert 2 is more optimistic than expert 1 we have the mirror image with only the advice of expert 1 being relevant.

The interesting case is when $\mu_1 > \mu_2$ and $\sigma_1^2 > \sigma_2^2$ (or both inequalities reversed). Figure 1 depicts the mean-variance functions $Q_j(w)$ $(j = 1, 2)$. Since $Q_j(0) = 0$, both functions start at the origin. The objective function of the robust investor is the minimum of $Q_1(w)$ and $Q_2(w)$. The optimal portfolio is located at the maximum of the robust objective function.

Figure 1: Robust Portfolio Choice.



The figure shows the mean-variance objective functions implied by the advice of two experts: expert one (dashed line) with largest expected returns and variances and expert two (dotted line). The solid line is the minimum of the two objective functions and presents the robust objective function. The horizontal axis shows the portfolio weight of the risky asset, the vertical axis shows the value of the mean-variance objective. Panel A depicts the situation where the robust optimum coincides with the optimum of expert one, in panel C with the optimum of expert two, and in panel B with the portfolio for which the objective values for both experts are equal.

Depending on the parameter values, three situations can occur as shown in figure 1. The relevant optimum may be the maximum according to expert 1 (panel A), the maximum according to expert 2 (panel C) or the point where $Q_1(w)$ and $Q_2(w)$ cross (panel B). Which situation will apply? This depends on the mutual relations between the means and variances proposed by the experts. More specifically it depends on the quotients $c_1 = 2\sigma_1^2/(\sigma_1^2 + \sigma_2^2)$ and $c_2 = 2\sigma_2^2/(\sigma_1^2 + \sigma_2^2)$. If $\mu_1 < c_1\mu_2$ then the robust portfolio coincides with the maximum according to expert 1; if $\mu_2 < c_2\mu_1$ then the robust portfolio coincides with the maximum according to expert 2; the portfolio is a convex combination of the optimal portfolios suggested by the experts if $c_1\mu_2 \leq \mu_1 \leq \frac{1}{c_2}\mu_2$. Theorem 2 in the appendix quantifies this convex combination.

The graphs show that a robust portfolio is not necessarily an extremely conservative portfolio. In the middle panel of figure 1 the robust portfolio $w_R = w_{12}$ allocates a larger share to the risky asset than w_1^*.

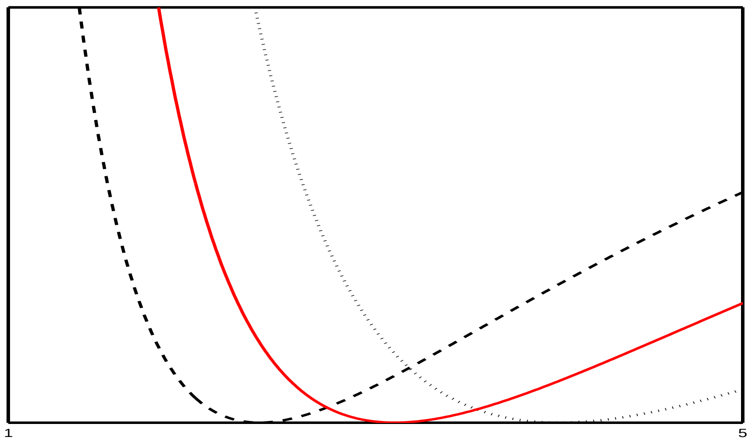
The loss function (2) for the single-asset example simplifies to

$$L(\mu, \sigma^2|w) \quad = \quad \frac{\mu^2}{2\gamma\sigma^2}\left(1 - \frac{\gamma\sigma^2 w}{\mu}\right)^2. \tag{5}$$

It is more convenient to consider the relative utility loss

$$\tilde{L}(\mu/\sigma^2|w) = \frac{L(\mu, \sigma^2|w)}{Q(w^*|\mu, \sigma^2)} = \left(1 - \frac{\gamma w}{\mu/\sigma^2}\right)^2, \tag{6}$$

Figure 2: Expected Loss



The figure shows the relative loss functions $\tilde{L}(\mu/\sigma^2)$ (vertical axis) as a function of μ/σ^2 (horizontal axis) for three portfolios: the optimal portfolios according to the advice of experts 1 and 2 (dashed lines) and the robust portfolio $w_R = w_{12}$ (solid line).

which is solely a function of μ/σ^2.

A typical situation is depicted in figure 2 which shows the loss function for the portfolios w_1^*, w_2^* and $w_R = w_{12}$ for the same parameters as figure 1. None of the three portfolios dominates the others for all parameter values. For small μ/σ^2 portfolio w_1^* is best, for large μ/σ^2 the best portfolio is w_2^*. The robust portfolio is best for intermediate parameter values. The worst robust portfolio performance is, by construction, at least as good as the worst case performance of either of its constituent models.

Figure figure 2 also shows that the robust portfolio never has a larger loss than both portfolios w_1^*, w_2^*. Indeed, w_{12} is a strictly convex combination of w_1^* and w_2^* and the loss functions (5) and (6) are strictly convex functions. This implies that the (relative) utility loss of the robust portfolio is strictly smaller than the largest loss of the two alternatives. Moreover, the robust portfolio may *outperform*, as in figure 2, both alternatives for a range of parameter values that fall in-between the advices of the experts. When we consider that the different expert advices follow from alternative beliefs about return models and that these different beliefs possibly outline a set of parameter values which contains the true parameter values (μ/σ^2), the robust portfolio is likely to outperform either of the portfolios based on a single advice.

We are unable to derive the analytical solution for the general robust portfolio choice problem with multiple assets and experts who propose general alternative return model priors. Yet for our empirical experiment we can compute the solution numerically (see e.g. Rustem et al. (2000)).

3 Model priors

So far we have not been explicit on how the experts arrive at their estimates (μ_j, Σ_j). In this section we describe the advice of the experts.

Each expert maintains a return model prior. A model prior is an established concept in Bayesian decision making. It is a subjective view on the distribution of alternative (plausible) models. The domain of the distribution conveys the variety of return models considered by the expert. If the expert confides in one particular model, the expert is a dogmatic believer and even sample information cannot change the investor's mind. We assume that experts have non-dogmatic return model priors. This means that the expert combines her return model prior with the sample information. Depending on the expert's confidence in the return model prior and the strength of sample information, she obtains a posterior return model by applying Bayes rule[1]. This posterior return model is the expert's advice to the investor. The investor relies - faute de mieux - on the advice of all consulted experts.

We test the performance of a robust approach empirically in the context of international portfolio choice. We take the point of view of a British mean-variance investor. She may invest in 81 portfolios from 9 European countries: Belgium, France, Germany, Italy, Netherlands, Spain, Sweden, Switzerland and Great-Britain. For each country the investor considers the market portfolio and 8 value and growth portfolios. The value and growth portfolios are based on four valuation ratios: book-to-market, earnings-to-price, cash earnings-to-price and dividend yield. The relevant asset returns for the British investor are in British pounds.

We continue this section by considering some typical model priors for this type of portfolio choice. These priors will serve as a basis for the expert advice in our empirical experiment. Firstly we consider a return model prior which hardly imposes any structure. This resulting 'unstructured' model suffers - by definition - no model misspecification. On the other hand, the absence of model structure makes precise estimation difficult. The may lead to considerable estimation uncertainty. Alternatively we consider some priors which introduce structure in the model: the Capital Asset Pricing Model (CAPM), the Fama & French factor model, and the international asset pricing model (IAPM). We derive the corresponding posterior models.

Unstructured model

The input for the mean-variance optimization for a British investor consists of: 1) the expected excess return vector in British pounds and 2) the return covariance matrix denominated in British pounds. Suppose the investor consults a single expert. Assume that this expert knows the true covariance matrix Σ and uses a non-informative prior to estimate the expected return vector. Then the posterior estimate of the expected return vector is the sample mean $\bar{y}$. Because the expert uses a non-informative prior, she does not impose any structure on the returns. Therefore model misspecification is no issue. On the other hand, due to the absence of structure many observations are required for a precise estimate of the expected return vector. If only a small number of observations is available, then there will be estimation uncertainty.

The expert's advice to the investor is $(\bar{y}, \Sigma)$. If the investor chooses the optimal portfolio according to this advice, her loss function becomes[2]

$$L(\mu|w_{MV}) = \frac{1}{2\gamma}(\mu - \bar{y})\Sigma^{-1}(\mu - \bar{y}) \tag{7}$$

So far we always conditioned on the advice. Integrating over the sampling variation the expected loss is

$$\mathsf{E}[L(\mu|w_{MV})] \quad = \quad \frac{1}{2\gamma}\mathsf{E}[tr(\Sigma^{-1}(\mu - \bar{y})(\mu - \bar{y})')] = \frac{N}{2\gamma T}. \tag{8}$$

The expected loss is not equal to zero. This unstructured model is - by definition - not misspecified. Hence the loss must result from estimation uncertainty. Note that the expected loss is constant and does not depend on the actual parameters (μ, Σ). A large investment universe (N) and a small number of observations increase the expected loss.

When the covariance matrix is unknown the expert uses the sample covariance matrix $\hat{\Sigma}$ as an estimate. In this case we cannot find a closed-form expression for the expected loss. Instead we compute it numerically in our empirical experiment.

[1]An additional virtue of non-dogmatic priors is that the posterior models of alternative experts will converge as more observations are added.

[2]In case we assume that the covariance matrix is known we omit Σ as an argument in the loss function.

Structured models

Experts advocate alternative structured models to explain international asset returns. We will consider four of these models: the CAPM with a global market index, the international Fama & French factor model, the IAPM with foreign exchange rate factors and a model, denoted by IFF, integrating all factors in the previous models.

The CAPM, developed primarily by Sharpe (1964), Lintner (1965) and Mossin (1966), employs one factor to explain excess returns: the excess market return. In this case we use the global market return denominated in British pounds.

Fama and French (1998) observe that value stocks with a high ratio of book-to-market equity have higher returns than growth stocks (with a low ratio of book-to-market equity) in markets around the world. Building on this observation Fama & French suggest the value premium as an additional factor to explain excess returns. In this case, the value premium is the difference in returns of stocks with high and low book-to-market equity. Also this additional factor will be denominated in British pounds.

The IAPM assumes that international markets are integrated. The global market risk premium (CAPM) and the exchange risk premium determine the excess returns. In this case we include the exchange rate effects of the German Mark and the US dollar. The corresponding exchange rate factors are defined as the (British pound to foreign currency) exchange rate return plus the difference between the foreign and domestic interest rate (denominated in their domestic currencies).

Each of these models will serve as a prior advocated by a particular expert. All these models use a linear factor structure for expected returns. The CAPM uses a single factor, the Fama & French model two factors, the IAPM three factors and the IFF model uses four factors. Let us consider the general model prior with k observed factors x_{jt} as described by Pàstor and Stambaugh (2000).

A k-factor model for excess returns y_t is

$$y_t = \alpha + Bx_t + u_t, \qquad \mathsf{E}[u_t u_t'] = D \tag{9}$$

where y_t and α are N-vectors of excess and exceptional returns respectively, B is a $(N \times k)$ matrix, x_t a vector of length k, and u_t contains error terms with mean zero and $(N \times N)$ covariance matrix D. The factor model is

$$x_t = \nu + e_t, \qquad \mathsf{E}[e_t e_t'] = \Psi \tag{10}$$

with ν a k-vector with factor means and e_t a vector of shocks with mean zero, uncorrelated with u_t, and covariance matrix Ψ.

A dogmatic believer of a pure factor model assumes $\alpha = 0$ and D diagonal. Expected excess returns under the factor model are

$$\mu = B\nu \tag{11}$$

and the covariance matrix

$$\Sigma = B\Psi B' + D. \tag{12}$$

However we consider non-dogmatic experts. Non-dogmatic experts combine their priors with sample information. We consider the same prior as Pàstor and Stambaugh (2000). For convenience we repeat the prior in appendix B. This prior features a strong belief in $\alpha = 0$ but is less informative for the residual's covariance matrix D. In particular we use $\sigma_\alpha = 0.1\%$ and $df = 82$ in the prior (29).

The experts combine their model prior with sample information to obtain a posterior return model. This posterior is their advice to the investor. Appendix B computes the posterior. The posterior parameters are indicated by a tilde: $\tilde{\alpha}$, $\tilde{B}$, $\tilde{\nu}$ and $\tilde{\Sigma}$.

When using a factor model, the MV investor chooses the portfolio

$$w^*_{FM} = \frac{1}{\gamma}\tilde{\Sigma}^{-1}(\tilde{\alpha} + \tilde{B}\tilde{\nu}) \tag{13}$$

If α, B, Σ and Ψ are known such that $w^*_{FM} = \frac{1}{\gamma}\Sigma^{-1}B\tilde{\nu}$, we can simplify the loss function (appendix A)

$$\mathsf{E}[L(\mu, \Sigma|\bar{y})] = \frac{1}{2\gamma}(\mu - B\nu)'\Sigma^{-1}(\mu - B\nu) + \frac{1}{2\gamma T}\mathsf{tr}(I + D(B\Psi B')^{-1})^{-1}$$

Observe that the second term in the last expression is smaller than $k/2\gamma T$. The second term is therefore also smaller than (8). This indicates a smaller loss through estimation errors compared to the unstructured model. The additional, positive term arises from model misspecification. The total loss can be both larger or smaller than the loss of the unstructured model. The restrictions of the factor model lead to less estimation risk, but when the factor assumption differs from the actual means, model misspecification will increase expected loss.

4 Empirical performance

4.1 Data

For our empirical experiment we need the historical excess returns (in British pounds) on the 81 portfolio from 9 countries. The historical returns on the portfolios (in US Dollars) are obtained from Fama & French database. This is available from the homepage of Kenneth French[3]. The 9 countries we use are all those with a complete dataset of historical monthly returns from January 1975 to December 2001. For each country we consider the market portfolio, four value portfolios and four growth portfolios. The value and growth portfolios are based on four valuation ratios: book-to-market, earnings-to-price, cash earnings-to-price and dividend yield. We need to convert these data to reflect the British investor's situation. For this purpose we use the British pound - US dollar exchange rate available from the Italian Foreign Exchange Office (UIC, *http://www.uic.it/*). To obtain excess returns we subtract the British riskfree interest rate. This interest rate -we use the London Euro-Currency interest rate - is available from Datastream (*http://www.datastream.com*).

We also need the historical factor values. We obtain the global market return (in US dollars) from the Fama & French database. The factor associated with the global value premium is the difference between the returns of high and low book-to-market (global) portfolios. We also obtain these data from the Fama & French dataset. The market premium and value premium factors are converted from US dollars to British pounds by using the exchange rate available from the UIC.

Finally we need the factors associated with the exchange rate risk premium. These factors consist of the exchange rate returns plus the difference between the foreign and domestic interest rate. We consider the British pound-German Mark and British pound-US Dollar exchange rate risk. We obtain the exchange rates from the UIC. The German interest rate (London Euro-Currency interest rate) is obtained from Datastream. We obtain the US riskfree interest rate from the Fama & French dataset.

4.2 Bootstrap experiment

Our aim is to compare the empirical ex-post performance of three types of portfolios: (i) the optimal portfolio based on the unstructured return model, (ii) the optimal portfolios according to the experts with alternative, structured return models and (iii) the portfolio of a robust investor who considers all these different structured return models.

The comparison is based on a bootstrap experiment. In the bootstrap experiment we repeatedly draw sets of T random observations from the Fama & French dataset (adapted to the British situation). On each set

[3] *http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html*

of observations we estimate the different return models. Based on the estimated return models, we construct the alternative optimal portfolios.

We will evaluate the portfolio performance ex-post on the entire dataset. Let us refer to the sample mean and covariance matrix taken over the entire dataset as the null model (μ_0, Σ_0). We are interested in the ex-post performance of a portfolio w given by

$$Q_0(w) = \mu_0' w - \frac{1}{2}\gamma w' \Sigma_0 w \tag{14}$$

with the risk aversion parameter[4] set to 5. We do not use the Sharpe ratio to measure performance: it does not take leverage into account and therefore is not an adequate measure to evaluate mean-variance performance in the presence of uncertainty.

An actual portfolio w is a function of the model prior as provided by an expert, the sample data and, for the robust portfolio, the preference for robustness. Given a bootstrap sample of T observations from the Fama & French dataset, w_u denotes the portfolio based on the sample estimates, the portfolio w_c is based on the CAPM prior (and the sample), w_d on the Fama & French prior, w_e in the IAPM prior, w_f is based on the model prior containing all factors. We will use j to refer to the acronyms c, d, e and f corresponding to the different models and $j = 0$ refers to the null model.

We are interested in comparing the *average* performance of the alternative portfolios. Therefore we consider K bootstrap samples and compute the average ex-post performance of the alternative portfolios. Apart from the ex-post performance of each portfolio we store, among others the ex-ante expected performance according to each model prior, the ex-post expected loss and some portfolio characteristics such as total risky investment and its cross-sectional standard deviation. We use w_j^k, $\hat{\mu}_j^k$ and $\hat{\Sigma}_j^k$ to denote the optimal portfolio, estimated expected return vector and the estimated covariance matrix at bootstrap $k \in \{1, .., K\}$ under return model $j \in \{u, c, d, e, f\}$. The optimal portfolio under the null model is denoted w_0. We report the statistics

ex-ante

excess return	r_j^k	$=$	$w_j^{k\prime} \hat{\mu}_j^k$
variance	$(s_j^k)^2$	$=$	$w_j^{k\prime} \hat{\Sigma}_j^k w_j^k$
mean-variance utility	Q_j^k	$=$	$r_j^k - \frac{1}{2}\gamma (s_j^k)^2$
Sharpe ratio	Sh_j^k	$=$	r_j^k / s_j^k

ex-post

excess return	$r_0(w_j^k)$	$=$	$\mu_0' w_j^k$
variance	$s_0^2(w_j^k)$	$=$	$w_j^{k\prime} \Sigma_0 w_j^k$
mean-variance utility	$Q_0(w_j^k)$	$=$	$r_0(w_j^k) - \frac{1}{2}\gamma s_0^2(w_j^k)$
Sharpe ratio	$Sh_0(w_j^k)$	$=$	$r_0(w_j^k)/s_0(w_j^k)$
expected loss	$L_0(w_j^k)$	$=$	$Q_0(w_0) - Q_0(w_j^k)$

portfolio statistics

sum	$=$	$\iota' w_j^k$
norm	$=$	$\lVert w_j^k \rVert.$
cross-sectional st.dev.	$=$	$\sqrt{\lVert w_j^k \rVert^2 / N - (\iota' w_j^k / N)^2}.$

For the robust portfolio w_R^k we (re)define the statistics

[4] Alternative values of γ affect the size of investment but do not affect the portfolio composition or, for robust portfolio choice, the decision whether or not to invest. Therefore the results for alternative values of γ will be scaled versions of the presented results.

worst model(s)	j_{wc}^k	$=$	$arg\min_{j\in\{c,d,e,f\}} Q_j^k(w_R^k)$
expected mv	Q_R^k	$=$	$\min_{j\in\{c,d,e,f\}} Q_j^k(w_R^k)$
expected return	r_R^k	$=$	$\min_{j\in\{c,d,e,f\}} r_j^k(w_R^k)$
expected st.dev.	s_R^k	$=$	$\min_{j\in\{c,d,e,f\}} s_j^k(w_R^k)$
expected sharpe	Sh_R^k	$=$	$r_R^k/s_R^k.$

Table 1: Model misspecification

		CAPM	FF	IAPM	IFF	ROB
	$w_u = w_0$	w_c	w_d	w_e	w_f	w_r
expected performance (ex-ante)						
excess return (r)	7.31	0.13	0.58	0.17	0.63	0.13
standard deviation (s)	12.09	1.63	3.40	1.84	3.55	1.62
mean-variance utility (Q)	3.66	0.07	0.29	0.08	0.32	0.07
Sharpe ratio (Sh)	0.60	0.08	0.17	0.09	0.18	0.08
performance under null (ex-post)						
excess return (r)	7.31	0.28	0.67	0.23	0.63	0.27
standard deviation (s)	12.09	1.61	3.37	1.83	3.53	1.60
mean-variance utility (Q)	3.66	0.21	0.39	0.14	0.32	0.20
Sharpe ratio (Sh)	0.60	0.17	0.20	0.12	0.18	0.17
expected loss (L)	0.00	3.44	3.27	3.51	3.34	3.45
portfolio						
sum	1.08	0.32	0.36	0.28	0.31	0.32
cross-sectional stdev	2.17	0.08	0.19	0.17	0.26	0.09
stringent to robust approach		√		√		

Notes: The table shows the ex-ante and ex-post performance of optimal mean variance portfolios ($\gamma = 5$) which are based on alternative return models corresponding to the acronyms u, c, d, e and f and the model robust portfolio w_R estimated on the entire Fama & French 81 portfolios dataset over the period January 1975 to December 2001. As estimated on the entire dataset, the results for the unstructured return model also present the results of the null model. All numbers are, except the Sharpe ratios and portfolio characteristics, given in percentages *per month.*

Table 1 reports the expected loss of the alternative portfolios (relative to the optimal portfolio for the null model) when the models are estimated on the *entire* Fama & French dataset. It provides a measure for model misspecification in the absence of estimation uncertainty. Naturally the unrestricted model leads to zero loss as it could only suffer from estimation uncertainty. Also note that the associate optimal portfolio leads to an 7% excess return per month. This incredible investment opportunity is, among others, due to the large asset universe. The non-zero loss of the portfolios associated with the alternative (strong) model priors suggest that these models are misspecified.

The ex-ante mean-variance utility of the portfolios based on structured models does not vary much, yet we observe a specific ranking. The CAPM employs one single factor which is also contained in the other models. Hence the CAPM imposes most structure on the return model. As a consequence the CAPM is the most restrictive in terms of portfolio choice and has the lowest mean-variance utility. On the other hand, the IFF model imposes least structure and has the highest mean-variance utility. The models FF and IAPM use a subset of the factors in the IFF model but include, at least, the global market return of the CAPM. Consequently the mean-variance utility falls in between that of the CAPM and IFF model. The robust approach is, by definition, ex-ante most conservative. In this case it leans on the CAPM which is ex-ante most restrictive. This dependence of portfolio choice under uncertainty on the most restrictive return model prior is also observed for a Bayesian approach (Lindleys paradox).

Notwithstanding the strong structure which the CAPM imposes, we will see that the CAPM does not lead to the largest effective model misspecification. With effective we want to indicate the consequences of model misspecification to portfolio choice. We measure the effective model misspecification by ex-post expected loss. The ex-post expected loss of the portfolio choice based on the CAPM is smaller than the portfolio choice based on the IAPM. We conclude that a better ex-ante fit to the data of a less restrictive model, i.e. comparing IAPM to CAPM, does not necessarily lead to better portfolio allocation. The model robust approach takes all four structured models into account and anticipates the most conservative performance. The robust portfolio actually has an smaller effective model misspecification than the worst of the structured portfolios. We will see that this phenomenon looms even larger in a practical situation where uncertainty plays a role.

We now turn to a more practical setting where we merely have a limited sample of observations to estimate our return model and make our portfolio choice. In other words, we introduce estimation uncertainty. Table 2 reports the averages and standard deviations of the above statistics for a bootstrap experiment with $K = 10,000$ samples and 12.5 years of observations. The table also reports the fraction of bootstrap samples for which the corresponding portfolio is active, i.e. with a non-zero investment in risky assets. The adjective 'stringent' in the table refers to the fraction of the bootstrap samples for which each return model is worst case for the robust portfolio. Moreover, $P(Q_0(w_j) \geq Q_j)$ denotes the fraction of the bootstraps for which the ex-post mean-variance performance of portfolio w_j is at least its ex-ante expectation. This statistic serves as measure of robustness of the expected performance.

Table 2: Portfolio performance of a model robust investor

		CAPM	FF	ICAPM	IFF	ROB
	w_u	w_c	w_d	w_e	w_f	w_r
expected performance (ex-ante)						
excess return (r)	51.01	0.28	1.05	0.62	1.44	0.28
	(16.26)	(0.33)	(0.64)	(0.46)	(0.74)	(0.33)
standard deviation (s)	31.58	1.98	4.35	3.28	5.18	1.98
	(4.78)	(1.31)	(1.42)	(1.28)	(1.38)	(1.31)
mean-variance utility (Q)	25.50	0.14	0.52	0.31	0.72	0.14
	(8.13)	(0.16)	(0.32)	(0.23)	(0.37)	(0.16)
Sharpe ratio (Sh)	1.58	0.10	0.22	0.16	0.26	0.10
	(0.24)	(0.07)	(0.07)	(0.06)	(0.07)	(0.07)
performance under null (ex-post)						
excess return (r)	19.47	0.27	0.74	0.25	0.72	0.27
	(7.57)	(0.25)	(0.40)	(0.33)	(0.48)	(0.25)
standard deviation (s)	96.93	2.32	5.66	4.62	7.27	2.32
	(33.61)	(1.58)	(2.08)	(1.94)	(2.31)	(1.58)
mean-variance utility (Q)	-243.62	0.07	-0.17	-0.38	-0.73	0.07
	(234.47)	(0.15)	(0.49)	(0.51)	(0.82)	(0.15)
Sharpe ratio (Sh)	0.21	0.11	0.13	0.05	0.10	0.11
	(0.06)	(0.08)	(0.05)	(0.07)	(0.05)	(0.08)
expected loss (L)	247.28	3.58	3.83	4.03	4.39	3.59
	(234.47)	(0.15)	(0.49)	(0.51)	(0.82)	(0.15)
ex-ante to ex-post						
$P(Sh_0(w_j) \geq Sh_j)$	0.0	56.6	14.9	9.7	2.6	56.6
$P(Q_0(w_j) \geq Q_j)$	0.0	55.5	14.0	9.1	2.3	55.7
portfolio						
active	100	100	100	100	100	100
sum	2.23	0.33	0.35	0.28	0.29	0.33
cross-sectional stdev	16.57	0.25	0.73	0.70	1.06	0.26
stringent		98.89	12.87	58.26	5.47	

Notes: The table shows the average ex-ante and ex-post performance of optimal mean variance portfolios ($\gamma = 5$) which are based on alternative return models corresponding to the acronyms u, c, d, e and f and the model robust portfolio w_r over 10.000 bootstrap samples of 150 random observations from the Fama & French 81 portfolios dataset over the period January 1975 to December 2001. The ex-post performance is evaluated on the null model. All numbers are averages (stdev's) over the bootstraps, and are, except the Sharpe ratios and portfolio characteristics, given in percentages per month.

The portfolio associated with the unstructured model anticipates an incredible performance of 26% per month with a monthly excess return of 51%. This follows from a very favorable estimate of the investment opportunity set. However the extremely favorable portfolios within this set are typically the result of estimation errors. The estimation errors are due to the combination of a small number of observations and a large asset universe (81 assets). Many assets implies many model parameters. The small number of observations makes precise parameter estimation difficult. The problem is that the optimization will select the extreme, typically leveraged, portfolios. However the extremely favorable investment opportunities actually do not exist (but are based on estimation errors). The selected portfolios are therefore suboptimal and (may) have a poor actual portfolio performance. Jobson and Korkie (1981) and Michaud (1998) describe such portfolio allocations as error-maximizing portfolios.

This also happens in the empirical experiment. There is considerable estimation uncertainty (compare table 2 to table 1 without estimation uncertainty). As a result the actual (ex-post) performance of the unstructured portfolio is a deplorable -244%. The sum of risky investment is 2.23. Hence, the portfolio is leveraged $(2.23 > 1)$.

The same phenomenon can be found in the performance of the structured portfolios. The portfolios associated with the least informative priors anticipate the highest performance but actually exhibit the poor ex-post performance. The IFF portfolio reports the worst performance among the structured portfolios. On the other hand, the CAPM portfolio is most structured. Yet is has the best actual (ex-post) performance of the structured porftfolios.

The portfolios associated with strong prior beliefs in alternative structured return models have moderate risky asset allocations. These structured portfolios have much better ex-post performance than the unstructured portfolio. The structured return models decrease estimation uncertainty and, in terms of ex-post portfolio performance, this decrease in estimation uncertainty outweighs the cost of the model misspecification.

We observe that the model robust portfolio performs very well compared to the other portfolios. The robust portfolio mimics the most conservative of the structured portfolios. Table 2 indicates that the CAPM (99%) and the IAPM model (58%) are typically stringent for the robust approach. These percentages do not sum to one as a combination of structured models generally determines the worst case. Again we observe the crucial dependence of portfolio choice under uncertainty on the most restrictive return model prior.

The reliability, measured by the fraction of bootstrap samples for which the portfolio attains, ex-post, the ex-ante expected performance is small: zero for the unstructured model, 2.6% for the IFF-based portfolio and 56% for the CAPM-based portfolio and the robust portfolio. Hence we cannot label the ex-ante performance estimates as robust. Estimation uncertainty and, for the robust approach, a limited number of alternative models lead to ex-ante overestimation of the performance. Note that the more reliable portfolios - the portfolios based on the CAPM and the robust portfolio - also have the best actual (ex-post) expected performance. This suggests that maximizing expected performance and robustness need not be conflicting objectives for portfolio choice under uncertainty.

Let us consider what happens when uncertainty decreases due to an increase in the number of observations. Figure 4 shows the portfolio performance for different numbers of observations. The performance of the portfolios associated with the unstructured return model naturally improves as estimation uncertainty decreases and will eventually converge to zero expected loss. The performance of the portfolios associated with the alternative return model priors improves only slowly. On the one hand model structuring pays off in an uncertain environment and improves performance considerably compared to an unstructured model. On the other hand the strong prior leaves little room for using the sample information. Therefore the effect of model misspecification in the prior is perceived even when the number of observations is large and uncertainty is small. The robust portfolio which strongly depends on the CAPM model, reacts similarly to an increase in the number of observations.

5 Model and Estimation Uncertainty

In this section we extend the robust approach to deal with estimation uncertainty. In addition to reporting the relevant aspects of the model, i.e. (μ_j, Σ_j), each expert now also provides the decision maker with an uncertainty set (in essence a confidence interval) around these estimates. The uncertainty set $\mathcal{U}_j$ of a particular expert j will contain the parameter values that are, in the expert's believe, plausible. We consider experts who exclusively consider uncertainty in the estimator for the expected return[5].

Investors who rely on one particular expert j and aim for an estimation robust portfolio use $\mathcal{U} = \mathcal{U}_j$ and solve

$$\max_w \min_{\mu \in \mathcal{U}} Q_j(\mu, w) \tag{15}$$

with

$$Q_j(\mu, w) \quad = \quad \mu' w - \frac{1}{2}\gamma w' \tilde{\Sigma}_j w$$

The specific form of the uncertainty set depends on the expert. Common in their approaches is that they derive this set from the posterior distribution of the parameters for each of the models. The uncertainty set should contain the parameter values which are plausible considering the observations and the return model. Namely in that case does a solution to (15) indeed present a portfolio choice that is robust to all plausible parameters. We fulfil this requirement on the uncertainty set by including parameter configurations with highest posterior probability density such that their cumulative density suffices for a considerable degree of robustness, say 90%. Appendix B.1 elaborates on the precise construction of the uncertainty set for alternative experts.

Apart from investors who base their portfolio on one of the expert models $j \in \{1, .., J\}$, we also consider an investor who adopts a robust approach to the alternative expert models. The model- *and* estimation robust investor solves

$$\max_w \min_j \min_{\mu \in \mathcal{U}_j} Q_j(\mu, w). \tag{16}$$

This robust investor considers alternative return models as well as an uncertainty set of parameter values for each return model. Hence the set of alternative expected return vectors and covariance matrices which the investor considers is larger than the set considered by the purely estimation robust or purely model robust investor. The model- and estimation robust investor will therefore also be more conservative.

5.1 Bootstrap experiment

We continue with an empirical study on the performance of an estimation robust approach. We consult the same experts as in section 4. The portfolios based on the advice of each individual expert as well as the model- and estimation robust portfolio are tested in a bootstrap experiment.

We repeat the bootstrap of section 4 but now consider the optimal portfolios w_j^k for $j = \{u, c, d, e, f\}$ which solve (15), with $\hat{\mu} = \hat{\mu}_j^k$, $\hat{\Sigma} = \tilde{\Sigma}_j^k$ and $\Omega = \Omega_j^k$. We choose $\Omega = \tilde{\Sigma}/T$ and use appendix B.1 to quantify the degree of robustness θ_j such that it corresponds to a 90% confidence level. The robust expected return which is corrected for estimation uncertainty replaces the naive expected return,

$$r_j^k(w_j^k) = w_j^{k\prime} \hat{\mu}_j^k - \theta_j \sqrt{w_j^{k\prime} \hat{\Omega} w_j^k}.$$

The model- and estimation robust portfolio w_R^k solves (16). Again we evaluate all portfolios on their (ex-post) performance under the null model.

[5] Errors in means are more critical than errors in variances, and errors in variances are more critical than errors in covariances (see Chopra and Ziemba (1993))

Table 3: Portfolio performance of a model- and estimation (mean) robust investor

		CAPM	FF	IAPM	IFF	ROB
	w_u	w_c	w_d	w_e	w_f	w_R
expected performance (ex-ante)						
excess return (r)	2.64	0.07	0.10	0.01	0.05	0.01
	(4.22)	(0.14)	(0.17)	(0.05)	(0.12)	(0.03)
standard deviation (s)	5.67	0.69	0.99	0.15	0.59	0.09
	(4.55)	(0.94)	(1.05)	(0.43)	(0.86)	(0.31)
mean-variance utility (Q)	1.32	0.03	0.05	0.01	0.03	0.00
	(2.11)	(0.07)	(0.09)	(0.02)	(0.06)	(0.01)
Sharpe ratio (Sh)	0.28	0.03	0.05	0.01	0.03	0.00
	(0.23)	(0.05)	(0.05)	(0.02)	(0.04)	(0.02)
performance under null (ex-post)						
excess return (r)	3.75	0.11	0.18	0.02	0.09	0.01
	(3.79)	(0.15)	(0.21)	(0.06)	(0.16)	(0.06)
standard deviation (s)	19.04	0.81	1.31	0.22	0.87	0.12
	(19.52)	(1.12)	(1.44)	(0.63)	(1.28)	(0.42)
mean-variance utility (Q)	-14.82	0.06	0.08	0.00	0.03	0.01
	(57.71)	(0.08)	(0.12)	(0.04)	(0.10)	(0.04)
Sharpe ratio (Sh)	0.21	0.17	0.16	0.13	0.12	0.14
	(0.06)	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)
expected loss (L)	18.48	3.60	3.57	3.65	3.63	3.65
	(57.71)	(0.08)	(0.12)	(0.04)	(0.10)	(0.04)
ex-ante to ex-post						
$P(Sh_0(w_j) \geq Sh_j)$	43.7	91.3	88.8	92.0	87.8	98.5
$P(Q_0(w_j) \geq Q_j)$	29.7	90.9	87.3	91.8	85.7	99.0
portfolio						
active	91	62	72	23	61	14
sum	0.42	0.13	0.08	0.01	0.03	0.01
cross-sectional stdev	3.25	0.09	0.17	0.03	0.13	0.02
stringent		2.94	8.61	12.36	5.78	

Notes: The table shows the average ex-ante and ex-post performance of optimal, estimation robust, mean variance portfolios ($\gamma = 5$) which are based on alternative return models corresponding to the acronyms u, c, d, e and f and the model robust portfolio w_R over 10.000 bootstrap samples of 150 random observations from the Fama & French 81 portfolios dataset over the period January 1975 to December 2001. All portfolios are estimation robust to uncertainty in the expected return vector. Moreover, the second column corresponding to w_u presents the unstructured portfolio which is also estimation robust to uncertainty in the covariance matrix.
The ex-post performance is evaluated on the null model. All numbers are averages (stdev's) over the bootstraps, and are, except the Sharpe ratios and portfolio characteristics, given in percentages per month.

Table 3 reports the performance of estimation robust portfolios for a bootstrap experiment with $K = 10.000$ samples with 12.5 years of observations each. The reported portfolios are robust to estimation uncertainty in the expected return vector.

Most striking is the the decrease in the absolute numbers when comparing tables 3 and 2. The ex-ante performance of the unstructured portfolio decreases from 25.5% to 1.3% for an estimation robust approach. The ex-ante performance for the CAPM decreases from 0.14% to 0.03%. Contrary to the decreased ex-ante expected performance, the ex-post performance of the estimation robust portfolios is considerably better. The ex-post expected loss corresponding to the unstructured model reduces to less than 5% of the original loss when the investor is estimation robust. The loss associated with the IFF is reduced 18% due to the estimation robust approach and the ex-post average mean-variance performance is now positive. On the other hand, the average loss of the CAPM increases slightly. The increase is (partly) due to a less active strategy: the CAPM investor is only active 62% of the time. When active, this investor obtains an average actual performance of $0.06\%/0.62 = 0.1\%$ (and an excess return of 2.15% per year). This is better than the average performance without estimation uncertainty.

The model- and estimation robust portfolio leans on the IAPM. Apparently the IAPM model is most restrictive. This is due to its relatively large estimation uncertainty. On the one hand, the IAPM imposes less model structure than the CAPM. Therefore, without estimation uncertainty, the IAPM would be less restrictive to robust portfolio choice than the CAPM. On the other hand, the IAPM suffers model estimation uncertainty than the CAPM. An estimation robust approach takes estimation uncertainty into account. Larger estimation uncertainty is more restrictive for the portfolio choice. In this case the combination of model specification and estimation uncertainty makes the IAPM the most restrictive model.

Also note that the Fama & French model presents a good mix of model structuring and estimation uncertainty. On the one hand, the value premium seems to contribute to model specification: the ex-ante performance for the FF portfolio is better than the performance of the CAPM portfolio. On the other hand, the number of parameters remains small and estimation uncertainty is limited: the ex-ante performance for the FF portfolio is better than the performance of the IFF portfolio.

Estimation robust portfolios are often passive portfolios: when a sample of 150 observations does not show significant investment opportunities, the estimation robust investor refrains from risky investment. For example an investor who relies exclusively on the CAPM model and employs an estimation robust approach, will only invest 62% of the time. Moreover when the investor enters the risky market, her total risky investment (long or short) is reduced as can be seen from the portfolio sum and the cross-sectional standard deviation of the portfolio. These measures, i.e. a preliminary test which considers whether sufficient investment opportunities exist and a reduction in risky investment whenever entering the risky market, contribute to the reliability of the portfolio choice.

An estimation robust approach considerably improves the reliability, measured by the fraction of bootstrap samples for which the portfolio attains, ex-post, the ex-ante expected performance. The reliability of a naive approach to estimation uncertainty is negligible (see table 2). The reliability of the portfolio based on the unstructured return model remains small (29%) when only estimation uncertainty in the expected returns is considered. The reason is that an estimate of the covariance matrix with dimensions 81×81 which is based on the limited number of observations is also highly uncertain. When we also adopt a robust approach for this uncertainty in the covariance matrix (with an approximate uncertainty set as outlined in appendix B.1), the reliability increases and the robust portfolio tends to zero investment in risky assets. This does not mean that an estimation robust approach is hopelessly conservative. This passive investment strategy is merely a result of too little information to make a reliable portfolio decision.

Factor models do not severely suffer from estimation uncertainty in the covariance matrix. These models effectively reduce the number of parameters to be estimated by imposing structure on the covariances. Consequently the number of parameters to be estimated for the corresponding covariance matrices is much smaller.

Figure 4 shows the portfolio performance for various degrees of uncertainty. The improved performance of the portfolio associated with the unstructured portfolio is most striking: the loss is reduced considerably and the rate of convergence to zero expected loss is increased. This effect is more pronounced when we account for estimation uncertainty in the covariance matrix. The losses corresponding to the structured models are also reduced but the effect of the strong model prior dominates the information in the data which would drive the loss quickly to zero.

When the number of observations increases and consequently uncertainty decreases the investor, whether relying on structured or unstructured models, becomes more active. With 120 observations, an investor who relies on the CAPM is active 4% of the time. With 300 observations, the same investor is active 98% of the time. Approximately the same holds for the estimation and model robust investor.

6 Model robust versus Estimation robust

The study on the effects of estimation robust portfolio choice allows us to compare the importance of a robust approach to model uncertainty and estimation uncertainty. Table 1 reports the expected loss due to model misspecification for each structured model. Table 2 shows the additional effect of estimation uncertainty for each model.

Let us start with the portfolio choice based on the Fama & French model. Without taking into account any uncertainty, the Fama & French model expects ex-ante performance 0.52 but realizes ex-post -0.17. Supplementing the Fama & French model with alternative models and using a model robust approach increases the ex-post performance to 0.07. On the other hand, an estimation robust approach based on the Fama & French model leads to a further increase in performance to 0.08. This is slightly larger than the increase due to model robustness.

Only the CAPM model shows a different pattern. Model robustness does not improve the performance. Moreover, the structure imposed by the CAPM reduces the effect of estimation uncertainty. This leads to a CAPM portfolio with a positive performance. Estimation robustness only makes the CAPM investor more passive. This effect may be the result of model misspecification in combination with estimation uncertainty.

These observations differ from the conclusion of Avramov (2002) who studies a Bayesian decision maker. He concludes that accounting for model uncertainty is more important than considering estimation uncertainty. A conceivable explanation for the difference between Avramov's (2002) conclusion and the observation in this study is the diversity of (structured) return models which is considered. Naturally the effect of a robust approach increases when we include models which lead to a greater dispersion of expected return forecasts. This could lead to an improvement in the performance of a model robust approach. It would be interesting to study the robust approach in the context considered by Avramov (2002). One could use the posterior model probabilities to select the alternative models for the uncertainty set.

The results also suggest two ways to reduce the effects of estimation uncertainty: imposing model structure and using an estimation robust approach. These approaches are not mutually exclusive. Model structuring deals with limited information by effectively reducing the number of parameters that need to be estimated. The drawback is that it imposes a model structure which may be wrong. When prior knowledge about the return model is not available one is thrown back on the unstructured model. Yet, with an estimation robust approach the resulting portfolio performance may be improved. An advantage of the estimation robust approach is that it adapts to the extend of uncertainty. The uncertainty set of parameters that need to be taken into consideration in the robust approach becomes smaller as more information is available and there is less uncertainty.

7 Conclusion

We have shown that the model robust portfolio is the optimal mean-variance portfolio corresponding to an endogenously determined combination of alternative return models. For some special cases we also showed that the robust portfolio is a convex combination of the optimal portfolios corresponding to the alternative models. In that case, the robust portfolio has smaller expected loss than the worst of the alternative portfolios. Moreover, the robust portfolio has least expected loss for a range of models in-between the alternative models. When we consider that experts propose sensible return models, then alternative expert's beliefs typically circumscribe a set of models which contains the true model. In this event, the robust portfolio is likely to outperform the optimal portfolios associated with alternative models.

We considered robust portfolio choice in an empirical setting with alternative return models. The models are based on strong beliefs in alternative asset pricing models. We observe that the model robust portfolio typically leans on the most restrictive return model: This is the CAPM if estimation uncertainty is not considered. The IAPM is most restrictive in the estimation robust approach (due to its larger estimation uncertainty).

The robust investor refrains from risky investment when investment opportunities are not (estimated) significant. Moreover when entering the risky market, the investor adapts the total risky investment to the level of uncertainty. We stress that it is not the aim of a robust investor to be conservative. Indeed, when the information is insufficient to be confident about significant investment opportunities, the robust investor remains passive. But if she is given an (sufficiently) accurate return model specification and more precise parameter estimates, the investor will not remain passive

We have seen that model robustness improves the portfolio performance but we also noted that a truly model robust portfolio needs to be based on more than just a few alternative return models.

A robust approach to estimation uncertainty also may be worthwhile: the estimation robust portfolios mostly outperform their non-robust unstructured, structured or even model robust counterparts. Estimation robustness adapts to uncertainty and therefore does not make portfolio choice too conservative: When the number of observations is small and estimation uncertainty is large, estimation robustness considerably improves the performance of the unstructured portfolio. When uncertainty decreases, the portfolio converges, faster than the structured portfolios, to the truly optimal portfolio.

A Proofs and Theorems

Theorem 1 *Consider a robust decision maker who obtains advice (μ_j, Σ_j) from J experts. Her optimal portfolio is*

$$w_R^* = \left(\sum_{j=1}^{J} \lambda_j \Sigma_j\right)^{-1} \left(\sum_j \lambda_j \mu_j\right)$$

where λ_j are constants satisfying $0 \leq \lambda_j \leq 1$ and $\sum \lambda_j = 1$.

Proof A formal representation of the robust portfolio problem is

$$\max Q_R, \tag{17a}$$

$$\text{subject to} \quad Q_j(w) - Q_R \leq 0, \qquad (j = 1, \ldots, J). \tag{17b}$$

Kuhn-Tucker conditions for the optimal portfolio are

$$1 - \sum_j \lambda_j = 0, \tag{18a}$$

$$\sum_j \lambda_j Q_j'(w) = 0, \tag{18b}$$

$$\lambda_j (Q_R - Q_j(w)) = 0, \qquad (j = 1, \ldots, J), \tag{18c}$$

$$\lambda_j \geq 0, \qquad (j = 1, \ldots, J), \tag{18d}$$

with $\lambda_j \geq 0$ a set of Lagrange multipliers and $Q_j'(w) = \mu_j - \gamma \Sigma_j w$ the first order derivative of the objective function with expert j's parameters. From (18a) and (18d) we deduce that the Lagrange multipliers must satisfy $0 \leq \lambda_j \leq 1$. From (18b) and the linearity of $Q_j'(w)$ we immediately obtain the result stated in the theorem. □

Theorem 2 *Suppose $\mu_1 > \mu_2$ and $\sigma_1^2 > \sigma_2^2$. Let w_j^* be the optimal portfolio according to expert j's advice. Further define $w_{12} > 0$ as the portfolio weight for which $Q_1(w) = Q_2(w)$. A robust investor will have an optimal portfolio*

$$w_R = \lambda w_1^* + (1-\lambda) w_2^* \tag{19}$$

with

$$\lambda = \begin{cases} 0 & \textit{if} \quad \mu_2 \le \frac{2\sigma_2^2}{\sigma_1^2+\sigma_2^2}\mu_1, \\ \frac{2(\mu_1-\mu_2)\sigma_1^2\sigma_2^2 - \mu_2\sigma_1^2}{(\mu_1\sigma_2^2-\mu_2\sigma_1^2)(\sigma_1^2-\sigma_2^2)} \in (0,1) & \textit{if} \quad \frac{2\sigma_2^2}{\sigma_1^2+\sigma_2^2}\mu_1 < \mu_2 < \frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2}\mu_1, \\ 1 & \textit{if} \quad \frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2}\mu_1 \le \mu_2 < \mu_1, \end{cases} \tag{20}$$

Proof With only two experts we only need to check three cases of combinations of active constraints in the Kuhn-Tucker conditions in the proof of theorem 1. We will show that these correspond to the three possible portfolios in (19). Under the assumption $\mu_1 > \mu_2$, we have $Q_1'(0) = \mu_1 > \mu_2 = Q_2'(0)$, implying that $Q_R(w) = \min(Q_1(w), Q_2(w)) = Q_2(w)$ for small w. Since both objective functions are quadratic, the difference between them is also quadratic, and they intersect at only two points: $w = 0$ and $w = w_{12}$. Therefore

$$Q_R(w) = \begin{cases} Q_2(w) & \text{if } 0 < w < w_{12}, \\ Q_1(w) & \text{if } w \ge w_{12}. \end{cases} \tag{21}$$

The maximum of $Q_R(w)$ is either at one of the interior maxima w_1^* or w_2^*, or at the intersection point w_{12}. Differentiating $Q_j(w)$, the interior maxima are easily found as

$$w_j^* = \mu_j / \gamma\sigma_j^2. \tag{22}$$

Since the expected returns are assumed to be positive, the optimal investment is always positive. We therefore do not need to consider portfolios with $w < 0$. The intersection point w_{12} follows from

$$Q_1(w) = Q_2(w) \quad \Leftrightarrow \quad (\mu_1 - \mu_2) w = \frac{1}{2}\gamma(\sigma_1^2 - \sigma_2^2) w^2,$$

so that

$$w_{12} = \frac{2(\mu_1 - \mu_2)}{\gamma(\sigma_1^2 - \sigma_2^2)}. \tag{23}$$

Let us first check if w_1^* can be a valid optimum. It is only a valid optimum if $w_1^* > w_{12}$, as otherwise $Q_R(w) = Q_2(w)$. Using the expressions for w_1^* and w_{12} this leads to the inequality

$$\frac{\mu_1}{\gamma\sigma_1^2} > \frac{2(\mu_1 - \mu_2)}{\gamma(\sigma_1^2 - \sigma_2^2)}, \tag{24}$$

which reduces to

$$\mu_2 > \frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2}\mu_1. \tag{25}$$

Since $\sigma_1^2 > \sigma_2^2$, we have $\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2} < 1$, and there exist pairs (μ_1, σ_1^2) and (μ_2, σ_2^2) for which the inequality holds. Moreover, if w_1^* is a valid local optimum, it is also the global optimum, since a local optimum $\tilde{w}$ of $Q_2(w)$ can only be valid if $Q_2(\tilde{w}) < Q_1(\tilde{w})$.

An analogous argument leads to the condition for w_2^* to be a valid optimum. Solving the inequality $w_2^* < w_{12}$ gives

$$\mu_2 < \frac{2\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\mu_1. \tag{26}$$

Again, since we have assumed $\sigma_1^2 > \sigma_2^2$, this optimum can occur for sufficiently small μ_2. Inequalities (25) and (26) can not hold simultaneously since

$$\frac{2\sigma_2^2}{\sigma_1^2 + \sigma_2^2} < \frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2}. \tag{27}$$

Therefore there also exists an interval

$$\frac{2\sigma_2^2}{\sigma_1^2+\sigma_2^2}\mu_1 < \mu_2 \leq \frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2}\mu_1, \tag{28}$$

for which the robust optimal portfolio is the corner solution w_{12}. □

B Factor model prior

Following Pàstor and Stambaugh (2000), we consider experts with a prior on $\alpha,\ B,\ \nu,\ D$ and Ψ:

$$\begin{aligned} D^{-1} &\sim W(df, H^{-1}) \\ \alpha \mid D &\sim N(0, \frac{\sigma_\alpha^2}{s^2}D) \\ p(B) &\propto 1 \\ p(\nu) &\propto 1 \\ p(\Psi) &\propto |\Psi|^{-(k+1)/2} \end{aligned} \tag{29}$$

with $H = s^2(df - N - 1)I_N$ and, following an empirical Bayes approach, the value of s^2 is set equal to the average of the diagonal elements of the sample estimate of D. $W(df, S)$ denotes the Wishart distribution with df degrees of freedom and parameter S.

Observe that the prior for D approximately contains the equivalent information of df observations and the expectation of D equals $s^2 I_N$. The strength of the expert's believe in the factor model is quantified by two coefficients: σ_α and df. A dogmatic follower of the factor model will set $\sigma_\alpha = 0$. Someone with serious doubts about the factor model will set σ_α large and df small.

To derive the posterior estimators, define the data matrices $Y = (y_1', \ldots, y_T')$, $X = (x_1', \ldots x_t')$ and $Z = (\iota_T\ X)$ with dimensions $(T \times N)$, $(T \times k)$ and $(T \times (k+1))$ respectively and the $((k+1)\times N)$ matrix $A = [\alpha\ B]'$ and denote $a = vec(A)$. Also define the statistics $\hat{A} = (Z'Z)^{-1}Z'Y$, $\hat{a} = vec(\hat{A})$, $\hat{D} = (Y - Z\hat{A})'(Y - Z\hat{A})/T$, $\hat{\nu} = X'\iota_T/T$ and $\hat{\Psi} = (X - \iota_T\hat{\nu}')'(X - \iota_T\hat{\nu}')/T$.

Pàstor and Stambaugh (2000) report the posterior

$$\begin{aligned} D^{-1} \mid X,\ Y &\sim W(T + df - k, (H + T\hat{D} + \hat{A}'Q\hat{A})^{-1}) \\ \alpha \mid D,\ X,\ Y &\sim N(I_N \otimes F^{-1}Z'Z)\hat{\alpha}, D \otimes F^{-1}) \end{aligned} \tag{30}$$

with $F = diag(s^2/\sigma_\alpha^2,\ 0_k') + Z'Z$ where 0_k presents a vector with k zeros and $Q = Z'(I_T - ZF^{-1}Z')Z$ and

$$\begin{aligned} \Psi^{-1} \mid X,\ Y &\sim W(T-1, \hat{\Psi}^{-1}/T) \\ \nu \mid \Psi,\ X,\ Y &\sim N(\hat{\nu}, \Psi/T) \end{aligned} \tag{31}$$

As B and ν are independent in the posterior, the posterior mean of the expected return vector is

$$\tilde{\mu} = \mathsf{E}[\mu|\ X, Y] = \tilde{\alpha} + \tilde{B}\tilde{\nu}. \tag{32}$$

where the tilde refers to the posterior expectation given the data. An analytical expression for the posterior variance

$$\tilde{\Sigma} = \mathsf{E}[\Sigma|\ X, Y] = \mathsf{E}[B\Psi B' + D|\ X, Y] = \mathsf{E}[B\Psi B'|\ X, Y] + \tilde{D} \tag{33}$$

is not available. We could (approximately) compute the posterior variance by Monte-Carlo simulation. We will use $\tilde{\Sigma} = \tilde{B}\tilde{\Psi}\tilde{B}' + \tilde{D}$ in our computations. This actually underestimates the posterior covariance matrix, but as the (co)variances of the elements of B (see (30)) are typically small, the underestimation is limited (about 1% in our empirical study).

The strength of the investor's belief in her favorite factor models is expressed by σ_α and df. We use $\sigma_\alpha = 0.1\%$ (annually for $T = 120$ observations) and $df = 81$ reflecting a strong prior on the expected return vector.

Consider the factor model (9). If α, B, Σ and Ψ are known such that $w^*_{FM} = \frac{1}{\gamma}\Sigma^{-1}B\tilde{\nu}$, we can simplify the loss function

$$\begin{aligned}
2\gamma \mathsf{E}[L(\mu, \Sigma|\bar{y})] &= \mathsf{E}[(\mu - \alpha - B\tilde{\nu})'\Sigma^{-1}(\mu - \alpha - B\tilde{\nu})] \\
&= \mathsf{E}[(\mu - \alpha - B\nu - B\bar{e})'\Sigma^{-1}(\mu - \alpha - B\nu - B\bar{e})] \\
&= \big((\mu - \alpha - B\nu)'\Sigma^{-1}(\mu - \alpha - B\nu) + \mathsf{E}[\bar{e}'B'\Sigma^{-1}B\bar{e}]\big) \\
&= \big((\mu - B\nu)'\Sigma^{-1}(\mu - B\nu) + \mathsf{E}[\mathsf{tr}(B'\Sigma^{-1}B\bar{e}\bar{e}')]\big) \\
&= \big((\mu - B\nu)'\Sigma^{-1}(\mu - B\nu) + \mathsf{tr}(B'\Sigma^{-1}B\Psi/T)\big) \\
&= (\mu - B\nu)'\Sigma^{-1}(\mu - B\nu) + \frac{1}{T}\mathsf{tr}(I + D(B\Psi B')^{-1})^{-1}
\end{aligned}$$

B.1 Estimation uncertainty

Consider an expert who adopts a non-informative prior on the expected returns and suppose Σ is known. The uncertainty of the expected excess return vector μ conditional on the observation $\bar{y}$, may be described by the uncertainty matrix $\Omega = \Sigma/T$ and

$$(\mu - \bar{y})'\Omega^{-1}(\mu - \bar{y}) \sim \chi^2(n).$$

Consequently the uncertainty set

$$\mathcal{U} = \{\mu : (\mu - \bar{y})'\Omega^{-1}(\mu - \bar{y}) \leq \theta^2\},$$

will contain μ with sufficient probability (90%) if $\theta^2 = \chi^2_{inv}(N, 0.90)$.

Alternatively, consider a factor model (9)-(10). From (30) and (31) it follows

$$\begin{aligned}
\Omega_a &= Var(a|X,\ Y) = \tilde{D} \otimes F^{-1} \\
\Omega_\nu &= Var(\nu|X,\ Y) = \hat{\Psi}/(T - K - 2)
\end{aligned}$$

and B and ν are independent in the posterior. Following a similar argumentation as for (B.1) we define the uncertainty sets for a and ν,

$$\begin{aligned}
\mathcal{U}_a &= \{(a - \tilde{a})'\Omega_a^{-1}(a - \tilde{a}) \leq \theta^2_\beta\} \\
\mathcal{U}_\nu &= \{(\nu - \tilde{\nu})'\Omega_\nu^{-1}(\nu - \tilde{\nu}) \leq \theta^2_\nu\}.
\end{aligned}$$

We choose θ_β and θ_ν such that the uncertainty sets cover 90% posterior probability[6].

Given a portfolio w, the innermost minimization in (16) reduces to

$$\begin{aligned}
&\min_{\alpha, B,\ \nu} w'(\alpha + B\nu) - \frac{1}{2}\gamma w'(B\tilde{\Psi}B' + \tilde{D})w \\
&vec([\alpha, B]') \in \mathcal{U}_a \\
&\nu \in \mathcal{U}_\nu
\end{aligned} \tag{34}$$

[6]This Carthesian product of uncertainty sets is not exactly the highest posterior density region associated with 0.90^2 cumulative posterior density but convenient for computations.

Unfortunately the uncertain parameters enter both in a bi-linear and quadratic term of the objective (34) and we are not able to solve this bi-linear optimization problem.

Only if we abstract from uncertainty in one of the two estimators and ignore the effect of B on the variance term, we can solve the problem exactly. As an alternative we consider the uncertainty in the expected return directly and apply the first-order Taylor expansion for standard errors on μ,

$$\begin{aligned}\tilde{\mu} - \mu &\approx (\tilde{\alpha} - \alpha) + (\tilde{B} - B)\tilde{\nu} + \tilde{B}(\nu - \tilde{\nu}) \\ &= \tilde{B}(\nu - \tilde{\nu}) + I_N \otimes \begin{pmatrix} 1 \\ \tilde{\nu} \end{pmatrix}' (a - \tilde{a})\end{aligned}$$

and consequently

$$\Omega = \mathsf{E}[(\mu - \tilde{\mu})(\mu - \tilde{\mu})'] = \frac{1}{T}\tilde{B}\tilde{\Psi}\tilde{B}' + \begin{pmatrix} 1 \\ \tilde{\nu} \end{pmatrix}' F^{-1} \begin{pmatrix} 1 \\ \tilde{\nu} \end{pmatrix} \tilde{D}$$

We derive an uncertainty set of the form

$$\mathcal{U} = \{\mu : (\mu - \tilde{B}\tilde{\nu})'\Omega^{-1}(\mu - \tilde{B}\tilde{\nu}) \leq \theta^2\}, \tag{35}$$

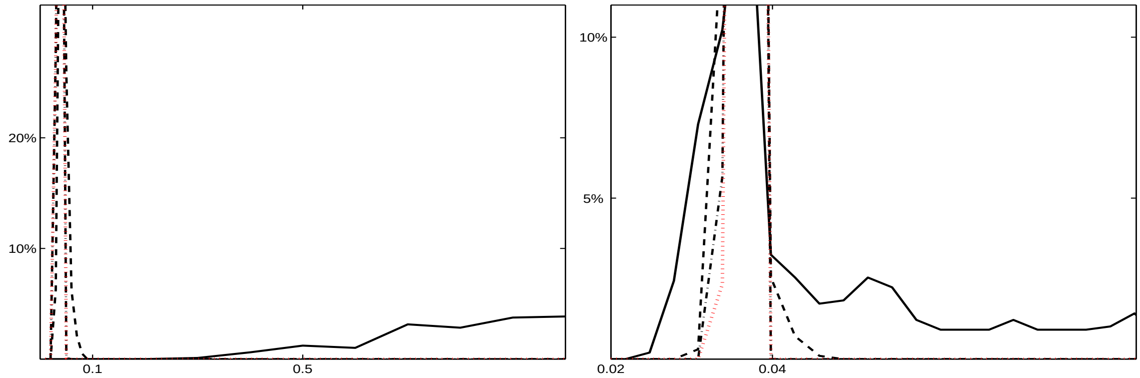
and we use[7] $\theta^2 = F_{inv}(K, T, 0.90)$ if $\sigma_\alpha \approx 0$ and $\theta^2 = \chi^2_{inv}(N, T, 0.90)$ otherwise to have a reasonable estimate of θ which leads to an uncertainty set with 90% posterior density, though not necessarily the highest posterior density region.

An investor may also be concerned with the uncertainty in the covariance matrix. An exact and workable characterization of this uncertainty is not available. Nevertheless Goldfarb and Iyengar (2003) provide an approximation of this set. We use this approximation in our computations.

[7]Ideally θ corresponds to the 90th percentile of $(\mu - \tilde{B}\tilde{\nu})'\Omega^{-1}(\mu - \tilde{B}\tilde{\nu})$ which we are unable to compute directly. The reported values are reasonable approximates to the desired percentile.

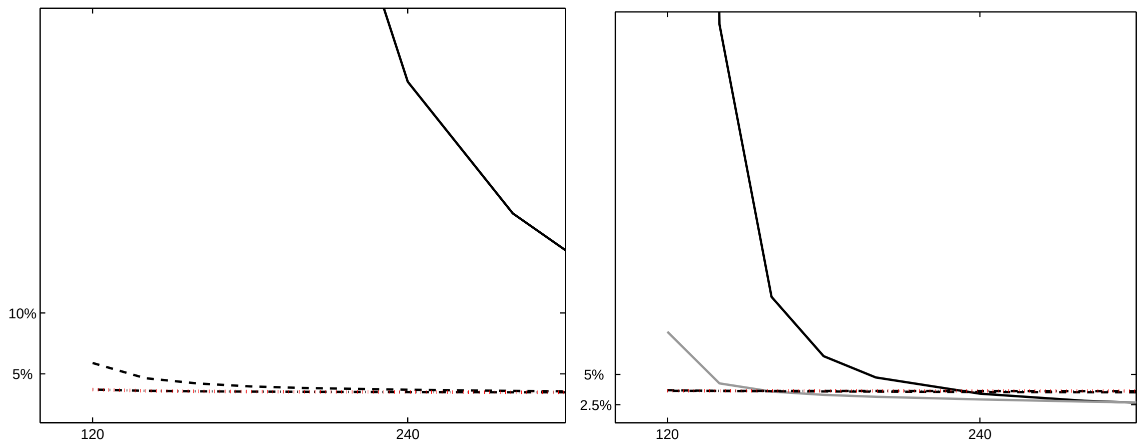
C Figures

Figure 3: Expected loss distribution.



Notes: The figure shows the ex-post expected loss distribution over 10.000 bootstrap samples of 120 random observations from the Fama & French 25 portfolios dataset over the period January 1975 to December 2001. The left panel concerns the unstructured portfolio. The right panel considers structured portfolios: CAPM prior (dash-dot), the IFF prior (dashed) and the model robust portfolio (thick dashed).

Figure 4: Expected loss as a function of uncertainty



Notes: The figure shows the portfolio performance in terms of expected loss (left axis, logarithmic scale) as a function of the number of observations, when the investor ignores estimation uncertainty (left-panel) and when the investor is robust to estimation uncertainty in the mean (right-panel). The different lines correspond to alternative model priors: unstructured model (solid), CAPM prior (dash-dot), the IFF prior (dashed) and the model robust portfolio (thick dashed). The grey line in the right panel corresponds to an unstructured model and a robust approach to estimation uncertainty in the expected return and the covariance matrix.

References

Avramov, D. 2002. Stock return predictability and model uncertainty. *Journal of Financial Economics* **64**(3), pp. 423–458.

Chopra, V. and Ziemba, W. 1993. The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management* **19**(2), pp. 6–12.

Fama, E. and French, K. 1998. Value versus growth: The international evidence. *Journal of Finance* **53**(6), 1975–99.

Goldfarb, D. and Iyengar, G. 2003. Robust portfolio selection problems. *Mathematics of Operations Research* **28**(1), pp. 1–38.

Jobson, J. and Korkie, B. 1981. Putting Markowitz theory to work. *Journal of Portfolio Management* **7**(4), pp. 70–74.

Jorion, P. 1986. Bayes–Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis* **21**(3), pp. 279–292.

Lintner, J. 1965. The valuation of risk assets and selection of risky investment in stock portfolios and capital budgets. *Review of Economics and Statistics* **47**, pp. 13–37.

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

Michaud, R. 1989. The Markowitz optimization enigma: is 'optimized' optimal?. *Financial Analyst Journal* **45**(1), pp. 31–42.

Michaud, R.: 1998. *Efficient asset management: a practical guide to stock portfolio optimization and asset allocation*. Harvard Business School Press.

Mossin, J. 1966. Equilibrium in capital asset market. *Econometrica* **34**, pp. 768–783.

Pàstor, L. and Stambaugh, F. 2000. Comparing asset pricing models: an investment perspective. *Journal of Financial Economics* **56**, pp. 335–381.

Rustem, B., Becker, R. and Marty, W. 2000. Robust min-max portfolio strategies for rival forecast and risk scenarios. *Journal of Economic Dynamics and Control* **24**, pp. 1591–1621.

Sharpe, W. 1964. Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance* **19**, pp. 425–442.

Stulz, R. 1995. The cost of capital in internationally integrated markets: the case of nestl. *European Financial Management* **1**, 11–12.